GENETIC DIVERSITY AND GENOTYPE × ENVIRONMENT INTERACTION OF ADVANCED ELITE SOYBEAN GENOTYPES IN UGANDA

BY

CLEVER MUKUZE

BSc. (Hons) Crop Science (University of Zimbabwe)

REG. NO: 2017/HD02/2474X

A RESEARCH THESIS SUBMITTED TO THE DIRECTORATE OF RESEARCH AND GRADUATE TRAINING FOR PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE AWARD OF MASTER OF SCIENCE IN PLANT BREEDING AND SEED SYSTEMS OF MAKERERE UNIVERSITY

DECEMBER, 2019

DECLARATION

I, Clever Mukuze, declare that the work presented in this thesis is my own research and has not been submitted for the award of degree in any other university.

Signed……. Date…….

This thesis has been submitted for being examined with our approval as the university supervisors.

Signed……..... Date……..
Prof. Phinehas Tukamuhabwa
School of Agricultural Sciences
Department of Agricultural production
Makerere University

Signed…… Date……
Prof. Patrick Rubaihayo
School of Agricultural Sciences
Department of Agricultural production
Makerere University

DEDICATION

I dedicate this work to my parents Mr. D. A. Mukuze and Mrs. E. Nemhari-Mukuze my brother Thomas, my sisters Juliana, Elita, Laiza, Plaxedes and Emily.

ACKNOWLEDGEMENTS

I am so grateful to The Almighty God for thus far He has brought me, for His protection, provision, favour and guidance. I am grateful to my supervisors Prof. Patrick Rubaihayo and Prof. Phinehas Tukamuhabwa for their dedication, continued support and guidance throughout my research project and thesis write-up.

I am very thankful to Intra-ACP mobility/Inter-University Cooperation to Train Crop Scientists for African Agriculture project funded by the European Union for granting me this study scholarship. I am grateful to RUFORUM for partial funding of my research project while the execution and interpretation is wholly my responsibility. I also extend my appreciation to the Makerere University College of Agricultural and Environmental Sciences (CAES) staff members for the technical knowledge and skills they have offered me. I also thank the Makerere University Soybean and Breeding Programme team and Intra-ACP mobility Administrator, Prof. Phinehas Tukamuhabwa, Mr. Tonny Obua, Mr. Senkubuge Yafesi, Mr. Kabayi Paul, Mr. George Yiga, Mercy Namara and Ruthie Najjuma Mutyaba for their support during my research.

Sincere gratitude to Dr. Mcebisi Maphosa, Dr. Hellen Kongai, Prof. Paul Gibson and Dr. Isaac Dramadiri. Thank you so much for your effort and technical advice to help shape my thesis amidst your busy schedules. Thank you for being mentors and role models, may the good Lord bless you exceedingly.

Special thanks to MUARIK Biotechnology Laboratory and NaCRRI Biosciences Laboratory administration for allowing me to carry out my research at their Institutes and to all the technicians that helped me during the laboratory work, thank you so much for the hard work and support that you offered me. Heartfelt appreciation to my family for the continued prayers, support and encouragement anytime I needed you. And to my dear friends and fellow graduate students, I am forever grateful for both the academic and moral support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

ANOVA…….Analysis of Variance

AMMI……Additive Main effect plus Multiplicative Interaction

CTAB…….Cytel trimethyl ammonium bromide

DARWin…….Dissimilarity Analysis Representation for windows

GenAIEx…….Genetic Analysis in Excel

GGE……Genotype main effect plus Genotype × Environment interaction

G × E……..Genotype × Environment Interaction

IPCA……Interaction Principal Component Analysis

MUARIK…….Makerere University Agricultural Research Institute, Kabanyolo

NaCRRI……..National Crops Resources Research Institute

NARO……..National Agricultural Research Organization

PCA……..Principal Component Analysis

PCoA……Principle Coordinate Analysis

PIC……...Polymorphic Information Content

PCR……..Polymerase Chain Reaction

RCBD……..Randomized Complete Block Design

RUFORUM…….Regional Universities for Capacity Building in Agriculture

SSR……..Simple Sequence Repeats

SVD…….Singular Value Decomposition

SVP……..Singular Value Partitioning

TVC……..Technology Verification Centre

UPGMA……Unweighted Pair-Group Method Arithmetic Average

ZARDI…….Zonal Agricultural Research and Development Institute

LIST OF APPENDICES

ABSTRACT

Soybean is an important food and cash crop in Uganda. Despite its importance, soybean production is highly affected by ever-changing biotic and abiotic stresses. The main objective of this study was to enhance soybean production and productivity through the use of adapted and diverse soybean germplasm in Uganda. The first study was specifically carried out to determine the degree of genetic diversity that existed among advanced elite soybean genotypes and released varieties. The study was carried out at Makerere University Agricultural Research Institute Kabanyolo (MUARIK) screen house and Biotechnology Laboratory. Thirty-four genotypes used in this study comprised of 23 advanced elite soybean genotypes and 11 released varieties. Thirty-one SSR markers were screened and only 21 showed polymorphism and were used for the study. A total of 59 alleles with an average of 2.85 alleles per locus were detected among the genotypes. The number of alleles varied from 2 (Satt126, Sat_409, Satt717, BE806308, Satt185, Satt264 and Sat_084) to 4 (Satt216, Satt431 and Satt411). The PIC values ranged from 0.208 on BE806308 to 0.741 on Satt411 with an average of 0.5870. Observed heterozygosity varied from 0.000 to 0.088 with an average of 0.010, while the expected heterozygosity ranged from 0.208 on BE806308 to 0.725 on Satt411, with an average of 0.548. Cluster analysis was performed and all the 34 genotypes were grouped into three major clusters, CI, CII and CIII with the results indicating that almost all released varieties belonged to cluster I. The principle coordinate analysis explained 28.9% of the total variation in which the PCo1 and PCo2 accounted for 17.2% and 11.7% of the total variation, respectively in discriminating between the released varieties and elite soybean genotypes. All the 21 markers used successfully distinguished advanced elite soybean genotypes and released varieties and showed the existence of moderate genetic variation among genotypes. The presence of moderate genetic variation is not much favorable in soybean breeding since there are dynamic pests, diseases and abiotic stresses which can accelerate vulnerability of closely related cultivars to outbreak of pests, diseases and abiotic stresses. Therefore, inclusion of more diverse germplasm and use of landraces in the soybean breeding programmme may provide the level of genetic variation necessary to cope with dynamic pests, diseases and abiotic stresses of Uganda.

The second study was specifically carried out to evaluate the yield, protein and oil content stability of advanced elite soybean lines across diverse environments. The experiment was

conducted at six locations namely; On-station (Kabanyolo), TVC (Iki-Iki and Nakabango) ZARDIs (Ngetta, Abi and Bulindi). Twenty-three advanced elite soybean genotypes and two commercial varieties used as checks were evaluated in a randomized complete block design (RCBD) replicated three times. Combined analysis of variance over locations and seasons was carried out for grain yield, protein and oil (%) content. The results for grain yield showed significant ($p<0.05$) differences for all the sources of variation except genotypes × season, interaction which was non-significant. The results for protein and oil content (%) showed non-significance ($p>0.05$) for all the sources of variation except location which was significant ($p<0.05$). The AMMI and GGE stability models were carried out for grain yield to determine wide and specific adapted genotypes. AMMI and GGE biplots explained 68% and 65.74% of the total interaction sum of squares, respectively and identified winning genotypes in each mega-environment as follows; BSPS 48A-28 for mega-environment I (Bulindi, Nakabango and Kabanyolo), BSPS 48A-28-1 for mega-environment II (Iki-Iki), Bulindi 18.4B for mega-environment III (Ngetta) and BSPS 48A-24-1 for mega-environment IV (Abi). Based on mean yield and stability performances, Maksoy 3N was the most stable genotype while BSPS 48A-28 was the highest yielding genotype with moderate stability. In conclusion, the study observed moderate genetic variation among the soybean genotypes evaluated. BSPS 48A-28 outperformed the released varieties (Maksoy 3N and Maksoy 4N) that were included in the evaluation trial. Superior and widely adapted genotypes like BSPS 48A-28 should be submitted for tests of novelty, distinctiveness, uniformity and stability (DUS) by the Ministry of Agriculture Animal Industries Department of the National Seed Certification Service. No need to test for protein and oil content on multi-location, since there was non-significant G × E interaction for those traits.

CHAPTER ONE

1.0 INTRODUCTION

1.1 Importance of soybean

Soybean [*Glycine max* (L.) Merr.] is a grain legume crop of great potential in Sub-Saharan Africa. It is an excellent source of protein and oil (approximately 40% protein and 20% oil content) for human food and animal feed (Ibanda *et al*., 2018). In Uganda and some parts of Sub-Saharan Africa, soybean is increasingly becoming a popular food and cash crop (Tukamuhabwa *et al*., 2016). This is evidenced by the rapid increase in number of industries involved in processing soybean in Uganda and within the region for food in the last decade (Tukamuhabwa and Obua, 2015).

Due to its nutritional superiority, soybean flour is often mixed with cereal flours mainly sorghum and maize to boost their nutritional value (Tukamuhabwa and Oloka, 2016). Soybean can substitute highly expensive source of proteins such as animal and fish protein because the amino acid profile of proteins found in soybean is similar to animal proteins (Merritt and Jenks, 2004). For this reason, soybean foods are highly recommended by nutritionists to nursing mothers, children and HIV patients (Tukamuhabwa and Obua, 2015).

Soybean oil is 85% unsaturated, contains linolenic acid (omega-3 fatty acid) and oleic acid, which have been shown to reduce the risk of heart disease by lowering serum cholesterol by 33% (Sacks *et al*., 2006). Besides that, soybean also contains flavones which increase artery and heart health (Clemente and Cahoon, 2009). Recent studies have shown that regular intake of soy food can reduce the risk of rectal cancer by 80%, mammary tumor by 40% and breast cancer by 50% (Tukamuhabwa and Obua, 2015). Soybean also improves soil fertility by adding nitrogen from the atmosphere hence help in replenishing soil fertility (Agoyi *et al*., 2017).

1.2 Status of soybean production in Uganda

Despite the aforementioned importance of soybean in Uganda, its production is very low in some parts, for example, the Northern region is the highest producer of soybeans with best productivity mean dry grain yield of 1804 kg/ha while West Nile reported the least yield of 247.1 kg/ha

(Tukamuhabwa *et al*., 2016). Farmers in Eastern, Western and Central region have mean dry grain yield of 1561.7kg/ha, 696.8kg/ha and 640 kg/ha, respectively (Tukamuhabwa *et al*., 2016). Consequently, the national Soybean Breeding Programme has been actively involved in developing varieties to meet the ever-changing needs of farmers and processors in the diverse environments of the country (Tukamuhabwa *et al*., 2012). In chronology of their release, the varieties which are already on the market are Maksoy 1N and Namsoy 4M (2004), Maksoy 2N (2008), Maksoy 3N (2010), Maksoy 4N and Maksoy 5N (2013) and Maksoy 6N (2017). These improved varieties have been adopted for commercial production in Uganda, and have led to soybean grain yield of between 2000–3000 kg/ha, providing income and an affordable source of protein for the country's rural population. However, new varieties are continuously being produced and, therefore, it is important for farmers to be kept updated on the new ones which have an advantage over the old varieties (Tukamuhabwa and Obua, 2015). The development of improved new varieties will be the only sustainable way to cope with the ever-changing biotic stresses like the outbreak of groundnut leafminer (Ibanda *et al*., 2018) and soybean rust (Maphosa *et al*., 2013; Gebremedhn *et al*., 2018) and abiotic stresses such as like extreme temperature and rainfall changes (Tukamuhabwa *et al*., 2016).

1.3 Production constraints in soybean

Tukamuhabwa *et al*. (2016) reported that diseases contributed a 9.1% reduction in soybean yield in Uganda. Major diseases which reduce soybean yield are soybean rust *Phakopsora pachyrhizi* (Maphosa *et al*., 2013; Gebremedhn *et al*., 2018) bacterial pustule (*Xanthomonas campestris* pv. *glycines*), bacterial blight (*Pseudomonas amygdali* pv. *glycinea*), frog-eye leaf spot (*Cercospora sojina),* red leaf blotch (*Phoma glycinicola*) and soybean mosaic virus disease (Tukamuhabwa and Obua, 2015). It is common to see overlapping disease infections in a soybean crop on many farmers' fields (Tukamuhabwa *et al*., 2016).

Namara (2015) and Tukamuhabwa *et al*. (2016) reported that soybean insect pests which contributed 22.6% yield reduction included pod feeders, affecting the harvestable product directly, and stem and root feeders. The foliage feeders include: grasshoppers (*Melanoplus spp*.), spotted cucumber beetle (*Diabrotica undecimpunctata* Howardi), striped blister beetle (*Epicauta vittata*), Japanese beetle (*Popillia japonica*), Mexican bean beetle (*Epilachna varivestis*), bean leaf beetle (*Cerotoma trifurcata*), green clover worm (*Plathypena scabra*), soybean looper

(*Pseudoplusia includens*), velvet bean caterpillar (*Anticarsia gemmatalis*), beet armyworm (*Spodoptera exigua*) and soybean aphid (*Aphis glycines*); pod feeders include corn earworm *(Heliothis zea*), green stink bug (*Nezara viridula*) and brown stink bug (*Euschistus servus*). Stem feeders include soybean stem borer (*Dectes texanus texanus*), lesser cornstalk borer (*Elasmopalpus lignosellus*), cut worms (*Agrotis spp*), grape colapsis (*Colaspis brunnea*) and white grubs (*Phyllophaga* and *Cyclocephala spp*) that feed on roots. The groundnut leafminer (*Aproaerema modicella*) (GLM) is principally a key pest of groundnuts but has recently become a major problem on soybean and is reported to be spreading widely and reaching out break densities, causing great losses on the crop in Uganda (Namara, 2015; Ibanda *et al*., 2018).

Low soybean yield in Uganda also is a result of abiotic stresses. Extreme changes in temperature and rainfall patterns, and low soil fertility contributed 26.1% and 5.2% of the soybean yield reduction, respectively (Tukamuhabwa *et al*., 2016). Therefore, widely adapted soybean varieties with stable yields and nutritional quality are necessary to sustain soybean production (Tukamuhabwa *et al*., 2012).

1.4 Genetic variability in soybean

Intensive breeding using closely related progenitors has been narrowing the genetic base of soybean (Bisen *et al*., 2015). Repeated use of closely related lines has narrowed the genetic base of soybean germplasm, consequently increasing its vulnerability to changes in biotic and abiotic stresses (Ssendege *et al*., 2015). To mention a few, (Namara, 2015; Ibanda *et al*., 2018) highlighted that; almost all released varieties in Uganda are now susceptible to groundnut leafminer, mainly due to narrow genetic base. Therefore, assessment of genetic diversity is very crucial for efficient selection of parental lines for designing future breeding efforts to improve soybean yield, quality, pest and disease management (Wang *et al*., 2006). Soybean, being a strictly self-pollinated crop, its genetic base is considered to be extremely narrow because of limited outcrossing (Hipparagi *et al*., 2017). To broaden the genetic base of soybean, understanding the genetic diversity is essential (Kumawat *et al*., 2015).

Generations of newly improved soybean cultivars can be enhanced by new sources of genetic variation, not only for the improvement of agronomic value, but also for genetic diversity (Bisen *et al*., 2015). Therefore genetic variation is very crucial for plant breeders to cope with ever-

changing biotic and abiotic stresses and also for variety protection (Ssendege *et al*., 2015). Studies have reported that only few accessions have contributed many alleles in current cultivars, leading to narrowing genetic diversity in soybean varieties, which is a major constraint to the improvement of soybean (Fu *et al*., 2007). Uganda is no exception to this; however, the introduction of exotic diverse germplasm into the breeding programme would improve genetic variability of local germplasm including alleles for high yielding, resistant to groundnut leaf miner, bruchids, rust and drought (Ssendege *et al*., 2015).

Knowledge of genetic diversity among accessions is crucial for designing proper future breeding efforts for improvement of yield, quality and diseases and pest resistance (Tantasawat *et al*., 2011). Several methods have been developed for assessment of genetic diversity among accessions such as the differences in morphological and agronomic traits, isozymes, pedigree information and DNA markers (Chakraborty *et al*., 2018). However, methods like morphological and agronomic traits assessment are highly affected by environmental factors (Gupta and Manjaya, 2017; Chauhan *et al*., 2015; Ghosh *et al*., 2014; Chakraborty *et al*., 2018). The use of pedigree information is also affected by uncertain or incomplete data and possible errors in data capture (Chakraborty *et al*., 2018; Oda *et al*., 2015). The limitation of data provided by the use of isozymes permits the use of DNA markers in genetic diversity studies (Chauhan *et al*., 2015). Among the different DNA markers are; amplified fragment length polymorphism (AFLPs), single nucleotide polymorphism (SNPs), restriction fragment length polymorphism (RFLPs), microsatellites or simple sequence repeats (SSRs) and random amplified fragment polymorphic DNAs (RAPDs) have been widely used in studying genetic diversity in soybeans, each with its own merits and demerits (Chakraborty *et al*., 2018). Among the DNA markers used for genetic diversity study, SSRs have been shown to be highly polymorphic as compared to RAPDs, RFLPs and AFLPs. Furthermore, SSRs have much greater capacity to effectively identify unique alleles among soybean germplasm than other marker systems (Tantasawat *et al*., 2011).

1.5 Genotype × Environment interactions

Uganda's climatic conditions are highly variable with mean annual rainfall of 510- 2160 mm, temperature 23- 28^{o}C and varied soil influenced by soil depth, texture, acidity and organic matter (Tukamuhabwa *et al*., 2012). Due to the variability in abiotic and biotic factors from location to

location, the crop remains exposed to the influence of huge genotype and environment interactions, leading to inconsistent genotypic responses (Bhartiya *et al*., 2017). The presence of a significant genotype × environment interaction (G × E interaction) for quantitative traits such as seed yield, protein and oil content can lead to the failure of genotypes to achieve the same relative performance in different environments (Gurmu *et al*., 2009).

The differential response of genotypes across several unrelated environments reduces the correlation between phenotype and genotype and thus reduces responses to selection and subsequently progress in plant breeding programme (Crossa *et al*., 2002; Yan and Kang, 2002). Furthermore, the presence G × E interaction needs new varieties to be tested in multi-environment trials (METs) for identification of multi-environmentally adapted genotypes (Bernado, 2002; Yan and Kang, 2002). The pattern of genotype response allows partitioning of test locations into mega environments and ideal environments based on their representative and discriminating ability (Yan *et al*., 2007). This is crucial in plant breeding in order to rationalise resources and confine genotype testing to locations with informative data facilitating a rapid response to selection. Yan and Tinker (2006) highlighted some objectives of MET analysis that included identification of mega-environments to reduce negative G × E interaction, identification of ideal testing locations within mega-environments and identification of superior genotypes.

Several options for dealing with significant G × E interaction have been reported that largely involve reducing or exploiting the interaction (Yan *et al.*, 2007; Yan and Tinker, 2006). Several statistical methods of analyzing G × E interaction have been reviewed (Westcott, 1986). However, not all ways of exploiting G × E interaction involve trying to reduce it (Bernado, 2002). Some methods, such as analysis of variance (ANOVA), are good at detecting G × E interaction but cannot reveal the pattern of the interactions (Gasura *et al*., 2015). Regression-based methods use environmental scores, which have less to do with genotype plus G × E interaction and explains only a small part of GGE (genotype main effect plus Genotype × environment interaction) (Yan *et al*., 2007). In the recent past, statistically effective methods, such as biplots based on Singular Value Decomposition (SVD) and Principal Component Analysis (PCA) have been developed for G × E interaction analysis (Gauch, 2006; Yan and Tinker, 2006).

Biplot analysis is a multivariate technique that graphically displays two-way data for visually interpreting the performance of genotypes in different environments (Gauch, 2013). The biplot is a scatter plot that approximates and graphical presentation of a genotype-environment two-way table. Approaches such as the additive main effect and multiplicative interaction biplot (AMMI) (Gauch, 2013) and the genotype main effect plus G × E interaction (GGE) biplot (Yan and Tinker, 2006) have been widely used to exploit significant G × E interaction in soybean METs data. Additive main effects and multiplicative interactions (AMMI) (Gauch, 2013) and genotype main effect plus G × E interaction (GGE) (Yan *et al*., 2007) statistical models effectively capture the additive (linear) and multiplicative (bilinear) components of G × E interaction and provide meaningful displaying and interpretation of multi-environment data set in breeding programmes.

The GGE biplot can be subjected to different ways of singular value partitioning (SVP) (Yan and Tinker, 2006). The biplot model that is fitted to residuals after the removal of the environmental main effect (environment-centered data) is called a GGE biplot (Yang *et al*., 2009). Yan and Hunt (2001) suggested that, for cultivar evaluation and recommendation, genotype and G × E interaction are the only two sources of variation that are crucial and must be considered simultaneously for appropriate genotype and test environment evaluation. Using a sites regression model (SREG), Yan *et al*. (2000) combined genotype main effect and genotype × environment interaction, denoted as G + G × E interaction or GGE and repartitioned this into crossover and non-crossover G × E interaction. In METs data, GGE biplot is useful in evaluating the genotype main effects plus the G × E interaction (Yan and Tinker, 2006). This approach has been commonly used for delineating soybean production mega-environments and soybean variety recommendations (Bhartiya *et al*., 2017).

AMMI, a multiplicative model that involves the product of a component that is due to genotype effect and the second component due to environmental effect in which the crop is grown (Bernado, 2002; Pinnschmidt and Hovmøller, 2002). The AMMI model specifically combines both the analysis of variance and PCA into a single model with additive and multiplicative parameters (Gauch, 2013). The model separates the variance due to genotype and environment (additive variance) from the interaction variance (G × E interaction variance) and applies PCA to the interaction portion from the ANOVA analysis (Bernado, 2002), to extract a new set of coordinate axes that account more effectively for the interaction patterns (Gauch, 2013). In

clarification of G × E interaction, AMMI summarizes patterns and relationships of genotypes and environments (Gauch, 2006). Furthermore, statistical model results from AMMI analysis are plotted in a graph showing the main and interaction effects for both genotypes and environments on the same scatter plot, with the noise rich residual discarded and the data separated into a pattern rich model to gain accuracy (Gauch, 2013).

1.6 Statement of the problem

Soybean production in Uganda is constrained by several biotic and abiotic stresses which cause significant yield losses and contribute to rampant malnutrition related disorders. Research carried out reported that none of the current released varieties is resistant to groundnut leafminers that attack soybean (Tukamuhabwa and Oloka, 2015). Ibanda *et al.* (2018) also reported that the epidemic of groundnut leafminer is sporadic and its severity varies from location to location. Furthermore, drought is undoubtedly a serious production constraint for smallholder farmers who grow soybean under rain-fed farming systems in most parts of the Eastern, Northern and West Nile regions (Namara, 2015). However, ongoing climatic changes and cyclic whether patterns such as the El-Nino phenomenon can greatly affect rain-fed agriculture, both positively and negatively regardless of traditional cropping seasons in Uganda. Low soil fertility has been reported for yield and nutritional quality loss in soybean (Tukamuhabwa and Oloka, 2015). Moreover, Obua, (2013); Maphosa *et al*. (2013); Tukamuhabwa *et al*. (2012); Gebremedhn *et al*. (2018) reported high yield loss in soybean due to rust. Under resource-limited farming systems, use of plant host resistance is the only sustainable approach to managing soybean rust. Yet, sustainable resistance is difficult to obtain at present, due to the high degree of genetic variability of the pathogen from location to location that causes resistance to break down in a short period after new resistant varieties have been released (Tukamuhabwa and Oloka, 2015). All these challenges call for assessment and mining of important alleles from diverse genotypes that have been developed by the Soybean Breeding Programme.

Since crop growing locations have no clearly defined demarcations and most farmers tend to influence each other in the choice of variety that is grown, the development of varieties with dynamic stability is strongly supported, rather than environment-specific varieties (Gasura *et al*., 2015; Bhartiya *et al*., 2017). Dynamic stability implies the crop does well in better environments rather that static stability were it gives average yield regardless the environment. Therefore, there

is need for advanced elite genotypes to be evaluated across the representative sites for informed variety recommendation to be done. In conclusion, widely adapted and diverse soybean varieties with stable yields and nutritional quality are necessary to sustain soybean production in Uganda (Tukamuhabwa *et al*., 2012).

1.7 Justification of the study

Knowledge on genetic diversity in available soybean genotypes could help the breeder to understand the structure of germplasm and predict which parental combinations would produce the best offsprings and facilitate to widen the genetic variation of breeding material for selection. Also, this study can generate knowledge of the pattern and magnitude of G × E interaction and stability of newly improved elite soybean genotypes and help the plant breeder in determining adapted genotypes in Uganda's variable climatic conditions. The identification of stable high yielding, protein and oil content lines would help the resource-limited farmers to cope with the ever-changing biotic and abiotic stresses, as well as improve their returns from selling high yielding soybeans with high protein and oil content. Also, the identification of stable high yielding, protein and oil content lines would help the plant breeder in future to incorporate those lines in the breeding programme. Accordingly, this study sought to evaluate advanced elite breeding lines in diverse areas of Uganda and evaluate their genetic, nutritional and agronomic attributes.

1.8 Main objective

To enhance soybean production and productivity through the use of adapted and diverse germplasm in Uganda.

1.8.1 Specific objectives

i. To determine the degree of genetic diversity of advanced elite soybean genotypes and released varieties.

ii. To evaluate the yield, protein and oil content stability of advanced elite soybean lines across diverse environments.

1.9 Hypotheses

i. There is narrow genetic variation among the released varieties and advanced elite soybean genotypes.

ii. There are stable high yielding advanced elite soybean genotypes with high protein and oil content.

CHAPTER TWO

2.0 LITERATURE RIVIEW

2.1 Genetic diversity in soybean germplasm

Simple sequence repeats have been widely used for genetic diversity studies among soybean accessions (Tantasawat *et al*., 2011). In recent genetic diversity and varietal identification study by Bisen *et al*. (2015) using sixteen polymorphic SSR markers, a total of 51 alleles with an average of 2.22 alleles per locus were detected. From that study, the polymorphic information content (PIC) varied from 0.049 (Sat_243 and Satt337) to 0.526 (Satt431) with an average of 0.199 among germplasm evaluated. Unweighted pairwise genetic similarity among genotypes was performed and varied from 0.56 to 0.97 with an average of 0.761. All the markers used were found polymorphic and successfully distinguish 12 out of 38 soybean genotypes evaluated.

Ssendege *et al*. (2015) evaluated genetic diversity of 92 soybean accessions using SSR markers. After two weeks, total genomic DNA was extracted from young leaves of seedlings using CTAB method as described by Deshmukh *et al*. (2007). Ten SSR primers pairs that were previously mapped on 10 linkage groups and highly informative (Song *et al*., 2013) were used to amplify the genomic DNA. The markers used were found polymorphic and successfully distinguished accessions with a total of 53 alleles detected. The PIC value was varied from 0.698 (Satt197) to 0.834 (Satt263). The dendrogram performed successfully clustered 92 soybean accessions into six distinct groups with a similarity coefficient level of 0.57.

Pagar *et al*. (2017) screened 52 SSR markers on 13 soybean genetic diversity study. From 52 markers evaluated, only 15 SSR primers (i.e. Satt184, Satt329, Satt335, BE806308, Satt077, Satt557, Satt194, Sat_409, Satt328, Satt406, Satt717, Satt329, Satt120, AW620 and Sct_189) showed polymorphic and successfully distinguished and discriminated all the soybean genotypes under study by generating unique allele for respective genotype. A total of 103 alleles detected with an average of 7.92 alleles per locus with a maximum of eleven bands (primer Sat_406) and minimum one polymorphic allele (Satt126). Unweighted pair wise coefficient of genetic similarity among all soybean genotypes ranged from 0.792 to 0.929.

Kumawat *et al.* (2015) carried a genetic diversity on 82 soybean accessions using 44 SSR markers. Only 40 markers were found polymorphic and successfully distinguished

accessions. The polymorphic information content among accessions was ranged from 0.101 (Satt484) to 0.742 (Satt396) with an average of 0.477. The pairwise genetic similarity varied from 0.28 to 0.90 among soybean accessions. The dendrogram performed based on Jaccard's similarity coefficient grouped 82 soybean accessions into three unique clusters.

Wang *et al*. (2010) assessed genetic diversity of 40 soybean accessions using 40 SSR markers and a total of 265 alleles with an average of 6.55 alleles per locus were detected among accessions. All the 40 markers were found polymorphic and successfully distinguished 40 accessions into seven distinct clusters based on similarity coefficient level of 0.755. Polymorphic information contents (PIC) ranged from 0.8500 (Satt197) to 0.5850 (Satt487), with an average of 0.7800. Out of 40 primers, eight markers recorded high number of alleles from loci Satt197, Satt281, Satt185, Satt373, Satt184 and Satt534; all had 8 alleles. Lower numbers of alleles were detected in loci Satt177, Satt226, Satt487, Satt022 and Satt514; all had 5 alleles and Satt236 had the lowest numbers of alleles. Shannon-Weaver's information indices ranged from 1.1528 to 2.0454 with an average of 1.6917, which corresponded to the primer pairs Satt574 and Satt487.

Ghosh *et al*. (2014) analyzed the genetic diversity among soybean 32 accessions using 15 SSR primer pairs. All of the 10 were found polymorphic and successfully distinguished soybean genotypes. A total of 8 alleles detected across 10 primer pairs. The polymorphic information content (PIC) among genotypes ranged from 0.21 (S26) to 0.83 (S27) with an average of 0.51. Pairwise coefficients of genetic similarity among genotypes varied from 0.76 to 1.00. Unweighted pair group method arithmetic average (UPGMA) cluster analysis grouped the genotypes into 2 major clusters and 6 sub-clusters.

Hu *et al*. (2018) evaluated the genetic diversity of 346 soybean accessions using 74 SSR primer pairs. A total of 924 alleles were detected among soybean accessions with an average of 12.49 per locus. Genetic diversity coefficient ranged from 0.02 (GMES3693) to 0.93 (att614), with an average of 0.71, and the polymorphic information content varied from 0.02 (GMES3693) to 0.93 (Satt614), with an average of 0.68. Principal coordinates analysis and neighbour-joining tree based on (Nei, 1972) genetic distance revealed that all the soybean accessions were grouped into two major clusters.

Mulato *et al*. (2010) assessed the genetic diversity of 79 soybean accessions using 30 markers (20 SSR and 10 EST-SSR primer pairs). All the markers were polymorphic and successfully distinguished soybean accessions. A total of 259 alleles were detected with an average of

8.63 number alleles per locus. The polymorphic information content values for the 30 used primers varied from 0.166 for marker PHYA1 to 0.921 for marker Sat_001, with an average of 0.626. Unweighted pair-group method arithmetic average (UPGMA) was performed using NTSYSpc software, version 2.2 (Rohlf, 2005) to generate the dendrogram which ultimately classified 79 soybean accessions into five clusters and several sub-clusters.

Guan *et al.* (2010) analyzed genetic relatedness among 244 soybean accessions using 46 SSR primer pairs and a total of 745 alleles were detected with per primer pair ranging from 7 (Satt230) to 34 by (Satt281) and an average of 16.2 numbers of alleles per locus. The values of each marker diversity varied from 0.61 to 0.92 in Chinese soybean and from 0.10 to 0.90 in Japanese soybean. The PIC values ranged from 0.64 to 0.79 among the accessions. From the observed 745 alleles, the allelic frequencies varied from 0.002 to 0.554 with an average of 0.06. All the markers used were found polymorphic and successfully distinguished and discriminated the 244 soybean accessions. The NTSYS-PC statistical package, version 2.1 (Exeter software, Setauket, NY) was used for Principal coordinate (PCO) analysis in order to show the distribution of the genotypes in scatter-plot based on their similarity matrix generated with Dice's method using the NTSYS-PC program (Rohlf., 2000) and the Nei's (1972) genetic distance matrices. The UPGMA dendrogram and the bootstrap test were performed using the DISPAN package (Ota, 1993) and grouped the 244 soybean accessions into four distinct clusters.

Chauhan *et al.* (2015) studied genetic diversity of 48 released varieties from different seed companies using 21 SSR primer pairs, with at least one primer pair from each linkage group. All the 21 primers were found polymorphic and successfully distinguished and discriminated all the varieties. A total of 84 alleles were detected among varieties with a frequency which ranged from 2 to 7 alleles per locus. Polymorphic information content ranged from 0.304 (Satt278) to 0.781 (SOYHSP176). Unweighted pair-group method arithmetic average (UPGMA) was carried out using NTSYS-PC statistical package, version 2.02e and the dendrogram constructed classified the soybean varieties into two major clusters with an average genetic similarity coefficient of 0.267.

Hua *et al*. (2005) conducted genetic diversity among 158 soybean accessions using 67 SSR primer pairs and a total of 460 alleles were detected with an average number of alleles per locus varying from 3 to 15 with an average of 6.9 across all accessions. The genetic similarity coefficient ranged from 0.101 to 0.672 with an average of 0.321 among all pairs of the 158

soybean accessions and UPGMA cluster analysis of the similarity classified all 158 soybean accessions into three distinct clusters.

Song *et al*. (2013) carried out genetic diversity study on 185 soybean accessions using 72 SSR markers. The total number of alleles detected from this study was 784 among accessions, with an average of 10.9 alleles per locus, ranging from three for eight markers (Sat_397, Sct_034, Sct_064, Satt227, Satt255, Satt328, Satt578 and Satt599) to 31 for Sat_210. The polymorphic information content ranged from 0.159 for marker Satt144 to 0.917 for Satt458 in among accessions. Unweighted pair-group method arithmetic average (UPGMA) cluster analysis was carried out using the software MEGA version 5.03 (Tamura *et al*., 2007) embedded in PowerMarker and the dendrogram constructed grouped the 185 soybean accessions into three distinct clusters containing several sub-clusters.

Shadakshari *et al*. (2011) studied genetic diversity on 50 soybean accessions using 11 SSR primer pairs and total of 33 alleles detected with an average number of alleles ranging from 2 to 4 per locus. All the markers used successfully distinguished and discriminated all the genotypes with a PIC varied from 0.280 (Satt448) to 0.843 (Satt207). The Unweighted pair-group method arithmetic average performed classified 50 accessions into 15 distinct clusters. Tantasawat *et al*. (2011) evaluated 25 soybean genotypes for their genetic diversity and relatedness with 12 SSR primer pairs and a total of 53 alleles were detected with an average of 4.82 alleles per locus. Only 11 markers were found polymorphic and successfully distinguished 23 of the 25 soybean genotypes. The PIC values varied from 0.13 (Satt285) to 0.88 (Satt173) with an average of 0.60. Cluster analysis results using UPGMA grouped all the genotypes into four distinct clusters. Further analysis was carried out using Principal coordinate analysis (PCoA) and confirmed the separation of genotypes into four groups.

2.2 Genotype × environment interactions for grain yield, protein and oil content in soybean.

Soybean has been reported to be sensitive to genotype × environment interactions (G× E interaction) by several researchers (Bhartiya *et al*., 2017; Krisnawati and Adie, 2018; Gurmu *et al*., 2009; Tukamuhabwa *et al*., 2012). Bhartiya *et al*. (2017) reported significant G × E interaction for soybean in a study on 36 soybean genotypes across 3 environments. The results from AMMI analysis revealed that environment, genotypes and genotype × environment interactions significantly affected grain yield with 9.76, 28.97 and 47.55% of the

total variation, respectively. Further analysis using GGE biplot was carried out to determine the most discriminating and representative test location and also to identify the winning genotypes across environments. The GGE biplot explained 74.40% of the interaction total sum of squares for grain yield in which 43.13% and 31.27% were accounted for by the first two principal components (PCI and PCII) respectively.

Tyagi *et al*. (2013) evaluated 40 soybean genotypes in eight environments. Combined analysis over locations revealed significant genotype, environments and G × E interaction for grain yield and protein content. Stability analysis results based on Eberhart and Russell (1966) regression model identified the most stable and high yielding genotypes. Liu *et al*. (2017) observed significant genotype, environments and G × E interaction for seed yield from 138 soybean genotypes evaluated in three locations for three seasons. Stability analysis using AMMI model (Guach, 2013) captured 70% of the total G × E interaction, and successfully discriminated genotypes, and identified the most stable and high yielding genotypes.

Popovic *et al*. (2013) reported significant year × genotype and genotype × environment interaction on grain yield, protein content and oil content on 10 soybean genotypes. Kumar *et al.* (2014) observed significant genotype, environments and G × E interaction for soybean grain yield evaluated on four locations. Stability analysis based on GGE model (Yan *et al*., 2007) effectively discriminated genotypes and identified the most stable and high yielding genotypes. Batista *et al*. (2015) reported significant genotype, environment and G × E interaction for grain yield, protein and oil content of 14 soybean genotypes evaluated on four locations. Based on regression analysis of Eberhart and Russell (1966), results showed genotypes that were most stable in terms of grain yield, protein and oil content. Njoroge *et al*. (2015) reported the presence of G × E interaction for grain yield and protein content on 15 soybean genotypes and the results of combined analysis of variance over locations revealed significant genotype, environment and G × E interaction effects for grain yield, protein and oil content. Atnaf *et al*. (2013) examined the nature and magnitude of G × E interaction on soybean for grain yield and to identify the winning genotypes and the results obtained from combined analysis of variance revealed that soybean grain yield was significantly ($p<0.001$) affected by genotypes (14.87%), G× E interaction (59.55%) and environments (25.58%). Stability analysis using GGE biplot was carried out to identify winning genotypes, most representative and discriminating location and results showed that, the first two principal components (PC1=41.6% and PC2=21.8%) of the GGE biplot explained only 63.4% of the GGE sum of squares using environment standardized model.

Tukamuhabwa *et al*. (2012) conducted a research to determine yield and stability of advanced generation soybean genotypes, the most ideal testing environment and to determine the presence of soybean production mega environments in Uganda. Results from combined analysis of variance over locations and seasons revealed significant interaction of genotypes and locations. The AMMI analysis showed the presence of significant scale G × E interaction for soybean grain yield. GGE biplot visual assessment identified the stable high yielding genotype and Nakabango as the most discriminating and representative test location. GGE biplot analysis explained 73.87% of total G × E interaction sum of squares with PCA1 (52.47%) and PCA2 (21.40%) respectively. From the environmental focusing plot, the five multi-locations tested were classified into two distinct mega environments for soybean production (Tukamuhabwa *et al*., 2012).

Gurmu *et al*. (2009) reported significant G × E interaction for grain yield, protein and oil content on study carried out on twenty soybean genotypes planted in six locations for one season. AMMI biplot was used to estimate the magnitude of G × E interaction, for grain yield in which the first two interaction principal component axes (IPCA 1 and IPCA 2) took the largest portion (66.15%) of the total sum of squares (36.36 and 29.79%, respectively). The IPCA 3 (17.52%) and IPCA 4 (13.02%) accounted for the remaining variation. The AMMI biplot for protein explained 43.13% and 25.0% (IPCA 1 and IPCA 2) of the interaction sum of squares respectively. AMMI model for oil content accounted (84.92%) of the interaction sum of squares with 34.45% and 25.7%, 24.77% for IPCA 1, IPCA 2 and IPCA 3 respectively.

Krisnawati and Adie (2018) reported the presence of G × E interaction for soybean grain yield. Twelve soybean genotypes were evaluated and the results of combined analysis of variance over locations showed significant G × E interaction for grain yield. Stability analysis using GGE biplot explained 75.96% in which PC1 and PC2 accounted for 57.41% and 18.55%, of the interaction sum of squares, respectively. Based on the GGE visual assessment, the soybean production environments were divided into three mega-environments in Indonesia. The which-won-where GGE scatter plot used effectively identified the wining genotypes per each location.

In recent studies on soybean multi-environment trials (METs) data conducted on five locations and three seasons in Zambia, Hampango *et al*. (2017) reported the presence of significant G × E interaction for protein and oil content. Results from combined analysis of

variance over locations showed significant G × E interactions for oil and protein content. Stability analysis using the AMMI model showed significant effects of environments, genotypes and G × E interaction which accounted for 49.41%, 27.15% and 23.44% of the total variation for oil content respectively. AMMI model explained 84.76% of the interaction sum of squares in which the IPCA1 and IPCA2 accounted 65.88% and 18.88% respectively. The AMMI biplot for protein content explained 81.11% of the interaction total sum of squares in which IPCA1 and IPCA2 accounted 62.17% and 11.94% respectively.

Chaudhary and Wu (2012) evaluated 15 soybean varieties for stability of their grain yield, protein content, and oil content at six different locations for one season. Mixed linear model based on the method of Moore and Dixon (2015) and AMMI model Gauch (2013) were applied to detect the presence G × E interactions and stability of each variety grain yield, protein and oil content. Based on AMMI analysis, results showed significant G × E interaction for grain yield, protein content, and oil content. In order to explain the magnitude of interaction of each genotype and environment, to identify varieties with specific environmental adaptation the AMMI2 biplot was applied which showed that biplot for grain yield explained 79.8% of the interaction total sum of squares in which PC1 and PC2 accounted for 50.4% and 29.4% of interaction respectively. For protein content, the AMMI2 biplot explained 68.7% which the PC1 and PC2 consisted of 39.5% and 29.2 % of interaction total sum of squares, respectively and for oil content accounted for 73.7% of the interaction in which PC 1 and PC2 comprised of 52.8% and 20.9% respectively. Nascimento *et al.* (2010) evaluated 15 soybean genotypes for three consecutive years on a total of 13 environments. Combined analysis of variance over years and locations revealed significant G × E interaction for protein and oil content. Based on Eberhart and Russell (1966) stability analysis model, results showed the most stable and ideal soybean genotypes in terms of protein and oil content.

In other related crops studied, Dolinassou e*t al.* (2017) reported significant G × E interaction for seed oil content on 12 peanut (*Arachis hypogaea* L.) varieties evaluated on 3 locations for two consecutive seasons. AMMI stability analysis based on Gauch (2013) model, explained 84.7% of the interaction total sum of squares and the most stable and high oil content peanut varieties across locations. Simion *et al.* (2018) reported significant genotypes, environments and G × E interaction for grain yield of 16 cowpea genotypes (*Vigna unguiculata*) evaluated on seven environments and stability analysis based on AMMI model (Gauch, 2013) and GGE model (Yan and Tinker, 2006) accounted for 79.33% and 62.47% of the total variation

respectively, and singled out the most stable and high yielding genotypes. The analysis by Simion *et al.* (2018) using GGE biplot managed to identify the most discriminating and representative test locations for cowpea in Ethiopia. Sousa *et al.* (2018) reported significant genotype, environment and G × E interaction for grain yield of 40 cowpea genotypes evaluated on three locations for three consecutive years. Stability analysis based on GGE model (Yan and Tinker, 2006) accounted for 66.05% (PCA1 and PCA2 explained 51.45% and 14.6 respectively) of the total variation due to G × E interaction, and identified the most stable and high yielding genotypes.

CHAPTER THREE

3.0 GENETIC DIVERSITY AMONG RELEASED AND ELITE SOYBEAN GENOTYPES IN UGANDA

3.1 Introduction

Repeated use of closely related lines has been narrowing the genetic variability of soybean germplasm, consequently increases its vulnerability to ever-changing biotic and abiotic stresses (Ssendege *et al*., 2015). Therefore, understanding genetic diversity of soybean is very crucial to broaden its genetic base (Kumawat *et al*., 2015). Several methods for genetic diversity studies in soybean accessions have been reviewed Bisen *et al*. (2015); Singh *et al*. (2010); Tantasawat *et al*. (2011); Song *et al*. (2013) and among the different methods, the use DNA markers has been considered more informative, reliable and stable as compared to the commonly used conventional methods like the use of phenotypic descriptors and pedigree analysis. Among various DNA markers used for molecular characterization and genetic diversity studies in soybean, simple sequence repeats (SSR) markers have been considered the molecular markers of choice because of their abundance, high polymorphism rate and high reproducibility (Kumawat *et al*., 2015). SSR markers have been used worldwide for soybean collections genetic diversity studies and high polymorphism at SSR loci have revealed both number of alleles per locus and the gene diversity (Moniruzzaman *et al*., 2019; Kumawat *et al*., 2015). Since the soybean germplasm in Uganda has been acquired from accessions of many different countries, their degree of genetic diversity were not understood (Ssendege *et al*., 2015) and therefore, the objective of this study was to determine the degree genetic diversity that exist among the advanced elite soybean genotypes and released varieties in Uganda.

3.2 Study site and materials

The study was carried out at Makerere University Agricultural Research Institute Kabanyolo (MUARIK) screen house and Biotechnology Laboratory. Thirty four genotypes were used in this study comprising of 23 advanced elite genotypes and 11 released varieties (Table 3.1).

Table 3.1: Description of genetic material which were used in the study

Entry	Name	Origin	Year of release
1	Duiker × 3N-5	Uganda-Makerere University	AYT
2	GC × 2N-1	Uganda-Makerere University	AYT
3	BSPS 48A-27-1	Uganda-Makerere University	AYT
4	BSPSS 48A-28-1	Uganda-Makerere University	AYT
5	NGDT8.11× 14.16B	Uganda-Makerere University	AYT
6	NII × GC 13.2	Uganda-Makerere University	AYT
7	BSPS 48A-25-1	Uganda-Makerere University	AYT
8	Nam II GC 17.3	Uganda-Makerere University	AYT
9	NII × GC 35.3-2	Uganda-Makerere University	AYT
10	NG 14.1 × UG5	Uganda-Makerere University	AYT
11	Nam 4M × 2N-2	Uganda-Makerere University	AYT
12	NII × 35.3-3	Uganda-Makerere University	AYT
13	G8586 × UG5	Uganda-Makerere University	AYT
14	NDGT 8.11× 3N-1	Uganda-Makerere University	AYT
15	BSPS 48A-28	Uganda-Makerere University	AYT
16	Bulindi 18.4B	Uganda-Makerere University	AYT
17	Maksoy 4N	Uganda-Makerere University	2014
18	BSPS 48A-24-1	Uganda-Makerere University	AYT
19	Bulindi 24.1A	Uganda-Makerere University	AYT
20	NII × GC 35.3-1	Uganda-Makerere University	AYT
21	NDGT 8.11 × 3N-2	Uganda-Makerere University	AYT
22	2N × GC	Uganda-Makerere University	AYT
23	Mak 3N × 1N	Uganda-Makerere University	AYT
24	NG 14.1 × NII-1	Uganda-Makerere University	AYT
25	Maksoy 3N	Uganda-Makerere University	2013
26	Maksoy 6N	Uganda-Makerere University	2017
27	Kabanyolo1	Uganda	1987
28	Maksoy 1N	IITA	2004
29	Maksoy 5N	Uganda-Makerere University	2013
30	Maksoy 2N	Uganda-Makerere University	2008
31	Namsoy 4 M	Uganda-NaCRRI	2004
32	Nam 1	Uganda-NaCRRI	1990
33	Namsoy 3	Uganda-NaCRRI	1995
34	Namsoy 2	IITA	1992

Source; (Namara, 2015)

3.2.1 DNA isolation

Total genomic DNA was collected from plants at two weeks per genotype following the cytel trimethyl ammonium bromide (CTAB) method of Maughan *et al*. (1995). At first, DNA quality and concentration was determined using NanoDrop ND-1000 spectrophotometer. The final concentration was adjusted to 50 ng/µl for amplification use in PCR analysis as described by Bisen *et al*. (2015).

3.2.2 Simple sequence repeats (SSR) analysis

A total for 31 SSR markers that were previously mapped and distributed on 20 linkage group by Cregan *et al*. (1999) were selected for initial screening. Gradient PCR was carried out for each primer with six randomly selected soybean DNA samples to standardize the annealing temperature for final amplification. Twenty one out of 31 SSR primers showed good amplification and were used for further study (Table 3.2). These 21 SSR primers were used for PCR analysis across 34 soybean samples. Most of the SSR markers used had an (AAT) motif due to their abundance and highly polymorphic nature in soybean genome (Narvel *et al.*, 2000). A total volume of 12 µl containing 2 µl genomic DNA (50 ng/µl), 5 µl of liquid premix, 4 µl of distilled water and 0.5 µl of each primer (10 nmol) was prepared for PCR (Bioneer, Inc, Republic of Korea). Amplification process was carried out in a thermocycler (G Storm, UK) with the following conditions; Initial denaturation at 95°C for 5 minutes followed by 35 cycles of 95°C for 2 minutes, annealing temperature at 48 - 60^{o}C for 50 seconds, extension at 72 °C for 50 seconds and a final extension at 72 °C for 7 minutes. The PCR products were fractionated by gel electrophoresis on 2% metaphor agarose gel (Lonza Bioscience, Singapore) stained with Gel RedTM Nucleic Acid Stain (10µl /100 µl of 1X TAE buffer) with a constant supply of 100 volts for 1 hour along with a 100-bp ladder as a size standard. Gel images were taken using a Bio Doc-ItTM Imaging System (Biotium, USA).

Table 3.2: SSR primers that were used for genotyping 34 soybean genotypes

SSR primer	Forward primer	Reverse primer	Linkage group
satt126	GCTTGGTAGCTGTAGGAA	ATAAAACAAATTCGCTGATAT	B2
satt329	GCGGGACGCAAAATTGGATTTAGT	GCGCCGAATAAAACGTGAGAACTG	A2
sat_409	GCGGAGGTTTGTGCATTTCTAGGTCTTC	GCGACGCGTATGTACATAAAATATGCTGTT	A2
satt717	GCGTTTTGTGATTTGTTTTCCTCATTTACT	GCGGCTATCAAACATTTTTACATGATGGTTA	A1
BE806308	GCGATTTGACCCCGTTCATACAT	GCGGCAGAAATCCGCTCTCTTTA	B1
satt173	TGCGCCATTTATTCTTCA	AAGCGAAATCACCTCCTCT	O
satt185	GCGCATATGAATAGGTAAGTTGCACTAA	GCGTTTTCCTACAATAATATTTCAT	E
satt409	CCTTAGACCATGAATGTCTCGAAGATA	CTTAAGGACACGTGGAAGATGACTAC	A2
SOYHSP176	TTTTTGTTTAAGTTACTGTACTGT	GCTAGTCTTCTACAACCTTCTA	F
satt411	TGGCCATGTCAAACCATAACAACA	GCGTTGAAGCCGCCTACAAATATAAT	E
satt431	GCGTGGCACCCTTGATAAATAA	GCGCACGAAAGTTTTTCTGTAACA	J
satt245	AACGGGAGTAGGACATTTTATT	GCGCCTCCTGAATTTCAAAGAATGAAGA	M
satt264	CCTTTTGACAATTATGGCATATA	GCATAGAAGGGCATCATTCAGAT	K
satt373	TCCGCGAGATAAATTCGTAAAAT	GGCCAGATACCCAAGTTGTACTTGT	L
satt440	TGAGAACGTTTGAAAAGAGAT	GAAGAGATTAAGCATAAAGAATACTT	I
satt406	GCGTGAGCATTTTTGTTT	TGACGGGTTTAATAGCAT	J
satt216	TACCCTTAATCACCGGACAA	AGGGAACTAACACATTTAATCATCA	D1b
sat_084	AAAAAAGTATCCATGAAACAA	TTGGGACCTTAGAAGCTA	N
satt211	GAAAAAGCCCACATCCAA	CATGGGCATGCAGTAACA	A1
satt126	GCTTGGTAGCTGTAGGAA	ATAAAACAAATTCGCTGATAT	B2
Sat_366	GCGGCACAAGAACAGAGGAAACTATT	GCGGACATGGTACATCTATATTACGAGTATT	J

Source; (Cregan *et al*., 1999)

3.3 Data analysis

The PCR products were analyzed by scoring the presence or absence of a band based on allele size for all 21 polymorphic primers. The SSR primer band appearing without ambiguity was scored as 1 (present) and 0 (absent) for each primer. To estimate genetic diversity, such as number of effective alleles, heterozygosity, fixation index, Shannon's information index, GenAIEx 6.51 software (Peakall and Smouse, 2012) was used. DARwin 6.0.21 software Perrier and Jacquemoud-Collet (2006) was used to determine genetic similarity among genotypes by estimating dissimilarity coefficients in a pairwise comparison using Jaccard's similarity coefficient. The hierarchical cluster analysis was performed using dissimilarity coefficients with unweighted pair-group method arithmetic average (UPGMA) clustering algorithim (Tantasawat *et al*., 2011). The effectiveness for cluster analysis was evaluated using 1000 bootstrapped replicates. The allelic diversity at each locus was determined as polymorphic information content (PIC) based on equation of Anderson *et al*. (1993),

$$PIC = 1 - \sum P_i^2$$

Where, P_i is the frequency of i^{th} allele in the set of genotypes analyzed, calculated for each SSR locus.

DARwin 6.0.21 software was used to perform principal coordinate analysis (PCoA) to depict multiple dimensions of the distribution of released and elite genotypes in a scatter plot Hipparagi *et al*. (2017) to complement the information obtained from hierarchical cluster analysis (Tantansawat *et al*., 2011).

3.4 Results

The results of expected genetic diversity parameters obtained at each locus across 34 soybean genotypes are presented in table 3.3.The results showed that out of 31 markers screened, only 21 were found polymorphic among 34 genotypes. A total of 59 alleles with an average of 2.85 alleles per locus were detected among the genotypes. The number of alleles varied from 2 (Satt126, Sat_409, Satt717, BE806308, Satt185, Satt264 and Sat_084) to 4 (Satt216, Satt431 and Satt411) and the frequency of major allele ranged from 0.324 on primer Satt431 to 0.882 on BE806308 with an average of 0.532. The fragment size of these 59 alleles varied from 100 to 375bp.

The PIC value ranged from 0.208 on BE806308 to 0.741 on Satt411 with an average of 0.5870 and the number of effective alleles varied from 1.658 with Satt717 to 3.642 with Satt411 with an average of 2.362. Shannon's information index ranged from 0.362 for BE806308 to 1.337 for Satt411 with an average of 0.894 (Table 3.3).

The observed heterozygosity varied from 0.000 to 0.088 with an average of 0.010, while the expected heterozygosity ranged from 0.208 on BE806308 to 0.725 on Satt411, with an average of 0.548. Fixation index varied from 0.862 to 1.000 with an average of 0.983 (Table 3.3).

Table 3.3: Estimated genetic diversity parameters obtained at each locus across 34 soybean genotypes

Marker	Na	Allele size range (bp)	Major allele freq	Ne	I	Ho	He	F	PIC
BE806308	2	200 – 210	0.882	1.262	0.362	0.000	0.208	1.000	0.208
Sat_084	2	151 – 180	0.618	1.895	0.665	0.000	0.472	1.000	0.472
Sat_366	3	185 – 205	0.471	2.762	1.058	0.088	0.638	0.862	0.587
Sat_409	2	150 – 200	0.559	1.867	0.657	0.000	0.464	1.000	0.583
Satt126	2	120 – 148	0.647	1.832	0.647	0.030	0.454	0.933	0.457
Satt173	3	210 – 297	0.441	2.689	1.036	0.000	0.628	1.000	0.638
Satt185	2	250 – 270	0.500	2.000	0.693	0.000	0.500	1.000	0.500
Satt211	3	100 – 180	0.618	1.934	0.760	0.000	0.483	1.000	0.493
Satt216	4	150 – 210	0.382	3.051	1.184	0.000	0.672	1.000	0.693
Satt245	3	180 – 211	0.559	2.439	0.991	0.000	0.590	1.000	0.590
Satt264	2	200 – 220	0.559	1.973	0.686	0.000	0.493	1.000	0.493
Satt285	3	200 – 250	0.647	1.984	0.855	0.000	0.496	1.000	0.525
Satt329	3	250 – 300	0.559	2.355	0.958	0.030	0.575	0.947	0.587
Satt373	3	225 – 290	0.441	2.847	1.073	0.000	0.649	1.000	0.649
Satt406	4	406 – 375	0.353	3.501	1.312	0.031	0.714	0.956	0.733
Satt409	3	183 – 200	0.441	2.604	1.027	0.000	0.616	1.000	0.721
Satt411	4	100 – 160	0.353	3.642	1.337	0.000	0.725	1.000	0.741
Satt431	3	205 – 250	0.324	2.985	1.096	0.033	0.665	0.950	0.720
Satt440	3	185 – 215	0.588	2.067	0.883	0.000	0.516	1.000	0.598
Satt717	2	240 – 250	0.706	1.658	0.586	0.000	0.397	1.000	0.431
SoyHSP176	3	100 -170	0.529	2.256	0.913	0.000	0.557	1.000	0.607
Total	59	-	11.177	49.601	18.778	0.213	11.513	20.648	12.026
Mean	2.81	-	0.532	2.362	0.894	0.010	0.548	0.983	0.573
SE	0.15	-	-	0.134	0.055	0.005	0.026	0.008	-

Na=number of alleles, Ne=number of effective alleles, I=Shannon's information index, Ho= observed heterozygosity, He=expected heterozygosity (gene diversity), F= fixation index, PIC= polymorphic information content, SE= standard error

Hierarchical cluster analysis

The results of the hierarchical cluster analysis which was performed to show the genetic relationship among the genotypes on the dendrogram which was constructed from genetic distance dissimilarity matrix (only bootstraps >50% were shown) are presented on Figure 3.1. The results showed that all the 34 genotypes were grouped into three major clusters, CI, CII and CIII, with cluster I comprising of 14 genotypes (Maksoy 3N, Maksoy 1N, Maksoy 2N, Nam 1, Nam 2, Namsoy 3, Namsoy 4M, Kabanyolo 1, Bulindi 18.4B, Bulindi 24.1A, GC × 2N-1, BSPS 48-24-1, NDGT.8.11 × 3N-2, NII×GC 35.3-1), cluster II comprising of 11 genotypes (Maksoy 4N, BSPS 48A-28, BSPS 48-27-1, BSPS 48A-28-1, BSPS 48A-25-1, NDGT 8.11×3N-1, NDGT 8.11×14.16B, NG 14.1×NII-1, NII×GC 13.2, Duiker×3N-5 and G8586×UG5) and cluster III comprising of 9 genotypes (Maksoy 6N, Maksoy 5N, NamII GC 17.3, Nam 4M× 2N-2, NG 14.1×UG5, NII×35.3-2, Mak 3N×IN and 2N×GC) genotypes, respectively. Cluster I was further partitioned into two distinct sub-clusters Ia and Ib with 11 and 3 genotypes, respectively. Almost all released varieties by Makerere University Soybean Research Centre Kabanyolo and Namulonge soybean breeding program (in red labels on the dendrogram) belong to sub-cluster Ia, while the sub-cluster Ib comprised of four advanced elite genotypes from Makerere University Soybean Research Centre Kabanyolo. Cluster II forms two distinct sub-clusters IIa and IIb with 7 and 4 genotypes respectively. Most of the advanced elite genotypes fall in cluster II, with only one released variety Maksoy 4N. Cluster III also consisted of two sub-groups IIIa and IIIb comprised of 7 and 2 genotypes, respectively. This cluster comprised of genotypes from Makerere University Soybean Research Centre Kabanyolo only, with two recently released varieties Maksoy 6N and Maksoy 5N belonging to this category.

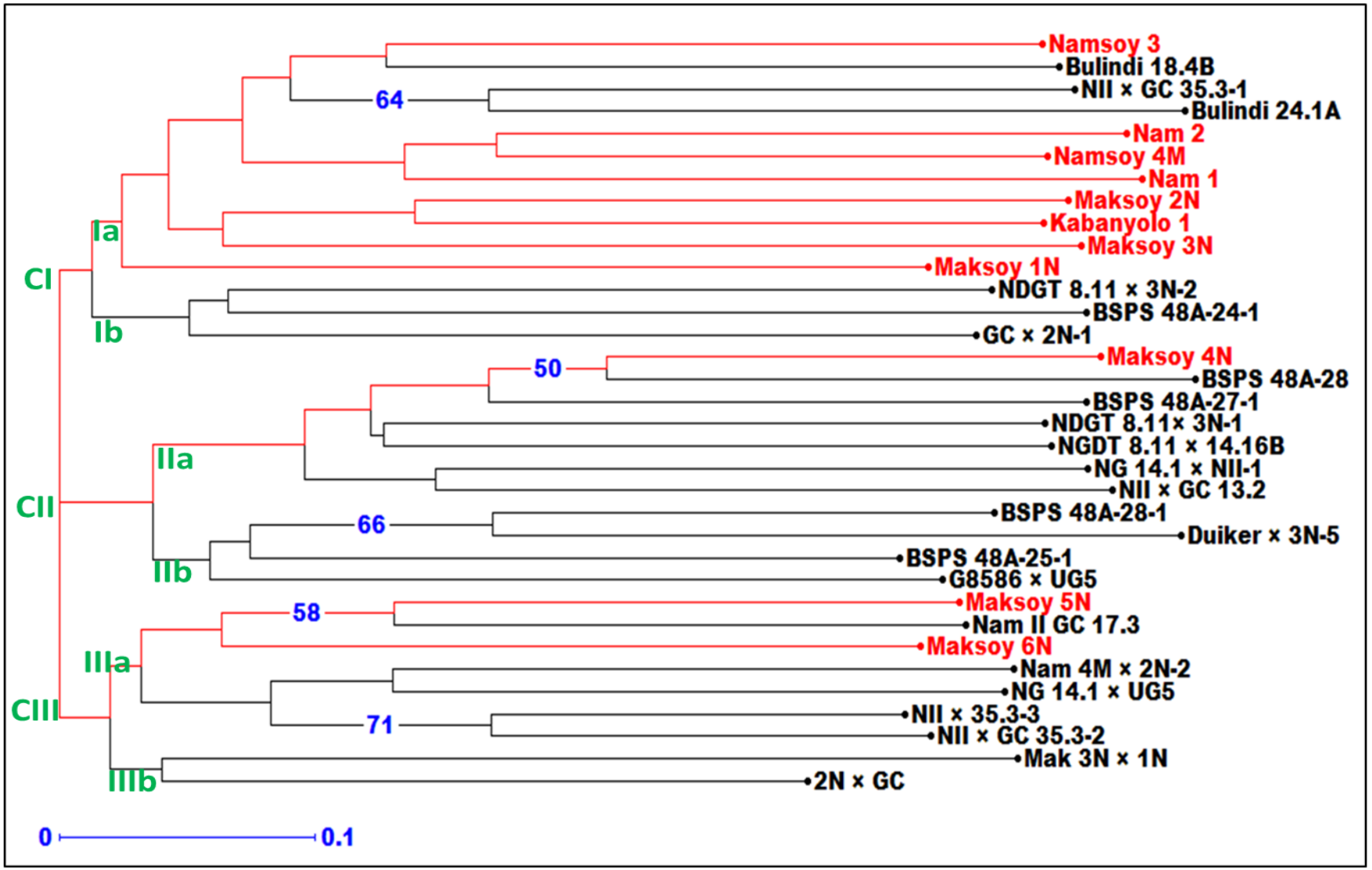

Figure 3.1: Dendrogram showing genetic diversity among 34 soybean genotypes based on UPGMA clustering of Jaccard's similarity coefficients. (Red labels = released varieties; black labels = advanced elite genotypes)

Principal coordinate analysis (PCoA)

The results of the Principle Coordinate Analysis (PCoA) performed to further assess genetic diversity among released varieties and elite soybean genotypes are presented in Figure 3.2. The results showed that three groups of genotypes were identified, with the first two axes discriminating the released varieties and advanced elite soybean genotypes. The PCoA explained 28.9% of the total variation with PC1 and PC2 accounted for 17.2% and 11.7% of the total variation, respectively. Red labels in Figure 3.2 represent released soybeans varieties and black label represent advanced elite genotypes.

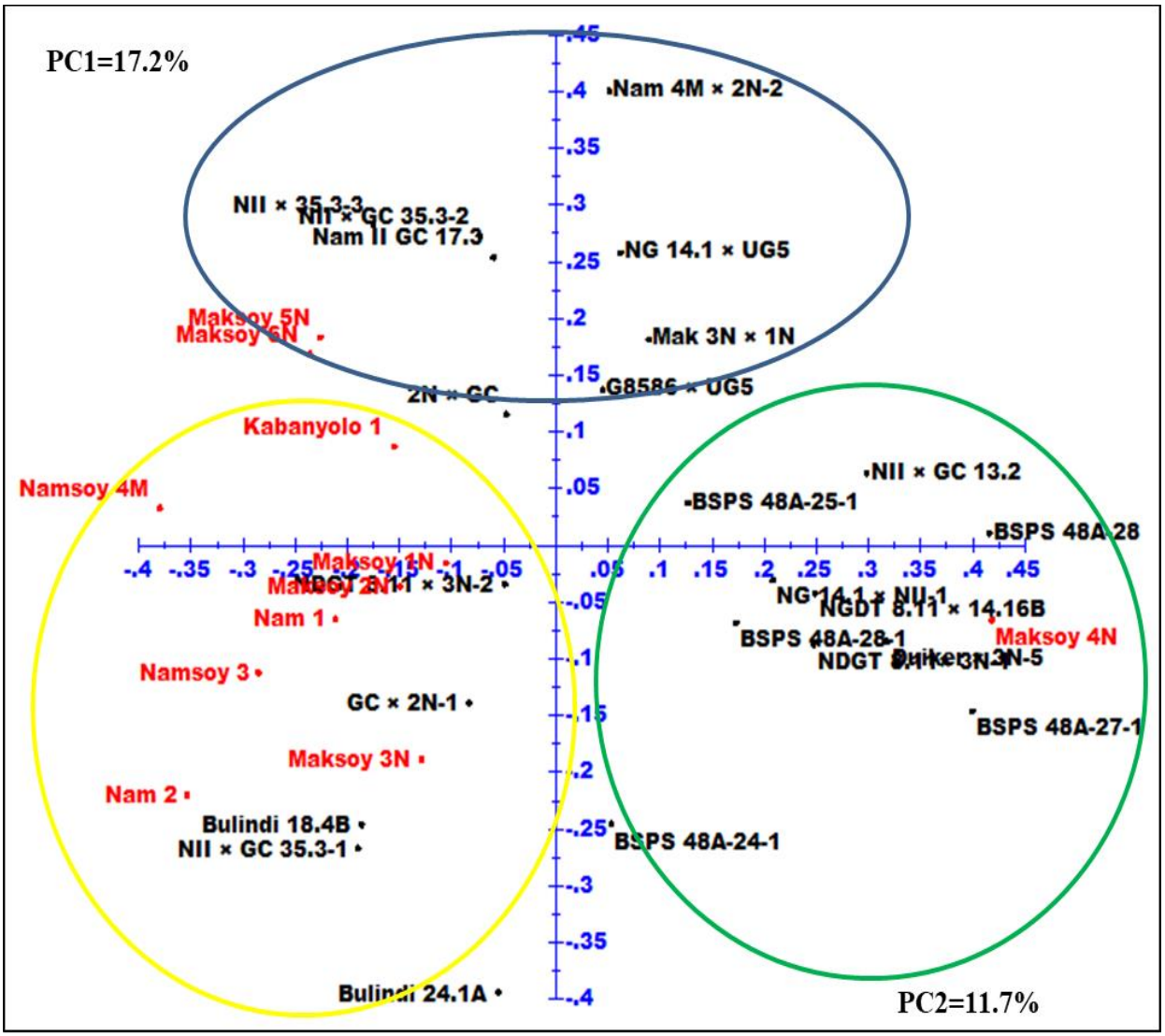

Figure 3.2: Principal coordinates analysis (PCoA) for 34 soybean genotypes on the basis of SSR marker data (Red labels released varieties, and black labels advanced elite genotypes**)**

The results of the rare allele sizes found among the genotypes on different locus are presented on table 3.4. The results showed that four unique or rare alleles were amplified by four SSR markers (Sat431, Sat_366, Satt126 and Satt406) on five different genotypes that are BSPS 48-28-1, Maksoy 4N, Maksoy 1N, NamII GC 17-3 and NDGT 8.11×3N-2. Sat_366 amplified one rare allele of size 185/205bp among three genotypes: Maksoy 1N, NamII GC 17-3 and NDGT 8.11×3N-2, while Satt431 amplified one rare allele of size 230/250bp on Maksoy 4N. Markers Satt406 and Satt126 amplified single rare allele of size of 300/375 and 120/148 respectively, on the same genotype BSPS 48-28-1.

Table 3.4: Allele size of SSR markers in soybean genotypes with unique alleles

Marker	Unique alleles	Allele size (bp)	Genotypes
Sat_366	1	185/205	Maksoy 1N NDGT 8.11 × 3N-2 Nam II GC 17.3
Satt431	1	230/250	Maksoy 4N
Satt406	1	300/375	BSPS 48A-28-1
Satt126	1	120/148	BSPS 48A-28-1

The results of gel electrophoresis showing different allele sizes amplified by marker Satt373 across all the genotypes are shown on Figure 3.3. The results showed that Satt373 had three allele sizes 225bp, 254bp and 290bp based 100bp ladder labeled M on Figure 3.3, with genotypes: 1; 3; 5; 6; 14; 15; 17; 24; 28; 30; 33 had the same allele size of 225bp, genotypes: 8;11;22;23;26;27 and 27 had the same allele size of 254bp and genotypes: 2; 4; 7; 9; 10; 12; 13; 16; 18; 19; 20; 21; 25; 31; 32 and 34 had the same allele size of 290bp. The numbering of genotypes on Figure 3.3 and Figure 3.4 had the same codes and labels as on Table 3.1.

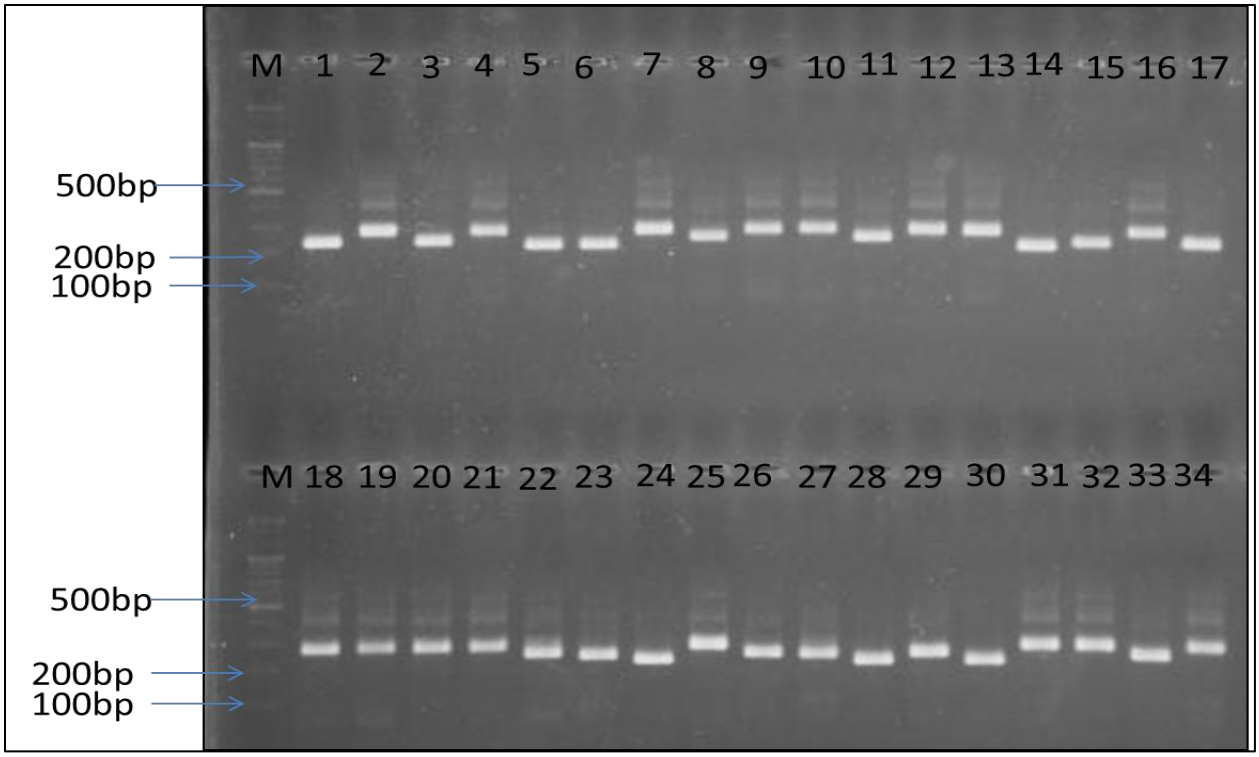

Figure 3.3: The gel electrophoresis DNA bands amplified by Satt373 marker across 34 soybean genotypes

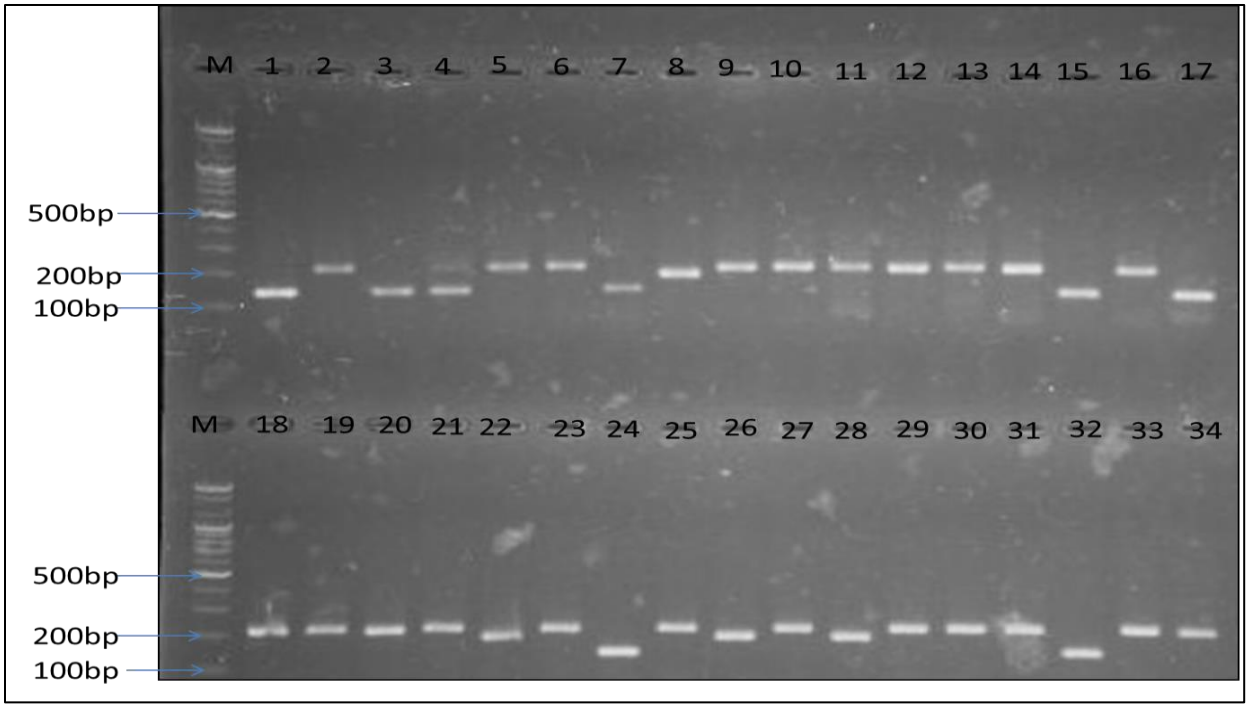

Figure 3.4: The gel electrophoresis DNA bands amplified by Satt216 marker across 34 soybean genotypes

3.5 Discussion

The assessment of genetic diversity among crop genotypes is not only significant for crop improvement but also important for precise efficient management and protection of the cultivars. The reliability, reproducibility and authentic results obtained from using SSR markers have made them widely preferred in genetic diversity studies. Out of 31 SSR primer pairs screened, only 21 primer pairs, distributed on 15 of 20 linkage groups of soybean (Cregan *et al*., 1999), amplified scorable bands. From this study, a total of 59 alleles with an average of 2.85 alleles per locus were detected among genotypes evaluated. The number of alleles varied from 2 (Satt126, Sat_409, Satt717, BE806308, Satt185, Satt264 and Sat_084) to 4 (Satt216, Satt431 and Satt411). However, the lower allele number indicates low allelic diversity in present set of soybean genotypes. Allelic richness (average number of alleles per locus) is an effective index for diversity evaluation but it is largely dependent on the sample size (Hipparagi *et al*., 2017). Hence to improve the allelic richness more landraces needs to be introduced into the breeding programme thus, enhancing genetic diversity (Widaningsih *et al*., 2014). The number of alleles observed in this study is comparable to those reported by Kumawat *et al*. (2015) where 2.9 alleles were detected per locus with an average polymorphic information content (PIC) value of 0.58. Similar results were also reported by Hipparagi *et al.* (2017) where 2.61 alleles with an average PIC value of 0.36 were detected among 75 soybean genotypes assayed by 21 SSR markers. However, lower allelic diversity in soybean was reported by Bisen *et al.* (2015) when they detected 2.22 alleles per locus with an average PIC value of 0.199 in 38 soybean genotypes using 16 SSR markers.

The polymorphic information content (PIC) (a measure of the allelic diversity of SSRs) observed in this study was 0.573. The high PIC value observed in this study was consistent with previous studies (Widaningsih *et al*., 2014; Kumawat *et al.*, 2015; Ghosh *et al*., 2014; Wang *et al.*, 2010). A total of 20 markers excluding BE806308 had the PIC values greater than 0.4 (Table 3.3) indicating that these markers were highly informative for discriminating and distinguishing released varieties and advanced elite genotypes. Also, these markers with high PIC values occurred on 14 separate linkage groups (Song *et al*., 2010), indicating that molecular polymorphism was spread across different regions of the genome. The PIC was highest for the SSR primer Satt411 (0.741) and was lowest for the primer BE806308 (0.208). Hence, primer

Satt411 is highly informative in the present study. This indicated that the primer Satt411 might be an effective and useful tool to determine the genetic differences among the soybean accessions and to study the phylogenetic relationship.

Gene diversity, also referred to as the expected heterozygosity (He) (Li *et al*., 2008), varied from 0.208 (BE806308) to 0.725 (Satt411) with an average of 0.548. This implied that there was moderate genetic variation existing among the evaluated genotypes. The gene diversity observed in this study was lower than the previous studies reported by Widaningsih *et al*. (2014) (0.66); Song *et al*. (2013) (0.648); Zhao *et al*. (2018) (0.879); Wang *et al*. (2015) (0.80). This may be ascribed to the emphasis on direct introductions, selection from introduced germplasm and single cross hybrids (some of which shared common parents) in the soybean breeding programme. Therefore, inclusion of more diverse germplasm and use of landraces in the soybean breeding programmme may provide the level of genetic variation necessary to cope with dynamic pests, diseases and abiotic stresses of Uganda (Ssendege *et al*., 2015). However, moderate genetic variation was also reported by Hipparagi *et al*. (2017) who observed gene diversity of 0.43 among 75 genotypes assayed by 21 SSR markers. The heterozygosity obtained from the present study was 0.010. The reason for low heterozygosity is to due the fact that soybean is strictly a self-pollinated crop which is expected to have low heterozygosity than mostly cross breeding crops *(Zhang et al*., 2013). Similar studies reported by Hipparagi *et al*. (2017) gave 0.11; and Zhao *et al*. (2018) reported 0.11 heterozygosity, so the heterozygosity observed in the present study was much lower than the value reported by the other workers but some others studies reported low heterozygosity in the crop, for example Song *et al*. (2013) (0.045) and Li *et al*. (2008) (0.014). The Shannon's information index of 0.894 was lower than the results obtained by Zhao *et al*. (2018) who reported Shannon's information index of 2.528. Moreover, Zhao *et al*. (2018) reported fixation index of 0.987 which is comparable to the one obtained in this study. The results of Shannon's information index and fixation index revealed a shift from Hardy-Weinberg equilibrium; hence the present findings conclude the presence of moderate genetic variation existing among the genotypes. The moderate genetic variation obtained in this study could be attributed to the fact that these genotypes might sharing some common parental lines since they were acquired from Makerere University soybean breeding programme and Namulonge vegetable oil breeding programme which are located in the same country.

The unique or rare alleles are mainly due to natural mutation and selection (Bisen *et al*., 2015; Kumawat *et al*., 2015). In this study, four unique alleles were detected by four SSR markers (Satt431, Sat_366, Satt126 and Satt406) on five genotypes: BSPS 48-28-1, Maksoy 4N, Maksoy 1N, NamII GC 17-3 and NDGT 8.11×3N-2 (Table 3.4). Three of the SSR markers amplifying unique alleles were known to be linked with three different phenotypic traits (Liu *et al*., 2011). The Satt406 was reported to be linked with photoinsensitivity trait. The Satt406 had amplified a unique allele in genotype BSPS 48-28-1. Another marker Satt126 known to be linked with yield QTLs (Du *et al*., 2009) had amplified a unique allele in genotype BSPS 48-28-1. The last marker Sat_366 known to be linked with protein QTLs (Bisen *et al.,* 2015) had amplified a unique allele in genotype Maksoy 1N. This genotype, Maksoy 1N is known for its high protein content. Therefore, the genotypes identified for photoinsensitivity, yield QTLs and protein QTLs marker alleles can be used in marker assisted introgression programme but further validation is required for marker traits linkage in segregating populations. In addition, these unique markers are highly reliable and informative in detecting genetic relationship among the soybean genotypes (Kumawat *et al*., 2015). Also, these unique or rare alleles found in those specific genotypes can be used for the identification of genes or alleles for improvement of protein and yield in future soybean breeding studies (Bisen *et al*., 2015). Kumawat *et al*. (2015) suggested that genotypes showed the presence of unique or rare alleles may serve as good sources for identification of new alleles of important genes. Therefore these five genotypes BSPS 48-28-1, Maksoy 4N, Maksoy 1N, NamII GC 17-3 and NDGT 8.11×3N-2 may serve as good sources for identification of new alleles of important genes.

The hierarchical clustering divided the soybean genotypes into three clearly distinct groups, with more than half of the released varieties grouped together indicating that most released varieties were developed from common parents. This implied that most of the released varieties will lose stability as well as resistance in case of changes in pests, diseases and abiotic stresses. The results of both the dendrogram and Principal Coordinate analysis (PCoA) were consistent with each other in grouping 34 soybean genotypes into three distinct classes (Figure 3.1 and 3.2). The first two axes of PCoA accounted 28.9% of the total variation implying that moderate genetic variation existed in the genotypes studied and hence the molecular markers grouping was more logical as compared to grouping based on origin and pedigree. Similar results were reported by Ssendege *et al*. (2015), where Maksoy 1N, Maksoy 2N and Maksoy 3N were grouped in the

same cluster which indicated that they were genetically related thus were likely to have shared similar pedigree. The results indicated the presence of moderate genetic variation among the genotypes evaluated. The presence of moderate genetic variation is not much favorable in soybean breeding since there are dynamic pests, diseases and abiotic stresses which can accelerate vulnerability of closely related cultivars to outbreak of pests, diseases and abiotic stresses (Moniruzzaman *et al*., 2019) while, the moderate genetic variation observed among soybean genotypes in present study necessitates the need of broadening genetic diversity by introducing more exotic germplasm along with utilization of wild relatives, the diverse genotypes identified in this study may serve as source of new alleles in soybean breeding programme of Uganda in order to cope with ever-changing outbreak of pests, diseases and abiotic stresses.

CHAPTER FOUR

4.0 GENOTYPE × ENVIRONMENT INTERACTIONS FOR GRAIN YIELD, PROTEIN AND OIL CONTENT OF ADVANCED ELITE SOYBEAN LINES

4.1 Introduction

Soybean is mainly grown in diverse environments in Uganda, and has been reported to be sensitive to genotype × environment (G × E) interaction (Tukamuhabwa *et al*., 2012). The presence of significant G × E interaction for grain yield, percentage protein and oil content in soybeans has been reported by several researchers (Tukamuhabwa *et al*., 2012; Gurmu *et al*., 2009; Atnaf *et al*., 2013; Bhartiya *et al*., 2017). Multi-environment yield trials are commonly used for identifying superior advanced soybean genotypes to be released as new varieties for specific or broad environments in Uganda soybean breeding programs (Tukamuhabwa *et al*., 2012). Although soybean is regarded as a resilient crop to various stresses (Tukamuhabwa *et al*, 2016), its grain yield, protein and oil content performances are greatly influenced by G × E interaction (Gurmu *et al*., 2009). The presence of G × E interaction reduces the correlation between phenotype and genotype and subsequently responses to selection during breeding programme (Gasura *et al*., 2015). Moreover, the presence of G × E interaction requires that new varieties be tested in several environments for identification of multi-environmentally adapted genotypes (Bernado, 2002) and identification of the best testing environment (Yan *et al*., 2010). Several statistical options for reducing or exploiting G × E interaction in soybean multi-environment trial data have been reviewed (Tukamuhabwa *et al*., 2012; Gurmu *et al*., 2009; Atnaf *et al*., 2013; Bhartiya *et al*., 2017) for instance joint regression model (Eberhart and Russell, 1966; Finlay and Wilkinson, 1963); AMMI model analysis (Gauch, 2013), and the GGE biplot model (Yan and Tinker, 2006). Therefore, the objectives of this study were to evaluate the yield, protein and oil content stability of advanced elite soybean lines and determine ideal testing locations across diverse environments for future breeding activities in Uganda.

4.2 Description of the study sites

The study was carried out at six locations; On-station (Kabanyolo), TVC (Iki-Iki and Nakabango) ZARDIs (Ngetta, Abi and Bulindi) that are located in different agro-ecological regions of Uganda (Table 4.1). These locations have different climatic conditions and therefore may influence the expression of soybean grain yield, protein and oil content differently. Also, these locations they represent major soybean growing areas of Uganda.

Table 4.1: Description of the six selected experimental sites used to evaluate grain yield and protein content and yield related traits during season 2018A and 2018B

Site	Coordinates	Altitude (m)	Mean annual Temperature (^{0}C)	Mean annual rainfall (mm)	Soil type
Nakabango	00^0 31'N $33^0$12'E	1178	26	1400	Crysalline basic
Iki-Iki	01^0 06'N 34^0 00'E	1156	28	1200	Sandy
Kabanyolo	00^0 28'N 32^0 37'E	1300	22	1255	Sand-clay loam
Bulindi	01^0 28'N 31^0 28'E	1230	23	1700	Sandy loam
Ngetta	02^0 17'N 32^0 56'E	1085	29	1483	Sandy loam
Abi	03^0 5'N 30^0 56E	1140	24	1250	Sandy-clay loam

Source: NARO Ngetta-ZARDI (2018)

4.3 Description of planting materials

The planting materials were accessed from MUARIK Soybean Breeding Programme. Twenty five genotypes were used in this study. Among the genotypes used, 23 were advanced elite genotypes (genotypes that have passed preliminary and intermediate tests; only waiting for yield, protein and oil content stability evaluation across soybean major growing areas of Uganda) and two were commercial varieties used as checks. The details on the characteristics of the genotypes used are presented on Table 4.2.

Table 4.2: Description of genetic material evaluated for grain yield performance, protein and oil content

Name	Pedigree	Protein content (%)	Oil content (%)	Yield (t/ha)
G1	Duiker × 3N-5	-	-	-
G2	GC × 2N-1	-	-	-
G3	BSPS 48A-27-1	-	-	-
G4	BSPSS 48A-28-1	-	-	-
G5	NGDT8.11×14.16B	-	-	-
G6	NII × GC 13.2	-	-	-
G7	BSPS 48A-25-1	-	-	-
G8	Nam II GC 17.3	-	-	-
G9	NII × GC 35.3-2	-	-	-
G10	NG 14.1 × UG5	-	-	-
G11	Nam 4M × 2N-2	-	-	-
G12	NII × 35.3-3	-	-	-
G13	G8586 × UG5	-	-	-
G14	NDGT 8.11× 3N-1	-	-	-
G15	BSPS 48A-28	-	-	-
G16	Bulindi 18.4B	-	-	-
G17	Maksoy 4N	38	21	2-3.5
G18	BSPS 48A-24-1	-	-	-
G19	Bulindi 24.1A	-	-	-
G20	NII × GC 35.3-1	-	-	-
G21	NDGT 8.11×3N-2	-	-	-
G22	2N × GC	-	-	-
G23	Mak 3N × 1N	-	-	-
G24	NG 14.1 × NII-1	-	-	-
G25	Maksoy 3N	36	22	2-3.5

Namara (2015)

4.4 Experimental design

A randomized complete block design (RCBD) with three replications was used. Two commercial varieties; Maksoy 3N and Maksoy 4N were used as checks. Each soybean line was planted in three rows with spacing of 60cm between and 5cm within rows. Weeding was done 3 times which kept the fields free of weeds.

4.5 Data collection

Data were collected on soybean rust, a major soybean disease in Uganda using a scale of 1-5 (Miles *et al*., 2006) for all the diseases where 1= no visible lesion, 2= few scattered lesions present, 3= moderate number of lesions on at least part of the leaf, 4= abundant number of lesions on at least part of leaf, and 5= prolific lesion development on most of the leaf. Days to 50% flowering, plant height and pod clearance was recorded as described by Obua (2013). Lodging was recorded on a scale of 1 to 5; with 1 being the most upright and 5 where soybean are completely prostrate. Number of pods per plant was recorded at physiological maturity. At harvest the genotypes were threshed and 100 seed weight and yield per plot determined and later corrected to 12% moisture content before determining yield per hectare (Tukamuhabwa *et al*., 2012). Protein content (%) was quantified using the data from first and second replications. The analysis described by Owusu-Apenten (2002) was used to quantify the protein content, whereas, the oil content was determined using Near infrared spectroscopic analysis as described by Sato (2010). Protein and oil content were determined at the Department of Biosciences at National Crops Resources Research Institute (NaCRRI) Namulonge.

4.6 Data analysis

Analysis of variance (ANOVA) was performed initially for each environment to determine performance of the genotypes in different environments. Combined analysis of variance over locations and seasons was conducted using mixed model as suggested by Moore and Dixon, (2015) (where genotypes and locations were fixed whereas seasons, all the interactions involving seasons, replications and error were considered random) in GenStat software version 18 (Genstat, 2016) to determine the performance of different genotypes across seasons and locations and establish G × E interaction for grain yield, oil and protein content. The following model for the combined analysis of variance was used as reported by Gasura *et al*. (2015);

$$Y_{ijkl} = \mu + r_1(pt)_{jk} + g_i + p_j + t_k + (gp)_{ij} + (gt)_{ik} + (pt)_{jk} + (gpt)_{ijk} + e_{ijkl}$$

Where: $Y_{ijkm(l)}$ is observed value of *i*th genotype in the *j*th location and the *k*th season in the *l*th replication, μ is the grand mean, $r_1(pt)_{jk}$ is the effect of the *l*th replication within locations and seasons, g_i, p_j and t_k are the main effects of the genotype, locations and seasons, $(gp)_{ij}$, $(gt)_{ik}$,

$(pt)_{jk}$ are the first order interactions and $(gpt)_{ijk}$ is the second-order interaction, and finally e_{ijkl} is the pooled error term.

The proper F-test for a mixed model in which genotypes and locations are considered fixed effects and seasons treated as random effects was applied as suggested by (Mclntosh, 1983) and recently by Moore and Dixon (2015). The assumption of sum to zero the effects of random interactions across each level of a fixed factor for combined experiments was used as described by Moore and Dixon (2015). In brief, the mean squares for genotypes, genotypes × locations, genotypes × seasons and genotypes × locations × seasons were tested against the pooled error mean square, while locations, seasons and locations × seasons were tested against the mean square of replications within locations and seasons (Mclntosh, 1983). The variance components due to genotypes (δ^2g), genotypes × location (δ^2gl), genotypes × seasons (δ^2gs), genotypes × locations × seasons (δ^2gls) and random error (δ^2error) were obtained by solving the equations formed by equating the mean squares to their respective expected mean squares (Moore and Dixon, 2015). The variance components due to environments (location × seasons combinations) were estimated by summation of δ^2l, δ^2s and δ^2ls, whereas the variance component attributed to genotype × environment (δ^2ge) was estimated by adding up δ^2gl, δ^2gs and δ^2gls (Mclntosh, 1983). The broad sense coefficients of genetic determination (BSCGD) (broad sense heritability based on fixed genotypes) on a single plot basis, single environment basis and across environments basis were obtained by solving the following equations as $\delta^2g/(\delta^2g + \delta^2gl + \delta^2gs + \delta^2gls + \delta^2error)$; $\delta^2g/ (\delta^2g + \delta^2gl + \delta^2gs + \delta^2gls + \delta^2error/ nr)$ and $\delta^2g/(\delta^2g + \delta^2gl/nl + \delta^2gs/ns + \delta^2gls/nls + \delta^2error/ nslr)$, respectively, where nr = number of replications, nl = number of locations, ns = number of seasons, nls = number of location × seasons combinations and nslr is the number of seasons × location × replications (Moore and Dixon, 2015).

Yield data was further subjected to two methods, GGE biplot (Yan and Tinker, 2006) and AMMI analysis (Gauch, 2006) for identification of stable soybean genotypes. Also, the GGE analysis was performed to determine the mega-environments and visualize the "which-won-where" pattern following the model for GGE biplot based on singular value decomposition (SVD) of *t* principal components was used as described by Yan and Tinker (2006).

GGE model: $Y_{ij} - \mu_i - \beta_j = \sum_{k=1}^{t} \lambda_k \alpha_{ik} \gamma_{jk} + \varepsilon_{ij}$

Where: Y_{ij} is the performance of genotype *i* in environment *j*, μ is the grand mean, *j* b is the main effect of environment *j*, *k* is the number of principal components (PC); λ_k is singular value of the k^{th} PC; and α_{ik} and γ_{jk} are the scores of i^{th} genotype and j^{th} environment, respectively for PC_k; ε_{ij} is the residual associated with genotype *i* in environment *j*.

For mega-environment delineation of test locations, the which-won-where scatter plot was generated by a polygon drawn by connecting genotypes that are furthest away from the biplot such that the polygon contained all other genotypes. Then the polygon was dissected by perpendicular lines drawn to the polygon sides and running from the biplot origin (Yan and Tinker, 2006). The genotype comparison biplot for visualization and comparing genotypes based on mean yield and stability was achieved by representing an average environment by an arrow. A straight line that dissecting the biplot origin to the average environment (average genotype axis) was drawn followed by a perpendicular line that passes through the biplot origin using the appropriate singular value partitioning (SVP) methods (Yan and Tinker, 2006). For the analysis of test locations, location comparison biplot was used for identification of ideal testing site (the most discriminating and representative locations) (Gasura *et al*., 2015). The environment vectors were drawn from the location comparison biplot origin to the markers of the environment (Yan and Tinker, 2006). The following AMMI model was used as described by Gauch, (2013).

AMMI Model: $Y_{ijk} = \mu + G_i + E_j + \sum_{k=1}^{m} \lambda_k \alpha_{ik} \gamma_{jk} + \rho_{ij}$

Where: Y_{ijk} = the yield of the i^{th} genotype in the j^{th} environment, G_i = the mean of the i^{th} genotype minus the grand mean, E_j = the mean of the j^{th} environment minus the grand mean, λ_k = the square root of the eigen value of the k^{th} IPCA axis, α_{ik} and γ_{jk} = the principal component scores for IPCA axis k of the i^{th} genotypes and the j^{th} environment, ρ_{ij} = the deviation from the model.

4.7 Results

The results of combined analysis of variance for grain yield, protein and oil content (%) are presented in table 3.3. The results for grain yield showed significant ($p<0.05$) differences for all components except genotypes × season interaction which was non-significant. The results for protein and oil content (%) showed non significance ($p>0.05$) for genotype, seasons, genotype × location interaction, genotype × season interaction and genotype × location × season interaction

except location which was significant ($p<0.05$). The broad sense coefficient of genetic determination (BSCGD) (equivalent to broad sense heritability of fixed genotypes) for grain yield on single plot basis, single environment basis and across environment basis were 3%, 6% and 40% respectively.

Combined analysis of variance results for agronomic traits are presented in table 3.4. The results showed significant difference for genotype across all the traits ($p< 0.001$) (100 seed weight, groundnut leafminer damage, stem lodging, number of pods, plant height, pod clearance, pod shattering, days to 50% flowering). Genotype × season was significant ($p<0.01$) 100 seed weight. Genotype × location interaction was significant ($p<0.001$) for the following traits 100 seed weight, groundnut leafminer damage, number of pods, plant height, pod clearance, pod shattering. However, stem lodging and days to 50% flowering showed non-significant interaction of genotype × location. Genotype × location × season interaction showed significant ($p<0.05$) for 100 seed weight, stem lodging and pod shattering, however, groundnut leafminer damage, number of pods, plant height, pod clearance, days to 50% flowering showed non-significant.

Table 4.3: Combined analysis of variance for grain yield, protein and oil content over locations and two seasons (2018 A and 2018B)

	GY (kg/ha)		Protein (%)	Oil (%)
Source of Variation	Df	MS	Df	MS	MS
Season	1	161551161***	1	0.971ns	2.192ns
Location	5	60563205***	3	72.012***	81.716*
Season. Location	5	10263901**	3	1.122ns	93.784**
Replication. Season. Location	24	2274739***	8	2.424ns	11.414**
Genotype	24	320102***	24	7.552ns	3.81ns
Genotype × Location	120	195916**	24	6.122ns	2.72ns
Genotype × Season	24	152514ns	72	5.443ns	4.419ns
Genotype × Location × Season	120	193393*	72	8.034ns	4.043ns
Pooled Error	576	142233	192	7.668	4.597
LSD		770.1		5.386	4.35
CV (%)		23.4		8.3	13.4
δ^2g		4940.81		-	-
δ^2gl		8947.17		-	-
δ^2gs		571.17		-	-
δ^2gls		17053.33		-	-
δ^2error		142233		-	-
H2 on single plot basis		0.03		-	-
H2 individual environment basis		0.06		-	-
H2 on across environment basis		0.41		-	-

***=p<0.001; **=p<0.01; *=p<0.05; ns=not significant; GY= grain yield; G= genotype; H2 = broad sense heritability; δ^2g= variance component due to genotype; δ^2gl= variance component due to genotype × location; δ^2gs= variance component due to genotype × season; δ^2gls= variance component due to genotype × location × season

Table 4.4: Combined analysis of variance for agronomic traits evaluated across six locations and two seasons (2018 A and 2018B)

		100SDWT	GLM	LODG	NPODS	PH	PODCL	PODSHAT	RUST	DT50%F
Source of variation	Df	MS	MS	MS	MS	MS	MS	MS	MS	MS
Season	1	4364.804***	1.778ns	1.068ns	2240.210*	548.260ns	152.289*	0.034ns	21.623***	3669.694***
Location	5	892.471***	26.374***	1.069*	15979.910***	22448.510***	1022.527***	4.474***	1.250*	6540.814***
Season. Location	5	498.108***	1.778*	1.768**	4202.310***	2904.020**	237.345***	0.034*	1.223ns	120.944***
Rep/season. Location	24	8.648	0.462	0.297	452.94	571.18	21.653	0.299	0.47	18.583
Genotype	24	33.899***	0.245***	0.305***	282.240***	2360.240***	92.170***	1.520***	1.510***	41.179***
Genotype × Season	24	4.244**	0.025ns	0.141ns	14.760ns	28.930ns	3.386ns	0.013ns	0.139ns	1.248ns
Genotype × Location	120	3.700***	0.183***	0.111ns	115.160***	173.560***	14.194***	0.518***	0.640***	1.707ns
G × Location × Season	120	2.568*	0.025ns	0.130*	17.48	34.960ns	3.246ns	0.013*	0.087ns	0.593ns
Pooled Error	576	2.147	0.058	0.1	66.88	58.61	6.602	0.136	0.194	1.709
Grand Mean		16.251	1.2189	1.08	29.5	69.22	12.876	1.154	1.477	42.752
LSD		2.489	0.437	0.527	14.574	14.31	4.305	0.605	0.726	2.485
CV (%)		9	19.7	29.3	27.7	11.1	20	31.9	29.8	3.1

***=p<0.001; **=p<0.01; *=p<0.05; ns=not significant, 100SDWT=100 seed weight (gm); GLM=groundnut leafminner, LODG= stem lodging; NPODS= number of pods; PH= plant height; PODCL= pod clearance; PODSH= pod shattering; DT50%F= days to 50% flowering

The means for agronomic traits are presented in table 4.5. The results showed that genotypes NG 14.1 × NII-1 and NDGT 8.11× 3N-1 had the highest 100 seed weight of 18.1g followed by Duiker × 3N-5 (17.4g), Maksoy 3N (17.3), BSPS 48A-28 and BSPS 48A-27-1 both had an average weight of 17.0g. NII × 35.3-3 had the lowest 100 seed weight with a mean of 14.4g and the overall mean for 100 seed weight across all the genotypes was 16.3g. NG 14.1 × UG5 had the longest days to 50% flowering with an average of 45 days followed by NII × GC 35.3-1, BSPS 48A-24-1, Maksoy 4N, BSPS 48A-28, and Nam II GC 17.3 all had the mean of 44 days. On the other hand, NGDT 8.1 × 14.16B had the shortest days (40 days which was two days less than the overall mean 42 days) to 50% flowering. The overall mean score for groundnut leafminner damage was 1.2 of which Mak 3N × 1N, Maksoy 4N, BSPS 48A-28, NDGT 8.11× 3N-1 and Duiker × 3N-5 had the lowest mean scores of 1.1 for groundnut leafminner damage. G8586 × UG5 had the highest mean score of 1.5 for groundnut leafminner damage. Duiker × 3N-5 had the highest mean score of 1.4 for stem lodging followed by BSPS 48A-27-1 with a mean score of 1.3 and GC × 2N-1, BSPS 48A-28-1, NII × GC 13.2, Nam II GC 17.3, NII × GC 35.3-2, NII × 35.3-3, G8586 × UG5, NDGT 8.11× 3N-1, BSPS 48A-24-1, NDGT 8.11 × 3N-2, Mak 3N × 1N and NG 14.1 × NII-1 had the lowest mean scores of 1.0 and the overall mean score for stem lodging across 25 genotypes was 1.1. Genotype NII × GC 13.2 had the highest mean number of pods of 35, followed by Duiker × 3N-5 and 2N × GC with 33 mean number of pods and genotype NDGT 8.11× 3N-1 had the lowest mean number of pods (23 pods) which was six less than the mean of all genotypes (29 pods) across six locations and two seasons. Genotype Duiker × 3N-5 had the highest mean of 85cm for plant height followed by Bulindi 24.1A and NG 14.1 × UG5 with plant height of 81cm and 80.5 cm respectively and Nam II GC 17.3 had the lowest mean plant height of 48.3 cm, and the overall mean plant height was 69.2 cm. Genotype NG 14.1 × UG5 had the highest mean height of 16.1 cm for pod clearance followed by Bulindi 18.4B and NDGT 8.11 × 3N-2 both had a mean of 15.1 cm, while Maksoy 3N had the lowest mean height of 10 cm for pod clearance. For pod shattering Nam II GC 17.3 had the highest mean score of 1.9, followed by NII × GC 13.2 with a mean score of 1.5, while NG 14.1 × NII-1, Mak 3N × 1N, BSPS 48A-24-1, BSPS 48A-28, NDGT 8.11× 3N-1, NG 14.1 × UG5, BSPS 48A-25-1, NGDT 8.1 × 14.16B and Duiker × 3N-5 had the lowest mean scores of 1.0 for pod shattering.

Table 4.5: Means for agronomic traits of 25 soybean genotypes evaluated in six locations across two seasons in Uganda (2018A and 2018B)

Genotypes	100SWT	DT50%F	GLM	LODG	NPODS	PH	PODCL	PODSH	RUST
Duiker × 3N-5	17.4	43.0	1.1	1.4	33.0	85.0	14.4	1.0	1.6
GC × 2N-1	15.5	42.0	1.2	1.0	27.0	71.6	12.1	1.1	1.5
BSPS 48A-27-1	17.0	43.0	1.2	1.3	28.0	70.7	14.7	1.1	1.2
BSPS 48A-28-1	16.3	42.0	1.2	1.0	32.0	64.3	10.2	1.3	1.6
NGDT 8.1× 14.16B	16.9	40.0	1.2	1.2	29.0	65.0	10.8	1.0	1.3
NII × GC 13.2	16.6	43.0	1.3	1.0	35.0	68.7	11.2	1.5	1.6
BSPS 48A-25-1	16.8	43.0	1.3	1.1	31.0	75.8	12.5	1.0	1.3
Nam II GC 17.3	15.8	44.0	1.2	1.0	27.0	48.3	13.3	1.9	1.3
NII × GC 35.3-2	15.0	43.0	1.2	1.0	30.0	73.3	13.2	1.1	1.5
NG 14.1 × UG5	16.2	45.0	1.2	1.1	31.0	80.5	16.1	1.0	1.4
Nam 4M × 2N-2	15.8	42.0	1.2	1.1	30.0	67.0	13.0	1.3	1.7
NII × 35.3-3	14.8	43.0	1.3	1.0	30.0	74.4	13.0	1.1	1.8
G8586 × UG5	15.3	43.0	1.5	1.0	29.0	52.9	11.3	1.3	1.6
NDGT 8.11×3N-1	18.1	42.0	1.1	1.0	23.0	68.4	13.3	1.0	1.4
BSPS 48A-28	17.0	44.0	1.1	1.1	32.0	74.3	14.2	1.0	1.1
Bulindi 18.4B	15.2	42.0	1.4	1.1	28.0	62.2	15.1	1.2	1.2
Maksoy 4N	16.9	44.0	1.1	1.1	30.0	72.5	14.4	1.1	1.3
BSPS 48A-24-1	15.4	44.0	1.2	1.0	29.0	71.6	12.1	1.0	1.4
Bulindi 24.1A	16.0	43.0	1.2	1.1	31.0	81.0	13.3	1.1	1.8
NII × GC 35.3-1	15.0	44.0	1.2	1.1	30.0	74.8	13.0	1.1	1.6
NDGT 8.11 × 3N-2	15.8	43.0	1.3	1.0	24.0	61.3	15.1	1.3	1.3
2N × GC	15.2	43.0	1.2	1.2	33.0	67.4	13.1	1.1	1.6
Mak 3N × 1N	16.7	41.0	1.1	1.0	29.0	65.8	10.9	1.0	1.7
NG 14.1 × NII-1	18.1	42.0	1.2	1.0	24.0	66.8	12.1	1.0	1.4
Maksoy 3N	17.3	41.0	1.2	1.1	31.0	66.7	10.0	1.1	1.9
Mean	16.3	42.0	1.2	1.1	29.0	69.2	12.9	1.2	1.5
CV (%)	9.0	3.1	19.7	29.3	27.7	11.1	20.0	31.9	29.8
LSD	2.5	2.5	0.4	0.5	14.6	14.3	4.3	0.6	0.7

100SWT=100 seed weight (gm); GLM=groundnut leafminner (scores), LODG= stem lodging (scores); NPODS= number of pods; PH= plant height (cm); PODCL= pod clearance (cm); PODSH= pod shattering (scores); DT50%F= days to 50% flowering, Rust (scores)

The results of mean % protein content are presented in table 4.6. The results showed that location Abi had the average protein content of 33.95%, with genotype NG 14.1× UG5 had the highest protein content of 36.24% , while Nam 4M× 2N-2 had the lowest protein content of 31.51%. The average protein content for Bulindi was 34.29% with genotype Duiker 3N-5 had the highest protein content of 39.96%, whereas NGDT 8.11 × 14.16B had the lowest mean of 31.51%. Iki-Iki had the mean protein content of 32.35% with genotype NG 14.1 × NII-1 had the highest protein content of 34.41%, while NDGT 8.11× 3N-1 had the lowest protein content of 31.41%. Nakabango had the mean protein content of 33.57% with genotype BSPS 48A-28 had the highest protein content of 35.06%, whereas Nam 4M × 2N-2 had the lowest protein content of 31.92%. The overall mean for protein content across seasons and selected locations was 33.54%, with genotypes 2N × GC, G8586 × UG5, Bulindi 24.1A, BSPS 48A-28-1 and Duiker × 3N-5 ranking the best five with the following protein content (%) of 34.67, 34.62, 34.45, 34.4, 34.3, respectively.

The results of mean % oil content analysis are presented in table 4.7. In Nakabango, the average oil content (%) was 16.79 with genotype BSPS 48A-28 having the highest oil content of 17.82% followed by NII × GC 35.3-1 (17.67%) 2N × GC (17.64%), whereas genotype BSPS 48A-25-1 had the lowest protein content of 13.79%. Location Abi had the mean oil content of 15.61% with genotype Duiker × 3N-5 having the highest oil content of 18.45%, while genotype NII × 35.3-3 having the lowest oil content of 13.29%. Location Bulindi had the average oil content of 16.71% with genotype G8586 × UG5 had the highest mean oil content of 19.16% followed by Duiker × 3N-5 (18.81%), NDGT 8.11× 3N-1 (18.44%) ,while genotype NII × 35.3-3 had the lowest mean oil content of 14.81%. Iki-Iki had the average oil content of 14.92% with genotype GC × 2N-1 had the highest oil content of 16.95%, whereas NG 14.1 × UG5 had the lowest oil of 12.76%. The overall mean for oil content across seasons and selected locations was 16.01%, with genotypes Duiker × 3N-5, NDGT 8.11× 3N-1, NGDT 8.11 × 14.16B, NII × GC 13.2 and Nam II GC 17.3 were ranked the best five with oil content (%) of 17.26, 16.62, 16.55, 16.52 and 16.43, respectively.

Table 4.6: Protein content (%) of 25 soybean genotypes evaluated in selected four locations across two seasons in Uganda (2018A and 2018B)

Genotype	Location					
	Abi	Bulindi	Iki-Iki	Nakabango	Mean	Rank
2N × GC	35.67	35.40	32.83	34.76	34.67	1
G8586 × UG5	35.00	36.36	33.24	33.87	34.62	2
Bulindi 24.1A	35.12	35.59	32.12	34.97	34.45	3
BSPS 48A-28-1	35.35	34.76	32.97	34.51	34.40	4
Duiker × 3N-5	35.57	36.96	31.51	33.16	34.30	5
BSPS 48A-27-1	34.28	35.28	34.04	32.81	34.10	6
NII × 35.3-3	32.62	36.22	31.97	34.83	33.91	7
Nam II GC 17.3	35.55	34.75	32.28	32.95	33.88	8
NG 14.1 × NII-1	33.61	33.13	34.41	34.16	33.83	9
Mak 3N × 1N	31.75	35.42	33.24	34.88	33.82	10
NG 14.1 × UG5	36.24	34.13	32.22	32.55	33.78	11
NDGT 8.11 × 3N-2	34.83	32.59	32.99	34.52	33.73	12
NII × GC 35.3-2	34.60	34.34	30.85	34.34	33.53	13
BSPS 48A-25-1	34.62	33.87	30.67	34.34	33.38	14
Maksoy 3N	34.03	34.20	32.55	32.59	33.34	15
Bulindi 18.4B	35.34	33.01	31.71	33.05	33.28	16
NGDT 8.11 × 14.16B	34.00	31.51	33.22	34.32	33.26	17
GC × 2N-1	34.14	34.25	31.83	32.51	33.19	18
BSPS 48A-24-1	33.97	34.24	31.97	32.20	33.10	19
NII × GC 35.3-1	31.54	35.27	32.68	32.79	33.07	20
BSPS 48A-28	32.27	32.50	31.87	35.06	32.93	21
Maksoy 4N	32.34	33.57	32.24	33.24	32.85	22
Nam 4M × 2N-2	31.51	35.33	32.29	31.92	32.76	23
NDGT 8.11× 3N-1	32.75	32.81	31.41	32.19	32.29	24
NII × GC 13.2	32.08	31.85	31.57	32.80	32.07	25
Mean	33.95	34.29	32.35	33.57	33.54	
CV	9.20	8.70	6.10	8.20	8.30	
LSD	6.28	5.98	3.98	5.51	5.39	

Table 4.7: Oil content (%) of 25 soybean genotypes evaluated in selected four locations across two seasons in Uganda (2018A and 2018B)

Genotype	Location					
	Abi	Bulindi	Iki-Iki	Nakabango	Mean	Rank
Duiker × 3N-5	18.45	18.81	14.68	17.12	17.26	1
NDGT 8.11× 3N-1	14.92	18.44	15.93	17.20	16.62	2
NGDT 8.11 × 14.16B	15.27	17.05	16.88	17.02	16.55	3
NII × GC 13.2	16.38	17.47	15.57	16.66	16.52	4
Nam II GC 17.3	15.34	17.74	15.86	16.78	16.43	5
GC × 2N-1	15.61	16.60	16.95	16.53	16.42	6
G8586 × UG5	15.81	19.16	14.86	15.45	16.32	7
NDGT 8.11 × 3N-2	13.94	17.82	15.46	17.94	16.29	8
Mak 3N × 1N	16.56	18.05	14.35	16.01	16.24	9
NII × GC 35.3-1	16.09	16.05	14.71	17.67	16.13	10
Maksoy 4N	16.76	16.09	14.18	17.39	16.10	11
2N × GC	14.92	15.82	15.66	17.64	16.01	12
Bulindi 18.4B	15.99	16.57	14.71	16.59	15.96	13
Maksoy 3N	15.61	16.69	14.91	16.29	15.87	14
NII × GC 35.3-2	17.21	16.53	13.94	15.51	15.80	15
NG 14.1 × NII-1	16.04	15.48	15.35	16.34	15.80	16
BSPS 48A-25-1	16.02	17.96	15.32	13.79	15.77	17
BSPS 48A-28	14.79	15.78	14.62	17.82	15.76	18
BSPS 48A-24-1	14.78	16.92	14.26	16.80	15.69	19
BSPS 48A-28-1	15.14	14.85	15.59	17.13	15.68	20
Nam 4M × 2N-2	15.38	15.61	13.97	17.56	15.63	21
BSPS 48A-27-1	14.86	16.29	14.17	17.07	15.60	22
Bulindi 24.1A	15.75	15.18	13.92	17.04	15.47	23
NG 14.1 × UG5	15.45	15.93	12.76	17.10	15.31	24
NII × 35.3-3	13.29	14.81	14.35	17.44	14.97	25
Mean	15.61	16.71	14.92	16.79	16.01	
LSD	4.57	5.86	3.89	3.03	4.35	
CV (%)	14.80	16.00	13.20	8.40	13.40	

The results of combined grain yield means of 25 soybean genotypes evaluated for two seasons across six locations are summarized in table 4.8. Genotype BSPS 48A-28 had the highest yield of 1767 kg¯ha followed by Maksoy 3N and Mak 3N × 1N both with average grain yield of 1725 kg¯ha, while NDGT 8.11× 3N-1 had the lowest grain yield mean of 1385 kg¯ha (Table 4.8). For locations, Bulindi had the highest mean yield (2650 kg¯ha) followed by Abi (1845 kg¯ha), Nakabango (1698 kg¯ha), Ngetta (1567 kg¯ha) and Kabanyolo (1017 kg¯ha) while Iki-Iki had the lowest mean yield of 889 kg¯ha.

Table 4.8: Grain yield performance in kg/ha of 25 soybean genotypes evaluated in six locations across two seasons in Uganda (2018A and 2018B)

Genotypes	Location							
	Abi	Bulindi	Iki-Iki	Kabanyolo	Nakabango	Ngetta	Mean Yield	Rank
BSPS 48A-28	1683	3006	843	1165	2069	1836	1767	1
Mak 3N × 1N	1809	2773	1073	1317	1841	1538	1725	2
Maksoy 3N	2001	2642	817	1041	1937	1912	1725	3
2N × GC	1578	2739	988	1346	1932	1678	1710	4
NDGT 8.11 × 3N-2	1942	2592	850	1189	2021	1621	1702	5
BSPS 48A-27-1	2139	2844	1092	924	1717	1369	1681	6
BSPS 48A-25-1	1696	2686	1027	1181	1709	1766	1678	7
BSPS 48A-24-1	2194	2729	747	1038	2002	1321	1672	8
Maksoy 4N	1926	3036	630	983	1926	1526	1671	9
NGDT 8.11×14.16B	1805	2540	986	1143	1598	1838	1652	10
Bulindi 18.4B	1380	2938	926	1260	1867	1515	1648	11
NII × GC 35.3-1	1915	3030	674	1064	1623	1492	1633	12
Nam II GC 17.3	1694	2376	1014	1189	1861	1610	1624	13
Duiker × 3N-5	2112	2712	937	883	1491	1519	1609	14
G8586 × UG5	1928	2578	978	861	1716	1578	1606	15
NII × 35.3-3	1757	2652	963	973	1635	1563	1590	16
NII × GC 35.3-2	1978	2578	943	1064	1303	1643	1585	17
Bulindi 24.1A	2016	2631	446	1040	1851	1448	1572	18
Nam 4M × 2N-2	1935	2617	799	1111	1359	1438	1543	19
BSPS 48A-28-1	1790	2313	1034	849	1513	1735	1539	20
NG 14.1 × NII-1	1648	2561	1049	793	1670	1464	1531	21
NG 14.1 × UG5	2121	2362	869	733	1440	1419	1491	22
GC × 2N-1	1861	2392	802	865	1447	1447	1469	23
NII × GC 13.2	1616	2520	876	759	1462	1579	1469	24
NDGT 8.11× 3N-1	1603	2417	858	662	1453	1319	1385	25
Mean	1845	2650	889	1017	1698	1567	1611	
CV (%)	25.4	20.5	26.8	31.4	19.1	17.6		
LSD	538	621.5	272.6	365.8	371.1	316		

The analysis of variance generated by AMMI for grain yield is summarized on table 4.9. The results of the AMMI model partitioned main effects into genotypes and environments. The AMMI analysis of variance for soybean grain yield (kg ha^{-1}) of 25 genotypes tested in six environments for two seasons showed that 42.6% of the total sum of squares was attributed to environmental effects ($p<0.05$)., while 1.1% to genotypic effects ($p<0.05$) and 3.3% to genotype × environment interaction effects ($p<0.05$). The G × E interaction was partitioned among the first two interaction principal component axis (IPCA) and were significant ($p<0.05$). The IPCA1 and IPCA2 managed to explain 68% of the total G × E interaction sum of squares in which 37.1% and 30.9% was accounted by IPCA1 and IPCA2 respectively.

Table 4.9: AMMI analysis of variance for 25 soybean genotypes tested over six locations across two seasons in Uganda (2018A and 2018B)

				Sum of squares explained		
Source of Variation	DF	SS	MS	%Total	% G×E	% G × E Cumulative
Total	899	710266583	790063			
Genotypes	24	7682457	320102***	1.1		
Environments	5	302816025	60563205***	42.6		
G×E Interaction	120	23509931	195916**	3.3		
IPCA 1	28	8733426	311908***	1.2	37.1	
IPCA 2	26	7253410	278977***	1	30.9	68
Residuals	66	7523095	113986	1.1		

***=p<0.001, **=p<0.01, *=p<0.05, ns=not significant, G × E = Genotype × environment interaction, IPCA 1, IPCA 2 = Interaction principal component axis one and two respectively

The results of the AMMI biplot analysis are presented on Figure 4.1. The AMMI biplot provided a visual expression of the relationships between the first interaction principal component axis (IPCA1) and means of genotypes and environments. The biplot showed that showed that the tested six environments were scattered without any definite grouping, with most of the genotypes clustered around the midpoint. AMMI biplot of IPCA1 scores versus grain yield means explained 37.1% of the total G × E interaction sum of squares. The genotype codes and names are the same as the ones presented on table 4.10.

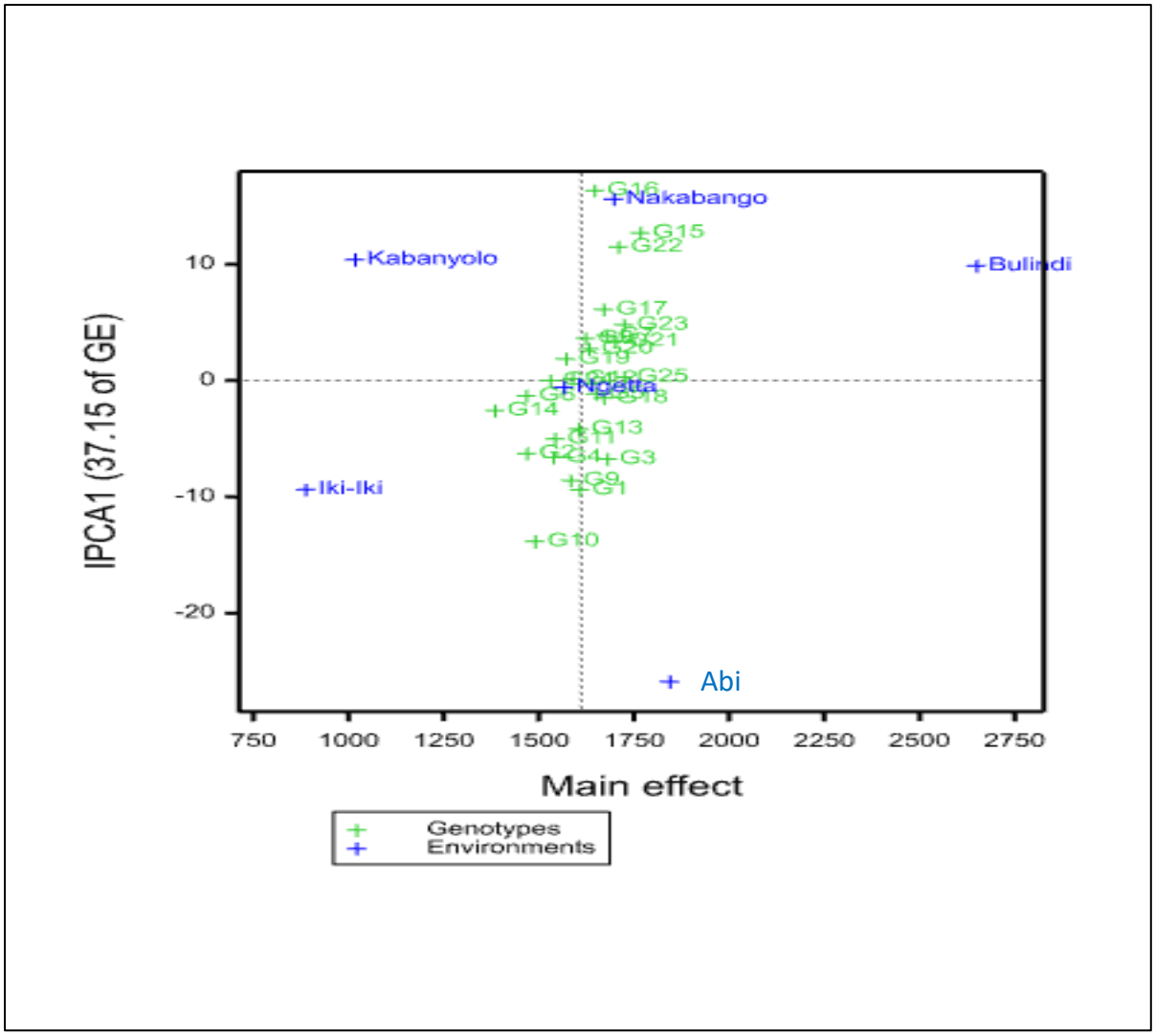

Figure 4.1: AMMI biplot of IPCA1 scores versus yield means for 25 soybean genotypes and six environments for 2018A and 2018B seasons

The results of AMMI mean grain yield values of 25 soybean genotypes averaged over six environments and two seasons are presented in table 4.10. The results showed that genotype BSPS 48A-28 had the highest yield of 1767 kg ha^{-1} and NDGT 8.11× 3N-1 had the lowest mean yield of 1385 kg ha^{-1}. Based on IPCA1, BSPS 48A-28 (G15), Bulindi 18.4B (G16) and 2N × GC (G22) showed highly positive interactions with the environment, while NG 14.1 × UG5 (G10), Duiker × 3N-5 (G1) and NII × GC 35.3-2 (G9) showed highly negative interactions with the environment. Genotypes NII × 35.3-3 (G12), NG 14.1 × NII-1 (G24) and Maksoy 3N (G25) had the lowest positive IPCA1 scores whereas BSPS 48A-24-1 (G18), NGDT 8.11 × 14.16B (G5) and NII × GC 13.2 (G6) had the lowest negative IPCA1 scores. Maksoy 3N (G25) (a commercial variety used as a check) had the ASV close to zero. On the basis of environmental index value in terms of negative and positive, Bulindi, Nakabango and

Kabanyolo were good environments, whereas, Ngetta, Iki-Iki and Abi were poor environments.

Table 4.10: AMMI mean yield and AMMI stability value (ASV) for the 25 soybean genotypes grown in six locations in Uganda during 2018A and 2018B seasons

Genotype codes	Genotype	Mean Yield	IPCA1	IPCA2	ASV	Rank
G15	BSPS 48A-28	1767	12.687	-1.615	15.360	1
G23	Mak 3N × 1N	1725	4.778	1.908	6.061	2
G25	Maksoy 3N	1725	0.288	-0.679	0.762	3
G22	2N × GC	1710	11.472	5.381	14.824	4
G21	NDGT 8.11 × 3N-2	1702	3.363	-2.451	4.734	5
G3	BSPS 48A-27-1	1681	-6.757	-6.020	10.121	6
G7	BSPS 48A-25-1	1678	3.881	7.216	8.597	7
G18	BSPS 48A-24-1	1672	-1.510	-13.859	13.977	8
G17	Maksoy 4N	1671	6.091	-12.857	14.802	9
G5	NGDT8.11×14.16B	1652	-1.152	8.327	8.442	10
G16	Bulindi 18.4B	1648	16.330	2.782	19.858	11
G20	NII × GC 35.3-1	1633	2.730	-9.766	10.304	12
G8	Nam II GC 17.3	1624	3.594	7.756	8.881	13
G1	Duiker × 3N-5	1609	-9.400	-3.311	11.792	14
G13	G8586 × UG5	1606	-4.166	0.741	5.070	15
G12	NII × 35.3-3	1590	0.141	2.830	2.835	16
G9	NII × GC 35.3-2	1585	-8.593	3.985	11.087	17
G19	Bulindi 24.1A	1572	1.856	-12.703	12.898	18
G11	Nam 4M × 2N-2	1543	-5.013	-1.325	6.180	19
G4	BSPS 48A-28-1	1539	-6.613	10.467	13.151	20
G24	NG 14.1 × NII-1	1531	0.011	4.915	4.915	21
G10	NG 14.1 × UG5	1491	-13.839	-1.798	16.760	22
G2	GC × 2N-1	1469	-6.278	1.182	7.651	23
G6	NII × GC 13.2	1469	-1.320	5.694	5.912	24
G14	NDGT 8.11× 3N-1	1385	-2.583	3.200	4.462	25
	Environment	Mean	IPCA1	IPCA2		
	Abi	1845	-25.885	-15.904		
	Bulindi	2650	9.852	-13.533		
	Iki-Iki	889	-9.373	19.591		
	Kabanyolo	1017	10.39	3.842		
	Nakabango	1698	15.595	-8.108		
	Ngetta	1567	-0.58	14.111		

The results of which-won-where biplot showed different winning genotypes in different environments are presented in Figure 4.2. The IPCA1 and IPCA2 explained 41.98% and 23.76% respectively and together accounted for 64.74% of the genotype main effect and G × E interaction for grain yield of the genotypes evaluated in 12 environments (evaluated in six locations for two seasons). The biplot was divided into seven sectors and four mega-environments and showed six vertex genotypes. The biplot identified winning genotypes in each mega-environment as follows; BSPS 48A-28 for mega-environment I (Bulindi, Nakabango and Kabanyolo), BSPS 48A-28-1 for mega-environment II (Iki-Iki), Bulindi 18.4B for mega-environment III (Ngetta) and BSPS 48A-24-1 for mega-environment IV (Abi). Genotypes within the polygon were less responsive than the vertex genotypes. The genotype codes and names on figure 4.2 are the same as the ones presented on table 4.10.

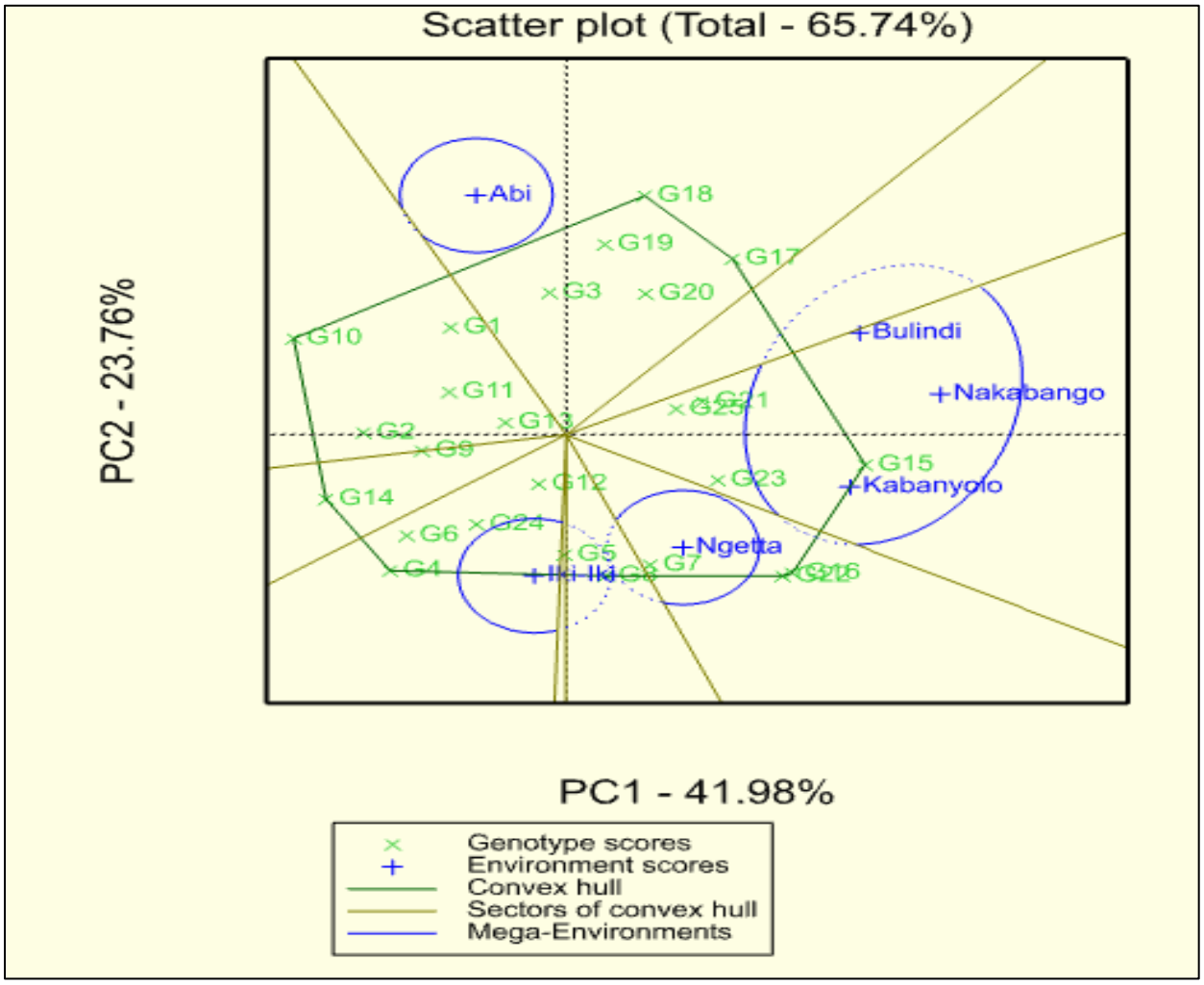

Figure 4.2: The which-won-where and mega-environment delineation biplot for the 25 soybean genotypes evaluated in six locations for two seasons (2018A and 2018B)

The results of genotype focused comparison biplot are presented in Figure 4.3. The results of the IPCA1 and IPCA2 explained 41.98% and 23.76% of the genotype main effect plus G × E sum of squares, respectively, while 35.26% was attributed to noise. Genotypes that were found at the center of the concentric circles were the ideal soybean genotypes (highest yielding and stable). The genotype comparison biplot showed G15 as the most stable and desirable genotype, whereas, G9 was the most stable and undesirable genotype. The biplot ranked genotypes based on both mean grain yield and stability performance in order to identify the highest yielding and stable genotypes. Based on mean yield performance and stability, the comparison biplot ranked G15>G16>G22>G17>G21, as ideal genotypes. The genotype codes on figure 4.3 are the same codes and names as presented on table 4.10.

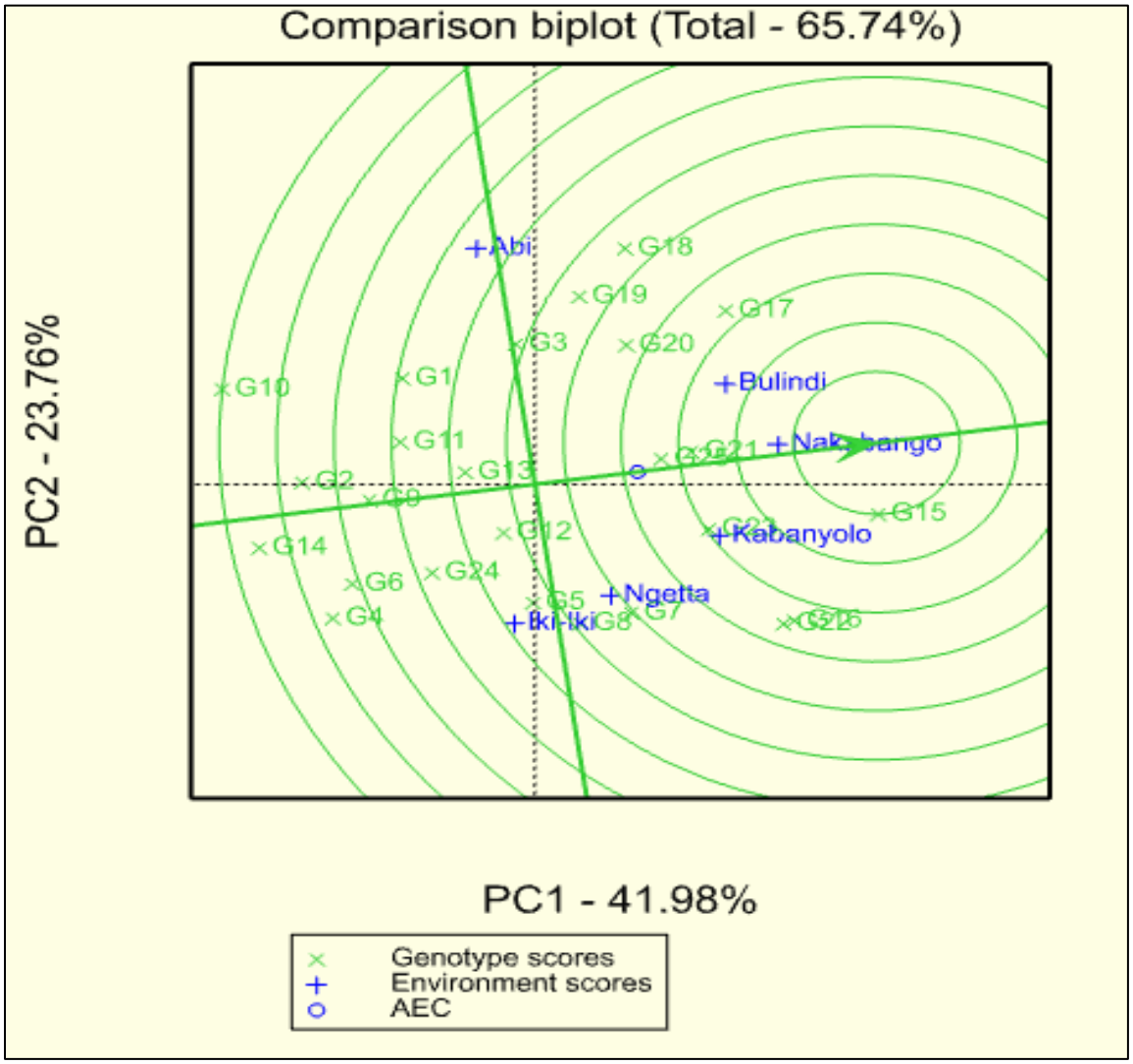

Figure 4.3: Genotype focused comparison biplot showing the best genotypes based on mean performance and stability

The results of genotype ranking plot are presented in Figure 4.4. The results of the IPCA1 and IPCA2 explained 41.98% and 23.76% of the genotype main effect plus G × E sum of squares, respectively, while 35.26% was attributed to noise. The biplot ranked genotypes based their mean yield and stability. The biplot shows that G15>G16>G22>G17>G21, the genotype codes on figure 4.4 are the same codes and names as presented on table 4.10.

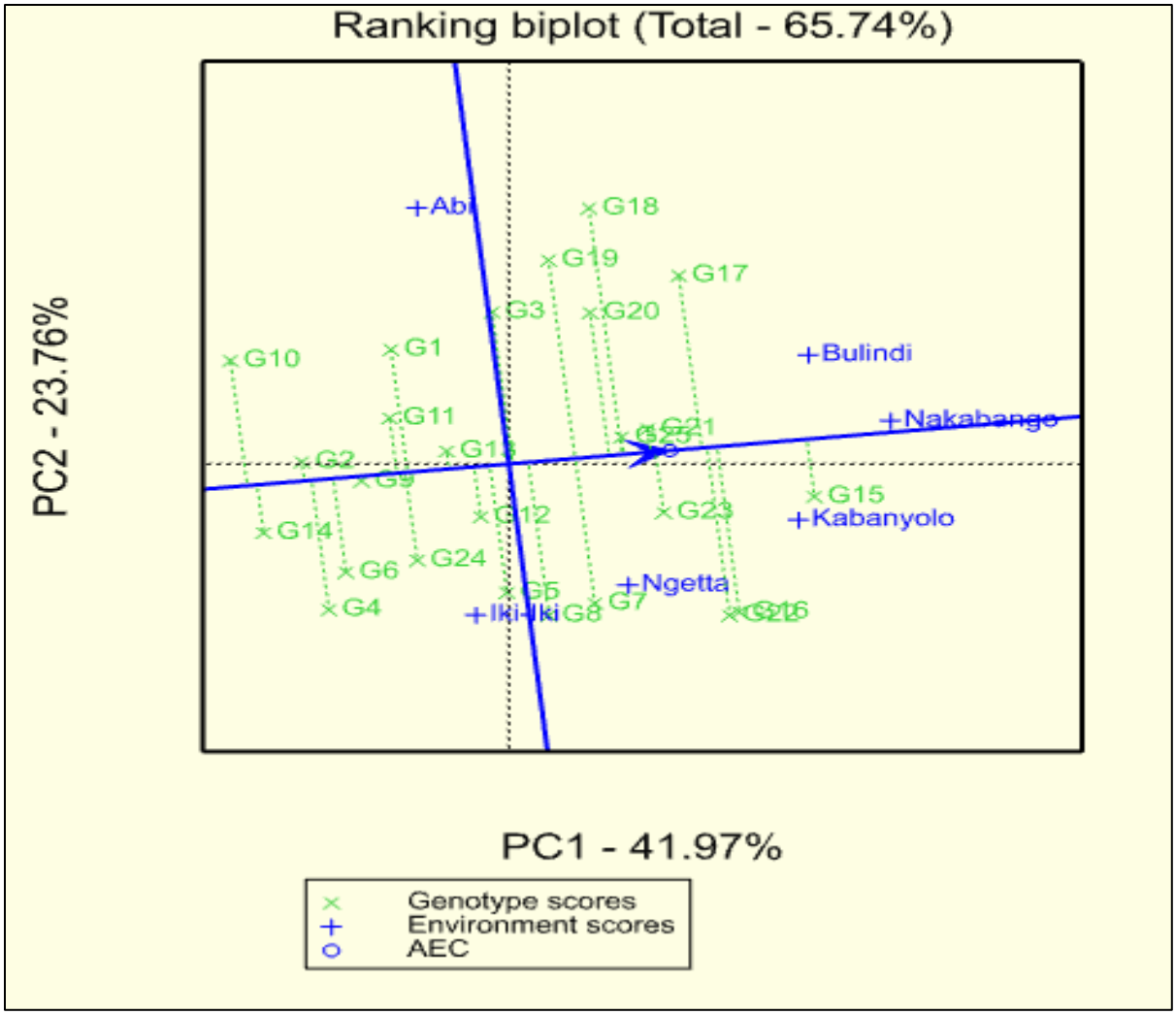

Figure 4.4: Ranking plot showing the best genotypes based on mean performance and stability

The results of environment focused comparison biplot are presented in Figure 4.5. The results showed that the ideal test location which was located near the center of the concentric circles was the most representative. The biplot showed Nakabango as the most representative testing location, while other test locations, Bulindi, Kabanyolo, Ngetta, Iki-Iki and Abi were not representative. The genotype codes on figure 4.5 are the same codes and names as presented on table 4.10.

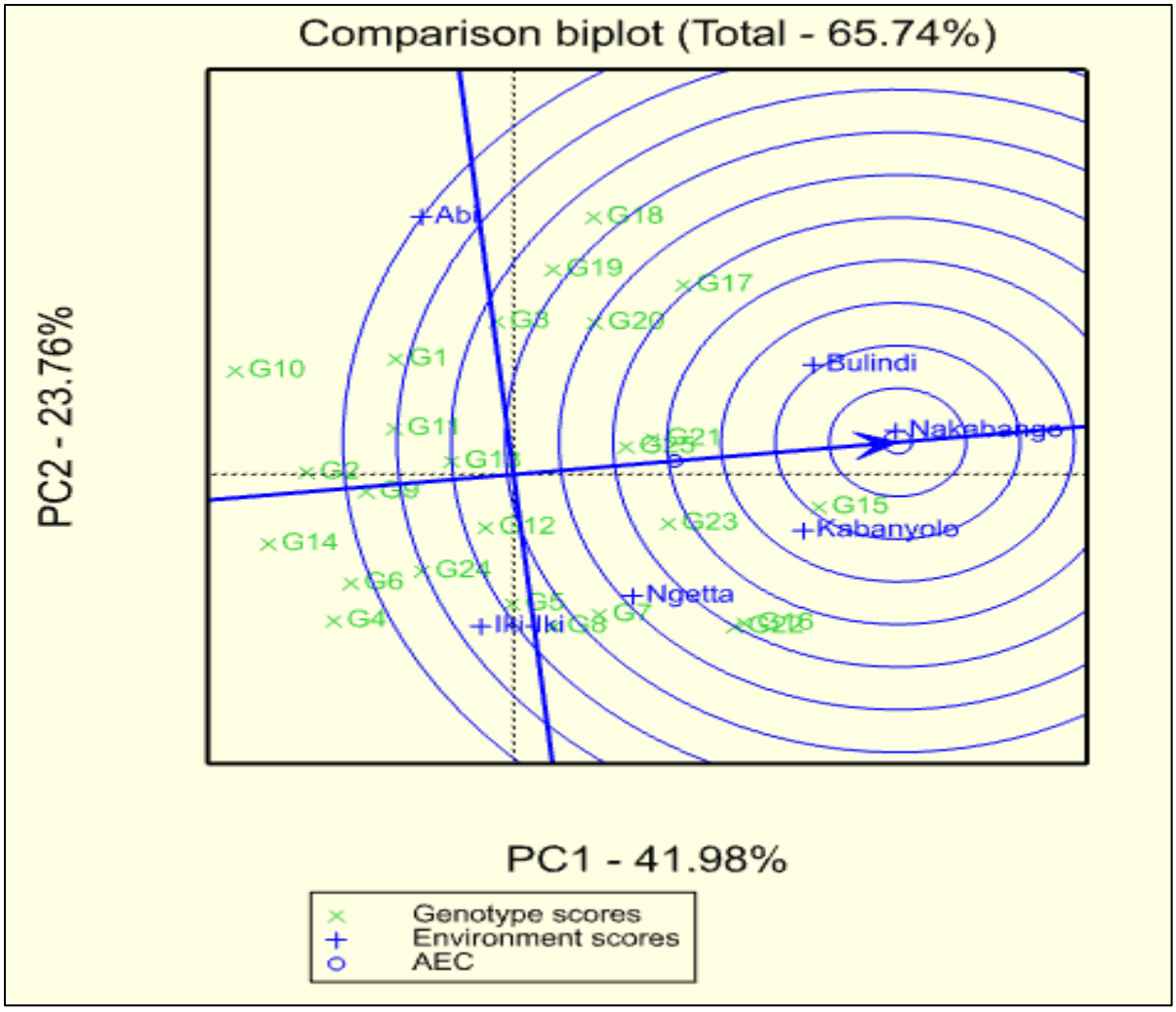

Figure 4.5: Environment focused comparison biplot showing the ideal testing location for soybean among the locations used in evaluations

The results of environment vector plot which shows the most discriminating test location are presented in Figure 4.6. The results showed that Abi, Nakabango and Bulindi had the longest vectors from the biplot origin. The angle between Abi and Bulindi is almost right angle and locations Ngetta and Iki-Iki had the shortest vectors from the biplot origin as well as the small angle between them. Abi, Nakabango and Bulindi were the most discriminating locations, while Ngetta and Iki-Iki were the least discriminating test locations. The genotype codes on figure 4.4 are the same codes and names as presented on table 4.10.

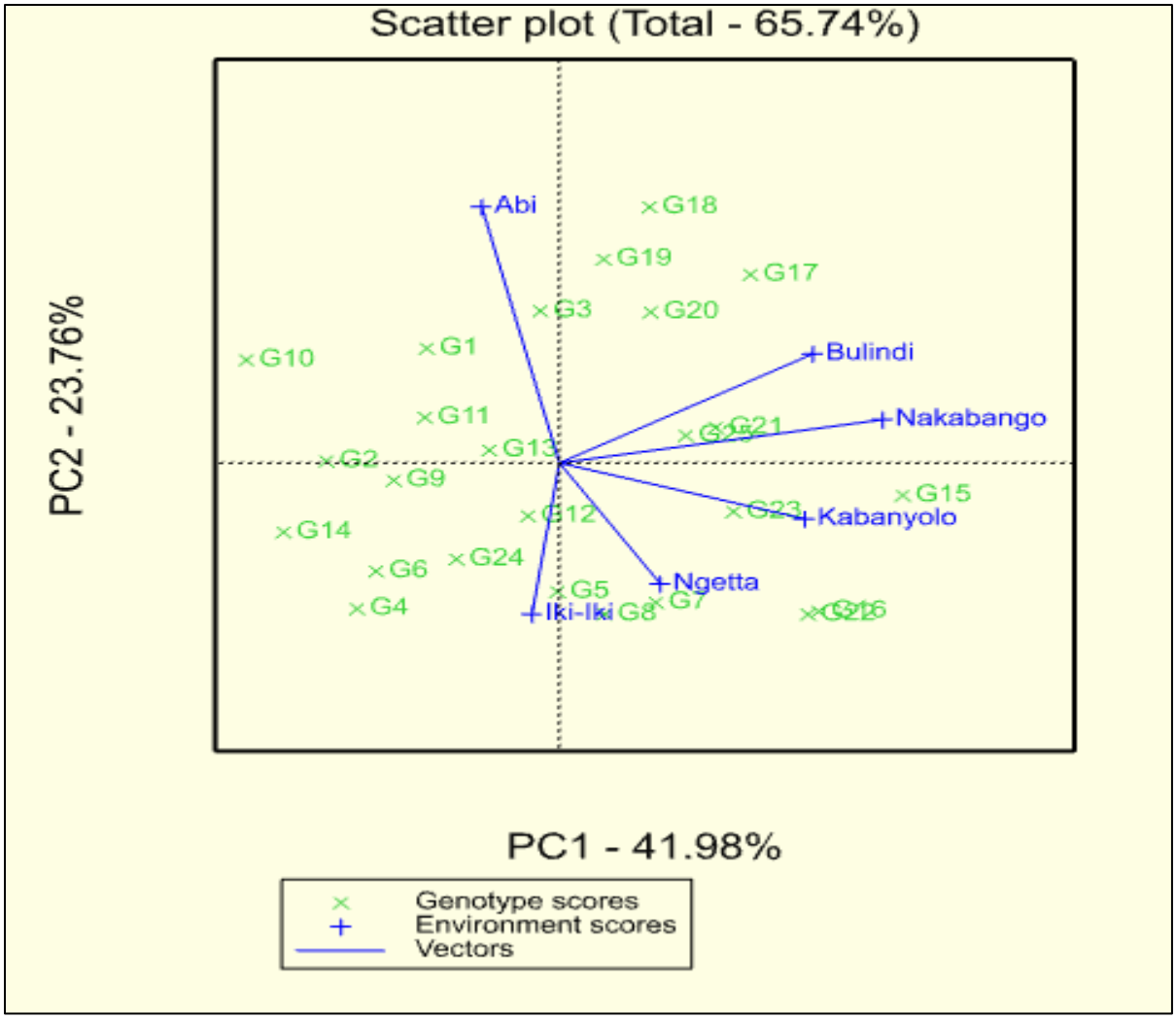

Figure 4.6: Environment vector plot showing discriminating ability of test locations used for soybean evaluations

4.6 Discussion

Nature of the G × E interaction and broad sense heritability

The presence of significant G × E interaction suggested differential responses of soybean genotypes across tested environments and implied the need to identify high-yielding and stable genotypes across the test environments. The large variance component due to environment alone justified the need to use both AMMI and GGE biplots, in which the GGE biplot captured much of the variation due to genotype plus G × E interaction as a fraction of the total variation (G + E + GE) (Yan *et al*. 2007). The large variance component due to locations and seasons depicted the locations used in this study as very diverse across seasons. Indeed, Uganda's climatic conditions are highly variable with mean annual rainfall of 510-2160 mm, also varied soil depth, texture,

acidity and organic matter (Wortman and Eledu, 1999). In these agro-ecological zones there was large amount of variability in both predictable factors (soil characteristics) and unpredictable factors (temperature and rainfall) (Table 4.1). Therefore, widely adapted soybean genotypes with dynamic yield stability are recommended to sustain soybean production country wide.

The large G × E interaction and error variance components found in the present study could reduce selection progress by complicating the identification and recommendation of superior genotypes for a target environment (Gasura *et al*., 2015). The results observed in this study, however, were inconsistence with Bhartiya *et al*. (2017) who reported G × E interaction almost doubled the genotypic main effects and five times larger than environmental effects. Large G × E interaction and residuals observed in multi-environment trials (METs) affect the repeatability of the experiment (Simion *et al*., 2018). Furthermore, the large G × E effects showed that the performance of soybean genotypes were different across different locations (Etnaf *et al*., 2013). The large G × E interaction and error variance components observed in this study might have contributed in low broad sense coefficient of genetic determination (which is equivalent to broad sense heritability based on fixed genotypes) (3%) on a single plot basis, but broad sense heritability across environments improved (41%) as the number of locations and seasons increased. Similar results were reported by Gasura *et al*. (2015) where broad sense heritability increased from 2.8% on single plot basis to 31.8% on across environments basis. Sousa *et al*. (2018); Gasura *et al*. (2015); Mare *et al.* (2017) suggested that large G × E interaction and error variance components raised up the cost of variety evaluation due to increase in numbers of replications, locations and seasons needed to improve broad sense heritability, and hence the selection efficiency. When G × E interaction is highly significant it should not be ignored, rather different methods should be applied and the G × E interaction should be exploited (Yan and Kang, 2002). In this study, most of the G × E interaction variation observed could be due to differences in location-intrinsic abiotic factors (predictable factors) such as soil nutritional status, rainfall patterns, and temperatures as well as biotic factors like outbreak of various pests and diseases found in specific locations (Ibanda *et al*., 2018). While, the aforementioned predictable factors could be easily managed, but modifying the growing environments to perfectly suit the crop growth conditions would further constrain the resource-limited farmers (Krisnawati and Adie, 2018). Therefore, the most inexpensive and farmer-friendly method is to use soybean varieties adapted to the target environments (Tukamuhabwa *et al*., 2012). Meanwhile, since crop

growing locations have no clearly defined demarcations and most farmers tend to influence each other in the choice of variety that is grown, the development of varieties adapted to a wide range of environments is strongly supported, rather than environment-specific varieties (Gasura *et al*., 2015; Bhartiya *et al*., 2017). Yan and Rajcan (2003) noted that a single year with several locations was good enough for predicting the performance, and recommending soybean cultivars for specific locations. Therefore, the presence of significance genotype × location × season interactions can be ignored since it makes it difficult for a breeder to recommend soybean cultivars based on seasons.

Evaluation of soybean genotypes across environments

The significant difference for grain yield and yield related traits observed among genotypes across environments indicated the presence of genetic and environmental causes of variation. The significant G × E interaction observed in this study also showed the significance of environmental effects in the expression soybean for grain yield. The results were in agreement with the previous studies of Bhartiya *et al*. (2017); Hampango *et al*. (2017) who reported significant G × E interaction for soybean grain yield. The absence of significant genotype, G × E interaction for protein and oil content observed in this study was inconsistent with previous studies by Gurmu *et al*. (2009) who reported the presence of significance genotype, G × E interaction for protein and oil content. This would imply that testing the set of genotypes at one location would be enough, hence saves both time and resources. From the results obtained, it showed that the performances of protein and oil content among genotypes were far below the standard as reported by Tukamuhabwa and Obua (2015) who reported 36% and 38% protein content for Maksoy 3N and Maksoy 4N, as well as 22% and 21% oil content respectively. The reason might be due to limited genetic variation among genotypes in terms of protein and oil content. Also, another reason might be due to the influence of environmental conditions since these two traits are highly sensitive to changes in abiotic conditions like soil nutrient status, temperature and rainfall amount received during the growing season.

The AMMI ANOVA results showed that the genotype main effect was smaller than the effect of the environment and G × E interaction. The dominant contribution of environment and G × E interaction effects for grain yield indicated that environment had greater effect on yield. Earlier studies by Gurmu *et al.*(2017) reported similar results. Yan and Kang (2002) suggested that,

when the G × E interaction is larger than the genotype main effect, it should not be ignored but rather exploited. In this study, the AMMI and GGE biplots explained 68% and 65.74% of the total interaction sum of squares, respectively and the G × E interaction explained only 3.3% of the total variance. Yan and Tinker (2006) pointed out the possibility of such complex situations which were usually common in MET data but, regardless of such complex situations, most G × E interaction patterns in the two-way data could be explained by the first and second interaction principal component.

Based on scatter biplot for mega-environments delineation, only four mega-environments with their winning genotypes located at the vertices of the polygon were identified. Locations Kabanyolo, Bulindi and Nakabango were classified on mega-environment I, in which BSPS 48A-28 (G15) was the winning genotype. Mega-environment II had Iki-Iki with BSPS 48A-28-1 as the winning genotype, Ngetta was classified on mega-environment III where genotype Bulindi 18.4B (G16) was the most adapted. Mega-environment IV had Abi found in the West Nile region where BSPS 48A-24-1 (G18) was the winning genotype. This indicated that Uganda had broad agro-ecological regions with unique environmental characteristics with specific high yielding genotypes. Similar results were reported by Tukamuhabwa *et al*. (2012) who found that Nakabango, Bulindi, Kabanyolo and Namulonge are on the same mega-environment (mega-environment I) whereas; Iki-Iki was found on its own mega-environment. Ngetta which was previously classified on mega-environment I, however, in this study, it was found on mega-environment III. Location Bulindi had the highest mean yield of 2650 kg/ha, while Iki-Iki had the lowest mean yield of 889 kg/ha. The reason is Bulindi received high rainfall (1700 mm/ annum) and the site has good soil types, with good nutritional status and water holding capacity (Table 4.1). The reason for low yielding in Iki-Iki might be the gradual changes in biotic and abiotic factors from time to time. On the other hand, Iki-Iki is characterised by poor sandy soils, with low water holding capacity (Tukamuhabwa *et al*., 2012). Also Iki-Iki is a hot spot for groundnut leaf miner, a new soybean pest which is devastating soybean in Uganda (Ibanda *et al.*, 2018). Despite the relatively low yield potential for soybean in Iki-Iki genotype BSPS 48A-28 (G15) managed to maintain its average performance implying that genotype this genotype had good dynamic stability. This is a good attribute for any commercial variety given the unpredictable nature of rainfall patterns in most parts of the country (Obua, 2013). The existence of crossover G × E interaction in this study indicated that genotypes evaluation and

recommendation typically based on any single location is unreliable because there is differential response of genotypes across locations (Mare *et al*., 2017). The presence of crossover interactions indicated the need to reduce or exploit G × E interaction (Bernado, 2002). When the crossover G × E interaction is non-repeatable over years, or seasons, Yan and Kang (2002) recommended that genotype evaluation should be based on mean performance and stability.

The genotype focused comparison biplot indicated that the most stable and high-yielding genotype was BSPS 48A-28 (G15) followed by NDGT 8.11 × 3N-2 (G21), Mak 3N × 1N (G23) and Maksoy 3N (G25). The reason might be that, BSPS 48A-28 (G15) had lowest groundnut leaf miner damage, rust scores, high number of pods, no pod shattering, late maturing and no lodging. It has almost all desirable attributes of a good cultivar. Based on its mean stability and high yield, genotype BSPS 48A-28 maintained its above average performance in the majority of the environments. Genotype Mak 3N × 1N (G23) was comparable in yield performance to the commercial variety Maksoy 3N which could be that this genotype was derived from a cross including Maksoy 3N. This Mak 3N × 1N (G23) genotype together with the top two genotypes BSPS 48A-28 (G15) and NDGT 8.11 × 3N-2 (G21), could be selected for national performance variety trials. Based on mean yield performance and stability, a commercial variety Maksoy 4N performed well although it was ranked fourth (since Maksoy 3N (G25) and Mak 3N × 1N (G23) had the same mean performance), outperformed by three experimental genotypes and Maksoy 3N a commercial variety. Genotype BSPS 48A-28 (G15) is a potential candidate for release since the variety release condition in Uganda advocate for broad instead of specific adaptation. Genotype like BSPS 48A-28 (G15) which performed well in specific locations could be targeted to those locations to maximize grain yield.

Evaluation of the test environments

The presence of G × E interaction for soybean yield justifies undertaking METs during cultivar selection and recommendations, however, this could highly increase the cost of evaluation but evaluation of the test environments would give the breeder a better understanding of the testing locations and probably help to minimize the cost of genotype evaluations (Mare *et al*., 2017; Gasura *et al*., 2015). Based on test location biplot, the vector length of the biplot approximates the standard deviation within each location, a measure of the discriminating ability of the location (Yan and Tinker, 2006). Nakabango, Bulindi and and Abi locations, which had the

longest vectors from the biplot origin, were the most discriminating testing locations and, therefore these three testing locations could be used jointly as discriminating locations for testing early generation breeding materials (Yan *et al*., 2007; Yan and Tinker, 2006). Tukamuhabwa *et al*. (2012) and Obua (2013) reported similar results, where they found Nakabango and Bulindi as the most discriminating testing locations for soybean in Uganda. Bulindi and Abi were discriminating genotypes but not representative and therefore, these two sites could be used together as "culling environments" for easily selecting against unstable genotypes during the breeding process (Yan and Kang, 2002). Nakabango was both discriminating and representative. The results are in agreement with the findings of Tukamuhabwa *et al*. (2012); Obua (2013) who reported Nakabango as the most discriminating and representative testing location for soybean in Uganda. Discriminating and representative test locations are useful for selecting superior genotypes while eliminating inferior ones (Etnaf *et al*., 2013).

CHAPTER FIVE

5.0 GENARAL DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS

5.1 General discussion

Gene diversity, varied from 0.208 to 0.725 with an average of 0.548 implying that there was moderate genetic variation existed among the evaluated genotypes and hence in the molecular markers grouping was more logical than the grouping based on origin and pedigree. Genetic diversity study among released varieties and advanced elite soybean genotypes based on the hierarchical clustering divided the soybean genotypes into three distinct groups, with more than half of the released varieties grouped in the same cluster. The results of both the dendrogram and Principal Coordinate analysis (PCoA) were consistent with each other in grouping the 34 soybean genotypes into three distinct classes in which all released varieties were found in cluster I showing that most released varieties were developed from common parent. This implies that most of the released varieties will lose stability as well as resistance in case of changes in pests, diseases and abiotic stresses. Therefore, inclusion of more diverse germplasm and use of landraces in the soybean breeding programmme may provide the level of genetic variation necessary to cope with dynamic pests, diseases and abiotic stresses of Uganda.

Soybean grain yield is highly sensitive to G × E interaction as reported by several studies (Gurmu *et al*., 2009; Tukamuhabwa *et al*., 2012; Obua, 2013; Atnaf *et al*., 2013; Bhartiya *et al*., 2017; Sousa *et al*., 2018). The presence of crossover G × E interaction for grain yield complicates the selection and recommendation of superior varieties in a breeding program. When G × E interaction is high, it should not be ignored, but rather exploited as suggested by Yan and Kang (2002). Therefore, there is need to use of stability models like AMMI and GGE to clearly dissect the nature and cause of that interaction especially in cases where significant G × E interaction for quantitative trait like grain yield is mainly due to variability in soil nutritional status, temperature, rainfall, pests and diseases severity and incidence. The most inexpensive and farmer-friendly method is to use soybean varieties adapted to the target environments (Tukamuhabwa *et al*., 2012). Meanwhile, since crop growing locations have no clearly defined demarcations and most farmers tend to influence each other in the choice of variety that is

grown, the development of varieties adapted to a wide range of environments is strongly supported, rather than environment-specific varieties (Gasura *et al*., 2015; Bhartiya *et al*., 2017). The absence of significant genotype, G × E interaction for protein and oil content observed in this study revealed that no genetic variation existed among genotypes and the absence of significant G × E interaction for quantitatives trait like protein and oil content indicated the consistency performance of genotypes across locations. This implies that there is no need to test protein and oil content in multi-location since there is no significant G × E interaction. The absence of significant genotype, G × E interaction for protein and oil content observed in this study were inconsistences with previous studies by Gurmu *et al*. (2009); Hampango *et al*. (2017), who reported the presence of significance genotype, G × E interaction for protein and oil content.

5.2 Conclusions

There is moderate genetic variation among almost all released varieties and advanced elite soybean genotypes, excluding the recently released varieties, Maksoy 4N, Maksoy 5N and Maksoy 6N, which are closely related to many of the advanced elite genotypes.

For specific adaptation, the biplot identified winning genotypes in each mega-environment as follows; BSPS 48A-28 for mega-environment I (Bulindi, Nakabango and Kabanyolo), BSPS 48A-28-1 for mega-environment II (Iki-Iki), Bulindi 18.4B for mega-environment III (Ngetta) and BSPS 48A-24-1 for mega-environment IV (Abi). Based on mean yield and stability (for broad adaptation) performances, Maksoy 3N was the most stable genotype while BSPS 48A-28 was the highest yielding genotype with moderate stability.

In conclusion, the study observed moderate genetic variation among the soybean genotypes evaluated. BSPS 48A-28 outperformed the released varieties (Maksoy 3N and Maksoy 4N) that were included in the evaluation trial.

5.3 Recommendations

(i) There is need for wider genetic base of the working germplasm and elite introductions for mid to long term soybean breeding.

(ii) Superior and widely adapted genotypes BSPS 48A-28 and NDGT 8.11 × 3N-2 as well as specific adapted genotypes 2N × GC; Bulindi 18.4B; NGDT 8.11 × 16.16B and BSPS 48A-24-1 should be submitted for tests of novelty, distinctiveness, uniformity and stability (DUS) by the Ministry of Agriculture Animal Industries Department of the National Seed Certification Service.

(iii) No need to test for protein and oil content on multi-location, since there was non-significant G × E interaction for those traits.

REFERENCES

Agoyi, E. E., Odong, T. L., Tumuhairwe, J. B., Chigeza, G., Diers, B. W., and Tukamuhabwa, P. (2017). Genotype by environment effects on promiscuous nodulation in soybean (*Glycine max* L. Merrill). *Agriculture and Food Security*, 6(1), 1-14. https://doi.org/10.1186/s40066-017-0107-7.

Anderson, J. A., Churchill, G. G., Autrique, J. E., Tanksley, S. D., and Sorrells, M. E. (1993). Optimizing parental selection for genetic linkage maps. *Genome*, 36: 18 1-186.

Atnaf, M., Kidane, S., Abadi, S., and Fisha Z. (2013). GGE biplots to analyze soybean multi-environment yield trial data in north Western Ethiopia. *Journal of Plant Breeding and Crop Science,* 5(12), 245-254. https://doi.org/10.5897/JPBCS13.0403.

Batista, R. O., Hamawaki, R. L., Sousa, L. B., and Hamawaki, O. T. (2015). Adaptability and stability of soybean genotypes in off-season cultivation. *Genetics and Molecular Research*, 14(3), 9633-9645. http://dx.doi.org/10.4238/2015.August.14.26.

Bernardo, R. (2002). Breeding for quantitative traits in plants. *Stemma Press*, St. Paul, MN. 369 pp.

Bhartiya, A., Aditya, J. P., Kumari, V., Kishore, N., Purwar, J. P., Agrawal, A., and Kant, L. (2017). GGE biplot and AMMI analysis of yield stability in multi-environment trial of soybean [*Glycine max* (L.) Merrill] genotypes under rainfed condition of north western Himalayan hills. *Journal of Animal and Plant Sciences*, 27(1), 227-238.

Bisen, A., Khare, D., Nair, P., and Tripathi, N. (2015). SSR analysis of 38 genotypes of soybean (*Glycine max* (L.) Merr.) genetic diversity in India. *Physiology and Molecular Biology of Plants*, 109-115. https://doi.org/10.1007/s12298-014-0269-8.

Chakraborty, S., Patel, D.A., Parmar, H., Dhaduk, H.L., and Sasidharan, N. (2018). Genetic diversity analysis in soybean (*Glycine max* (L.) Merrill.) using SSR markers. *Journal of Pharmacognosy and Phytochemistry*, 7(3): 2380-2384.

Chaudhary, K. R., and Wu, J. (2012). Stability analysis for yield and seed quality of soybean [*Glycine max* (L.) Merril] across different environments in eastern South Dakota.

Conference on Applied Statistics in Agriculture. https://doi.org/10.4148/2475-7772.1033. 65

Chauhan, D. K., Bhat, J. A., Thakur, A. K., Kumari, S., Hussain, Z., and Satyawathi, C. T. (2015). Molecular characterization and genetic diversity assessment in soybean [*Glycine max* (L.) Merr.] varieties using SSR markers. *Indian Journal of Biotechnology*, 504-510.

Clemente, T. E., and Cahoon, E. B. (2009). Soybean Oil : Genetic Approaches for Modification of Functionality and Total Content. *Plant Physiology*, 1030-1040. https://doi.org/10.1104/pp.109.146282.

Cregan, P. B., Jarvik, T., Bush, A. L., Shoemaker, R. C., Lark, K. G., Kahler, A. L., and Specht, J. E. (1999). An Integrated Genetic Linkage Map of the Soybean Genome. *Crop Science*, 1464-1490.

Crossa, J., Cornelius, P. L., and Yan, W. (2002). Biplots of Linear-Bilinear Models for Studying Crossover Genotype. *Crop Science*, 619-633.

Deshmukh, V. P., Thakare, P.V., Chaudhari, U.S., and Gawande, P.A. (2007). A simple method for isolation of genomic DNA from fresh and dry leaves of Terminalia arjuna (Roxb .) Wight and Argot. *Electronic Journal of Biotechnology*, 10 (3). https://doi.org/10.2225/vol10.

Dolinassou, S., Baptiste, J., Noubissié, T., Mamoudou, M., Nguimbou, M., and Njintang, N.Y. (2017). Genotype × environment interactions and oil content stability analysis of peanut (*Arachis hypogaea* L.) in Northern Cameroon. *Journal of Plant Breeding and Crop Science*, 45-53. https://doi.org/10.5897/JPBCS2017.0647.

Du, W., Wang, M., Fu, S., and Yu, D. (2009). Mapping QTLs for seed yield and drought susceptibility index in soy- bean (*Glycine max* L.) across different environments. *Journal of Genetics and Genomics*, 36(12), 721–731. https://doi.org/10.1016/S1673-8527(08)60165-4

Eberhart, S. A., and Russell, W. A. (1966). Stability Parameters for Comparing Varieties, (3).

Finlay, K. W. and Wilkinson, G. N. (1963). The analysis of adaptation in a plant breeding programme. Journal of Agricultural Research, 14: 742-754.

Fu, Y., Peterson, G. W., and Morrison, M. J. (2007). Genetic diversity of Canadian soybean cultivars and exotic germplasm revealed by simple sequence repeat markers. *Crop Science Society of America,* https://doi.org/10.2135/cropsci2006.12.0843.

Gasura, E., Setimela, P., and Souta, C. (2015). Evaluation of the performance of sorghum genotypes using GGE biplot. *Canadian Journal of Plant Science*, (613), 150713053357002. https://doi.org/10.4141/CJPS-2015-119. 66

Gauch H. G., Jr. (2013). A simple protocol for AMMI analysis of yield trials. *Crop Science Society of America*, 53: 1860-1869. https://doi.org/10.2135/cropsci2013.04.0241.

Gauch, H. G. (2006). Statistical Analysis of Yield Trials by AMMI and GGE, *Crop Science Society of America,* 1488-1500. https://doi.org/10.2135/cropsci2005.07-0193.

Gebremedhn, H. M., Weldekidan, M. B., Msiska, U. M., Asmamaw, F. A., Odong, T. L., and Tukamuhabwa, P. (2018). Inheritance of soybean resistance to soybean rust in Uganda's soybean germplasm. *International Journal of Agronomy and Agricul*tural *Research*. (12 (1), 26-36.

Genstat. (2016). Genstat Release 13.3 (PC/Winbars Vista). VSN International Limited.

Ghosh, J., Ghosh, P. D., and Choudhury, P. R. (2014). An Assessment of Genetic Relatedness between Soybean [*Glycine max* (L.) Merrill] Cultivars Using SSR Markers. *American Journal of Plant Sciences*, 3089-3096.

Guan, R., Chang, *Æ*. R., Li, *Æ*. Y., Wang, L., Liu, *Æ*. Z., and Qiu, *Æ*. L. (2010). Genetic diversity comparison between Chinese and Japanese soybeans (*Glycine max* (L.) Merr.) revealed by nuclear SSRs. *Genetic Resource and Crop Evolution*, 229-242. https://doi.org/10.1007/s10722-009-9465-8.

Gupta, S. K., and Manjaya, J. G. (2017). Genetic Diversity and Population Structure of Indian Soybean [*Glycine Max* (L .) Merr .] Revealed by Simple Sequence Repeat Markers. *Journal of Crop Science and Biotechnology*, (10), 221-222.

Gurmu, F., Hussein, S., and Laing, M. (2017). Genotype-by-environment interaction and stability of sweetpotato genotypes for root dry matter, β-carotene and fresh root yield. *Open*

Agriculture, 473-485. https://doi.org/10.1515/opag-2017-0052.

Gurmu, F., Mohammed, H., and Alemaw, G. (2009). Genotype x environment interactions and stability of soybeanfor grain yield and nutrition quality. *African Crop Science Journal*, 17(2), 87-99.

Hampango, M., Cheelo, P., Lungu, M., Mwala, M., and Agrama, H. (2017). AMMI determination of stability for oil and protein content in soybean (*Glycine max* L. Merrill) seed in Zambia. *Journal of Experimental Agriculture International*, 17(5), September, 1-16. https://doi.org/10.9734/JEAI/2017/35583.

Hipparagi, Y., Singh, R., Choudhury, D. R., and Gupta, V. (2017). Genetic diversity and population structure analysis of Kala bhat (*Glycine max* (L.) Merrill) genotypes using SSR markers. Hereditas, 1-11. https://doi.org/10.1186/s41065-017-0030-8.

Hu, Z., Kan, G., Zhang, G., and Zhang, D. (2018). Genetic diversity analysis using simple sequence repeat markers in soybean. *Plant Genetic Resources*, 12, 87-90. https://doi.org/10.1017/S1479262114000331.

Hua, X. I. E., Rongxia, G., and Ruzhen, C. (2005). Genetic diversity of Chinese summer soybean germplasm revealed by SSR markers. *Chinese Bulletin*, 50(6). https://doi.org/10.1360/982004-805.

Ibanda, A. P., Karungi, J., Malinga, G. M., and Adjumati, G. (2018). Influence of environment on soybean [Glycine max (L.) Merr.] resistance to groundnut leaf miner, [(*Aproaerema modicella* (Deventer)] in Uganda. *Journal of Plant Breeding and Crop Science* , https://doi.org/10.5897/JPBCS2018.0764.

Krisnawati, A., and Adie, M. M. (2018). Yield stability of soybean promising lines across environments. IOP Conference Series: *Earth and Environmental Science*, 102(1). https://doi.org/10.1088/1755-1315/102/1/012044.

Kumar, A., Kumar, S., and Kapoor, C. (2014). GGE Biplot Analysis of Genotype × Environment Interaction in Soybean Grown in NEH Regions of India. *Environment and Ecology*, 32, 1047-1050.

Kumawat, G., Singh, G., Gireesh, C., and Shivakumar, M. (2015). Molecular characterization and genetic diversity analysis of soybean (*Glycine max* (L .) Merr .) germplasm accessions in India. *Physiology and Molecular Biology*, 21, 101-107. https://doi.org/10.1007/s12298-014-0266-y. 68

Li, Y., Guan, R., Liu, Z., Ma, Y., Wang, L., Li, L ., and Qiu, L. (2008). Genetic structure and diversity of cultivated soybean (*Glycine max* (L .) Merr .) landraces in China. *Theoretical and Applied Genetics*, (2008) 857-871. https://doi.org/10.1007/s00122-008-0825-0.

Liu, W., Young, M., Yang, K., Kang, J., Lee, K. V. Y., Srinives, P., and Lee, Y. S. (2011). QTL identification of flowering time at three different latitudes reveals homologous genomic regions that control flowering in soybean, 545–553. https://doi.org/10.1007/s00122-011-1606-8

Liu, Z., Fan, X., Huang, W., Yang, J., Zheng, Y., and Qiu, L. (2017). Stability analysis of seven agronomic traits for soybean [(*Glycine max* (L.) Merr.] using the AMMI model. *Plant Production Science*, 1008, 1-8. https://doi.org/10.1080/1343943X.2017.1358095.

Maphosa, M., Talwana, H., and Tukamuhabwa, P. (2013). Assessment of Comparative Virulence and Resistance in Soybean Using Field Isolates of Soybean Rust. *Journal of Agricultural Science*, 5(5), 249-257. https://doi.org/10.5539/jas.v5n5p249.

Mare, M., Manjeru, P., Ncube, B., and Sisito, G. (2017). GGE biplot analysis of genotypes by environment interaction on *Sorghum bicolor* L. (Moench) in Zimbabwe. *African Journal of Plant Science*, 11(7), 308-319. https://doi.org/10.5897/AJPS2017.1538.

Maughan, P. J., Maroof, M. A. S., and Buss, G. R. (1995). Microsatellite and amplified sequence length polymorphisms in cultivated and wild soybean. *Genome*, 723, 715-723.

Mclntosh, M. S. (1983). Analysis of combined experiments. *Agronomy Journal*, 75: 153-155.

Merritt, R. J., and Jenks, B. H. (2004). Fifth International Symposium on the Role of Soy in Preventing and Treating Chronic Disease Safety of Soy-Based Infant Formulas Containing Isoflavones, *The Clinical Evidence* 1, 1220-1224.

Miles, M. R., Frederick, R. D., and Hartman, G. L. (2006). Evaluation of Soybean Germplasm

for Resistance to soybean rust (Phakopsora pachyrhizi). *Plant Health Progress*, 1(January), 1-6. https://doi.org/10.1094/PHP-2006-0104-01-RS.

Moniruzzaman, M., Saiem, R. M., Emon, R. M., Haque, M. S., Saha, N. R., Malek, M. A., and Khatun, K. (2019). Genetic diversity analysis of soybean genotypes using SSR markers for salinity tolerance. *Progressive Agriculture*, 3030(1), 1-9.

Moore, K. J., and Dixon, P. M. (2015). Analysis of Combined Experiments Revisited. *Agronomy Journal*, 763–771. https://doi.org/10.2134/agronj13.0485. 69

Mulato, B. M., Möller, M., Zucchi, M. I., Quecini, V., and Pinheiro, J. B. (2010). Genetic diversity in soybean germplasm identified by SSR and EST-SSR markers Genetic diversity in soybean germplasm identified by SSR and EST-SSR markers. *Brasília*, (August 2014). https://doi.org/10.1590/S0100-204X2010000300007.

Namara, M. (2015). Resistance of soybean germplasm to the groundnut leaf miner (Aproaerema modicella) in Uganda. RUFORUM, *www.http://repository.ruforum.org*/documents/ resistance-soybean germplasm-groundnut-leaf-miner-*aproaerema-modicella*-uganda.

NARO Ngheta-Zardi. (2018). Weather conditions. National Agricultural Research Organisation (NARO), *Ministry of Agriculture, Animal Industry and Fisheries,* Entebbe, Uganda.

Narvel, J. M., Fehr, W. R., Chu, W., Grant, D., and Shoemaker, R. C. (2000). Simple Sequence Repeat Diversity among Soybean Plant Introductions and Elite Genotypes. *American Crop Science*, 412–417.

Nascimento, M., Finoto, E. L., Sediyama, T., and Cruz, D. C. (2010). Adaptability and stability of soybean in terms of oil and protein content. *Crop Breeding and Applied Biotechnology*, 48-54. https://doi.org/10.12702/1984-7033.v10n01a07.

Nei, M. (1972). Genetic Distance between Populations. *American Naturalist*, 106, 283-292. https://doi.org/10.1086/282771.

Njoroge. J. N., Owouche. J.O., and Oyoo. M. E. (2015). Evaluation of soybean [Glycine max (L.)Merr.] genotypes for agronomic and quality traits in Kenya. *African Journal of Agricultural Research*, 10(12), 1474-1479. https://doi.org/10.5897/AJAR2014.9168.

Northeast China Based on Nuclear Microsatellite Variation. *International Journal of Genomics*, 2018 https://doi.org/10.1155/2018/8561458.

Obua, T. (2013). Soybean rust diversity and adaptation of elite soybean lines to the Ugandan environment. *http://repository.ruforum.org/sites/default/files/*Thesis-Obua%20Tonny.pdf.

Oda, M. C., Sediyama, T., Matsuo, É., Cruz, C. D., and Barros, E. G. De. (2015). Phenotypic and molecular traits diversity in soybean launched in forty years of genetic breeding. *Agronomy Science and Biotechnology*, 1(1), 1-9.

Ota, T. (1993). DISPAN: Genetic distance and phylogenetic analysis. Pennsylvania State University, Pennsylvania. 70

Owusu-Apenten, R. (2002). Food Protein Analysis: Quantitative Effects On Processing, *CRC Press*. ISBN 9780824706845 - CAT#DK1906.

Pagar, T. A., Akhare, A. A., Gahukar, S. J., Mujahid, S., and Gawande, A. M. (2017). DNA fingerprinting of soybean (Glycine max L.) genotypes by using simple sequence repeats (SSR) markers. *International Journal of Chemical Studies extension*, 5(5), 674-679.

Peakall, R., and Smouse, P. E. (2012). GenAIEx V5 : Genetic Analysis in Excel. Populations Genetic Software for Teaching and Research GenAlEx 6.5 : genetic analysis in Excel. Population genetic software for teaching and research, https://doi.org/10.1093/bioinformatics/bts460.

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

Pinnschmidt, H. O., and Hovmøller, M. S. (2002). Genotype × environment interactions in the expression of net blotch resistance in spring and winter barley varieties. *Euphytica*, 125: 227-243.

Popovic, V., Miladinovic, J., Tatic, M., Djekic, V., Dozet, G., Đukić, V., and Grahovac, N. (2013). Stability of soybean yield and quality components. *African Journal of Agricultural Research*, 8(45), 5651-5658. https://doi.org/10.5897/AJAR12.1146.

Rohlf, F.J. (2000). NTSYS-pc: numerical taxonomy and multi- variate analysis system, version

2.1. Owner manual.

Rohlf, F.J. (2005). NTSYS-Pc: numerical taxonomy and multivariate analysis system. Version 2.2. Setauket: Exeter Software.

Sacks, F. M., Lichtenstein, A., Horn, L. Van, Harris, W., Kris-etherton, P., and Winston, M. (2006). Soy Protein , Isoflavones , and Cardiovascular Health An American Heart Association Science Advisory for Professionals From the Nutrition Committee. *AHA Science Advisory*, 1034–1044. https://doi.org/10.1161/circulationaha.106.171052.

Sato, T. (2010). Estimation of Fatty Acid Composition in Soybean Powder by Examining Near Infrared Spectroscopic Patterns. *National Agriculture and Food Research Organization* (NARO), Japan. www.intechopen.com. 71

Shadakshari, T. V, Senthil, N., Kalaimagal, T., Boranayaka, M. B., Gowda, R. K., and Rajesha, G. (2011). Genetic diversity analysis in soybean [*Glycine max* (L.) Merrill] using simple sequence repeat (SSR) markers. *International Journal of Plant Sciences*, 149-153.

Simion, T., Mohammed, W., and Amsalu, B. (2018). Genotype by environment interaction and stability analysis of cowpea [*Vigna unguiculata* (L.) Walp] genotypes for yield in Ethiopia. Journal of Plant Breeding and Crop Science, 249-257. https://doi.org/10.5897/JPBCS2018.0753.

Song, J. Y., Piao, X., Choi, Y., Lee, G., Chung, J., Lee, J., and Lee, M. C. (2013). Evaluation of Genetic Diversity and Comparison of Biochemical Traits of Soybean (*Glycine max* L.) Germplasm Collections. *Journal of Plant Breeding and Biotechnology*, (4), 374-384. http://dx.doi.org/10.9787/PBB.2013.1.4.374.

Song, Q., Jia, G., Grant, D., and Nelson. R.T. (2010). Abundance of SSR motifs and development of Candidate Polymorphic SSR markers in Soybean. *Crop Science*, 50:1950-1960.

Sousa, M. B. E., Damasceno-silva, K. J., and Moura, M. D. E. (2018). Genotype by environment interaction in cowpea lines using GGE biplot method. https://periodicos.ufersa.edu.br/index. 2125, 64-71. http://dx.doi.org/10.1590/1983-21252018v31n108rc.

Ssendege, G., Obua, T., Kawuki, R., Maphosa, M., and Tukamuhabwa, T.P. (2015). Soybean genetic diversity and resistance to soybean rust disease in Uganda. *Agricultural Journal*, 10(3-6), 17-23.

Tamura, K., Dudley, J., Nei, M., and Kumar, S. (2007). MEGA4 : Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4. 0, 0-3. *Journal of Molecular Biology and Evolution*, https://doi.org/10.1093/molbev/msm092.

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

Tukamuhabwa, P. Assimwe, M. Nabasirye, M, Kabayi, P. and Maphosa, M. (2012). Genotype by Environment Interaction of Advanced Generation. *African Crop Science Journal*, 20(2), 107-115. 72

Tukamuhabwa, P., and Obua, T. (2015). Soybean production guide in Uganda, Makerere University Agricultural Research Institute, Kabanyolo (MUARIK), Makerere University, Kampala.

Tukamuhabwa, P., and Oloka, H. K. (2016). Soybean research and development in Uganda: a case of paradigm shift in an African University. Makerere University Agricultural Research Institute, Kabanyolo (MUARIK), Makerere University, Kampala.

Tukamuhabwa, P., Obaa, B., Obua, T., Namara, M., Okii, D., Kabayi, P., and Sciencee, E. (2016). Status of Soybean Production and Impact Indicators of New Soybean Varieties in Uganda, 1-30.

Tyagi, S. D., Sethi, T. P.J., and Tyagi, M.V. (2013). Assessment of Genotype × Environment Interactions and Stability Analysis in Soybean [*Glycine max* (L.) Merrill] genotypes. *Indian Journal of Ecology*, 40, 268-271.

Wang, L., Guan, R., and Qiu, S. L. (2006). Genetic Diversity of Chinese Cultivated Soybean Revealed by SSR Markers. *American Crop Science*, https://doi.org/10.2135/cropsci2005.00.

Wang, Y. H., Zhang, X. J., and Fan, S. J. (2015). Genetic diversity of wild soybean populations

in Dongying, China , by simple sequence repeat analysis. *Genetics and Molecular Research*, 14(3), 11613-11623.

Wang. M., Run-zhi, L., Wan-ming, Y., and Wei-jun, D. (2010). Assessing the genetic diversity of cultivars and wild soybeans using SSR markers. *African Journal of Biotechnology*, 9 (31), 4857-4866.

Westcott, B. (1986). Some methods of analysing genotype - environment interaction. *Heredity*, 56, 243-253.

Widaningsih, N. A., and Purwanto, E., and Nandariyah. (2014). The Use of DNA Microsatellite Markers for Genetic Diversity Identification of Soybean (Glycine max (L) Meriil .) as a Supplementary Method in Reference Collections Management. *Indonesian Journal of Biotechnology*, 19(2), 136-145. 73

Wortmann, C.S., and Eledu, C.A. (1999). Uganda's agroecological zones: A guide to planners and policy makers. Centro Internationale de Agricultural Tropical (CIAT), Kawanda, Uganda.

Yan, W., and Hunt, L. A. (2001). Interpretation of genotype × environment interaction for winter wheat yield in Ontario. *Crop Science*, 41(1), 19-25. https://doi.org/10.1128/AAC.00481-06

Yan, W., and Kang, M. S. (2002). GGE biplot analysis. A grafical tools for breeders, geneticists, and agronomists. *Crc Press*, (1338), 224. https://doi.org/10.1201/9781420040371.

Yan, W., and Rajcan, I. (2003). Prediction of Cultivar Performance Based on Single-versus Multiple-Year Tests in Soybean. *American Crop Science Journal*. 555, 549-555.

Yan, W., and Tinker, N. (2006). Biplot analysis of multi-environment trial data: Principles and applications. *Canadian Journal of Plant Science*, 86(3), 623-645. https://doi.org/10.4141/P05-169.

Yan, W., Glover, K. D., and Kang, M. S. (2010). Statistical test of genotype-by-environment interaction patterns observed from a biplot. *Crop Science*, 50, 1121-1123.

Yan, W., Hunt, L. A., Sheng, Q., and Szlavnics, Z. (2000). Cultivar evaluation and mega-

environment investigation based on the GGE biplot. *American Crop Science*, 40: 597-605.

Yan, W., Kang, M. S., Ma, B., Woods, S., and Cornelius, P. L. (2007). GGE biplot vs. AMMI analysis of genotype-by-environment data. *American Crop Science Journal*, 47(2), 643-655. https://doi.org/10.2135/cropsci2006.06.0374.

Yang, R., Crossa, J., Cornelius, P. L., and Burgueño, J. (2009). Biplot Analysis of Genotype × Environment Interaction : Proceed with Caution. *American Crop Science Journal*, (October). https://doi.org/10.2135/cropsci2008.11.0665.

Zhang, G., Xu, S., Mao, W., Hu, Q., and Gong, Y. (2013). Determination of the genetic diversity of vegetable soybean [*Glycine max* (L.) Merrill.] using EST-SSR markers. *Biomedical and Biotechnology*, 14(4), 279-288. https://doi.org/10.1631/jzus.B1200243.

Zhao, H., Wang, Y., Xing, F., Liu, X., Yuan, C., Qi, G., and Dong, Y. (2018). The Genetic Diversity and Geographic Differentiation of the Wild Soybean [*Glycine* soja (L.) Merrill.] in Northeast China Based on Nuclear Microsatellite Variation. *International Journal of Genomics*, 2018 https://doi.org/10.1155/2018/8561458.

APPENDICES

Appendix 1: ANOVA table for grain yield (kg/ha) of 25 soybean genotypes evaluated in Bulindi across two seasons of 2018A and B

Source of Variation	DF	SS	MS	VR	F PR.
Season	1	25	25	0	0.998
Replication. Season	4	21972205	5493051	18.68	<0.001
Genotype	24	6118519	254938	0.87	0.644
Season × Genotype	24	7097639	295735	1.01	0.467
Residual	96	28229503	294057		
Mean		2650			
CV		20.5			
LSD		621.5			

Appendix 2: ANOVA table for grain yield (kg/ha) of 25 soybean genotypes evaluated in Iki-Iki evaluated across two seasons of 2018A and B

Source of Variation	DF	SS	MS	VR	F PR.
Season	1	66967004	66967004	294.8	<0.001
Replication. Season	4	908656	227164	4.02	0.004
Genotype	24	3310601	137942	2.44	0.001
Season × Genotype	24	3016203	125675	2.22	0.003
Residual	96	5431388	56577		
Mean		889			
CV		26.8			
LSD		272.6			

Appendix 3: ANOVA table for grain yield (kg/ha) of 25 soybean genotypes evaluated in Kabanyolo evaluated across two seasons of 2018A and B

Source of Variation	DF	SS	MS	VR	F PR.
Season	1	7506254	7506254	16.26	0.016
Replication. Season	4	1846666	461666	4.53	0.002
Genotype	24	4943863	205994	2.02	0.009
Season × Genotype	24	2622391	109266	1.07	0.389
Residual	96	9783231	101909		
Mean		1017			
CV		31.4			
LSD		365.8			

Appendix 4: ANOVA table for grain yield (kg/ha) of 25 soybean genotypes evaluated in Nakabango across two seasons of 2018A and B

Source of Variation	DF	SS	MS	VR	F PR.
Season	1	33111504	33111504	19.59	0.011
Replication. Season	4	6761996	1690499	16.12	<0.001
Genotype	24	7264968	302707	2.89	<0.001
Season × Genotype	24	2103238	87635	0.84	0.684
Residual	96	10067980	104875		
Mean		1698			
CV		19.1			
LSD		371.1			

Appendix 5: ANOVA table for grain yield (kg/ha) of 25 soybean genotypes evaluated in Ngheta across two seasons of 2018A and B

Source of Variation	DF	SS	MS	VR	F PR.
Season	1	61320591	61320591	64.93	0.001
Replication. Season	4	3777837	944459	12.43	<0.001
Genotype	24	3646800	151950	2	0.01
Season × Genotype	24	2539338	105806	1.39	0.132
Residual	96	7297100	76011		
Mean		1567			
CV		17.6			
LSD		316			

Appendix 6: ANOVA table for grain yield (kg/ha) of 25 soybean genotypes evaluated in ABI across two seasons of 2018A and B

Source of Variation	DF	SS	MS	VR	F PR.
Season	1	43965289	43965289	9.1	0.039
Replication. Season	4	19326381	4831595	21.96	<0.001
Genotype	24	5907637	246152	1.12	0.339
Season × Genotype	24	9488752	395365	1.8	0.024
Residual	96	21116999	219969		
Mean		1845			
CV		25.4			
LSD		537.5			

Appendix 7: ANOVA table for percentage oil content of 25 soybean genotypes evaluated in Bulindi across two seasons of 2018A and B

Source of Variation	DF	SS	MS	VR	F PR.
Season	1	168.644	168.644	4.71	0.162
Replication. Season	2	71.567	35.784	4.99	
Genotype	24	141.008	5.875	0.82	0.697
Season × Genotype	24	155.538	6.481	0.9	0.597
Residual	48	344.506	7.177		
Mean		16.71			
CV		16			
LSD		3.809			

Appendix 8: ANOVA table for percentage oil content of 25 soybean genotypes evaluated in Nakabango across two seasons of 2018A and B

Source of Variation	DF	SS	MS	VR	F PR.
Season	1	12.465	12.465	1.55	0.339
Replication. Season	2	16.073	8.036	4.01	
Genotype	24	78.284	3.262	1.63	0.075
Season × Genotype	24	37.922	1.58	0.79	0.732
Residual	48	96.186	2.004		
Mean		16.79			
CV		8.4			
LSD		2.013			

Appendix 9: ANOVA table for percentage oil content of 25 soybean genotypes evaluated in Iki-Iki across two seasons of 2018A and B

Source of Variation	DF	SS	MS	VR	F PR.
Season	1	82.361	82.361	169.89	0.006
Replication. Season	2	0.97	0.485	0.13	
Genotype	24	87.453	3.644	0.94	0.553
Season × Genotype	24	88.051	3.669	0.95	0.545
Residual	48	186.023	3.875		
Mean		14.92			
CV		13.2			
LSD		2.799			

Appendix 10: ANOVA table for percentage oil content of 25 soybean genotypes evaluated in ABI across two seasons of 2018A and B

Source of Variation	DF	SS	MS	VR	F PR.
Season	1	20.073	20.073	14.84	0.061
Replication. Season	2	2.705	1.352	0.25	
Genotype	24	102.848	4.285	0.80	0.714
Season × Genotype	24	74.868	3.12	0.59	0.921
Residual	48	255.911	5.331		
Mean		15.61			
CV		14.8			
LSD		3.283			

Appendix 11: ANOVA table for percentage protein content of 25 soybean genotypes evaluated in Iki-Iki across two seasons of 2018A and B

Source of Variation	DF	SS	MS	VR	F PR.
Season	1	0.064	0.064	0.05	0.839
Replication. Season	2	2.396	1.198	0.30	0.742
Genotype	24	76.691	3.195	0.81	0.709
Season × Genotype	24	69.708	2.904	0.73	0.791
Residual	48	189.787	3.954		
Mean		32.35			
CV		6.1			
LSD		2.827			

Appendix 12: ANOVA table for percentage protein content of 25 soybean genotypes evaluated in Nakabango across two seasons of 2018A and B

Source of Variation	DF	SS	MS	VR	F PR.
Season	1	0.055	0.055	0.13	0.753
Replication. Season	2	0.845	0.423	0.06	0.941
Genotype	24	101.03	4.21	0.55	0.944
Season × Genotype	24	240.163	10.007	1.3	0.214
Residual	48	368.688	7.681		
Mean		33.57			
CV		8.3			
LSD		3.94			

Appendix 13: ANOVA table for percentage protein content of 25 soybean genotypes evaluated in Bulindi across two seasons of 2018A and B

Source of Variation	DF	SS	MS	VR	F PR.
Season	1	4.012	4.012	15.89	0.058
Replication. Season	2	0.505	0.252	0.03	0.970
Genotype	24	193.933	8.081	0.87	0.630
Season × Genotype	24	309.06	12.877	1.39	0.162
Residual	48	443.335	9.236		
Mean		34.29			
CV		8.9			
LSD		4.321			

Appendix 14: ANOVA table for percentage protein content of 25 soybean genotypes evaluated in Abi across two seasons of 2018A and B

Source of Variation	DF	SS	MS	VR	F PR.
Season	1	0.207	0.207	0.03	0.886
Replication. Season	2	15.645	7.823	0.80	0.455
Genotype	24	201.457	8.394	0.86	0.653
Season × Genotype	24	106.452	4.436	0.45	0.981
Residual	48	470.529	9.803		
Mean		33.95			
CV		9.2			
LSD		4.451			