Genetic characterization and analysis for yield and quality related traits in the Zimbabwean flue-cured tobacco germplasm

By

Justify Gotami Shava

A thesis submitted in partial fulfilment for the requirements of the degree of Doctor of Philosophy (D. Phil) in Agriculture (Plant Breeding)

Department of Crop Science

Faculty of Agriculture

University of Zimbabwe

May 2019

DECLARATION

I *Justify Gotami Shava* hereby declare that the thesis which I here submit for the degree of Doctor of Philosophy (D. Phil) at the University of Zimbabwe is my own independent and original work which has not been previously submitted at any other university or faculty within this university by me or any other person. I, therefore, cede copyrights of this thesis in favour of the University of Zimbabwe.

Signature…… Date…….*30/05/19*……

Justify Shava

Signature… Sue Kageler …… Date…….*30/05/19*…….

Dr. S. J. Kageler - Crop Science Department, University of Zimbabwe

Signature……..……

Date…….*30/05/19*…….

Dr. S. Dari - Crop Science Department, University of Zimbabwe

Signature……..…… Date…….

Dr. D. Rukuni - Tobacco Research Board, Zimbabwe

Signature……..…… Date…….

Dr. F. Magama (Tobacco Research Board, Zimbabwe)

Signature……..…… Date…….

Dr. E. Gasura (Chairperson – Crop Science Department, University of Zimbabwe)

ABSTRACT

Tobacco (*Nicotiana tabaccum* L.) is a crop with commercial value in most countries of the world and not much work has been done to improve its performance. Information on availability of genetic diversity expressible in the phenotype, understanding gene action influencing key traits and adaptation of tobacco varieties to different growing environments are a prerequisite for effective tobacco breeding and subsequent production. This study was conducted (a) to determine the extent of genetic diversity within the Zimbabwean flue-cured tobacco (*Nicotiana tabacum* L.) germplasm, (b) to determine the nature of gene action for key traits in tobacco as well as (c) to investigate genotype × environment interaction effects in tobacco production. To determine the molecular diversity of flue-cured tobacco germplasm, a sample of 93 accessions were extracted from the Tobacco Research Board germplasm collection. Genomic DNA was extracted from leaves using a Direct Zol™ Plant DNA extraction kit. Microsatellite amplifications were performed using 10 SSR markers and analysis of molecular variance (AMOVA) was used to evaluate the partitioning of diversity within and between populations and between sample groups using Arlequin. Morphological diversity among tobacco germplasm was determined by evaluating 196 accessions in a 14×14 α-Lattice design with two replicates at Kutsaga research station for two seasons. Data analysis was done using Genstat Version 17. Genetic analyses for growth rate, ripening rate and flue-cured leaf colour composition of the flue-cured tobacco was done using Griffing's Method I, Model I; and Hayman's $W_r - V_r$ graphic analysis. A ten-parent full diallel cross including reciprocals and parentals was laid out in an α-Lattice experimental design with two replicates over a two-year period at Kutsaga research station. To investigate genotype × environment interaction effects, experiments were set out in a randomised complete block design with three replicates and twelve genotypes were evaluated across eight locations and five parameters were measured. Data was subsequently subjected to Additive Main Effects and Multiplicative Interaction (AMMI) Model and Genotype×environment interaction (GGE) biplot analysis. Results from genetic and morphological characterization showed that the studied germplasm was morphologically closely related, despite originating from diverse parts of the world since the majority of the accessions had similarity coeficiencies of above 60%. However, there were some genotypes like Coker 347, which were different from the rest of the genotypes and could contribute private alleles for future use in breeding programmes to expand the tobacco genetic base. Results from genetic analysis showed significant General Combining Ability effects for dark cured leaf colour ($p<0.05$) suggesting the influence of additive gene action on the trait. Reciprocal effects were not significant for all traits meaning that the parents can be interchanged in hybridization programmes. The Hayman Wr-Vr graph indicated partial dominance mode of inheritance for traits measured except dark cured leaf colour. Heritability was moderate for ripening rate and low for growth rate and dark cured leaf colour implying that slow progress can be made when selecting for the later two traits in tobacco improvement programmes. Significant genotype × interaction effects were only recorded for the traits: quality grade, grade index, and mass at untying ($P<0.01$) and not for saleable yield and dark cured leaf colour. The comparison biplot showed that the genotype KE1×AW3R was the most stable for quality. Generally, these results show that although the flue-cured tobacco germplasm in Zimbabwe is not diverse molecularly and morphologically, the little diversity present could, however, be sufficient for exploitation in future variety development programmes. The predominance of additive gene action for most traits implies that selection progress is possible although absence of dominance gene action makes heterosis breeding less important. Significant genotype × interaction for some traits implied that breeders can either select varieties that are specifically adapted to the environments they will be grown or that are stable in performance across tobacco growing environments of Zimbabwe like KE1 × AW3R K326R1-1-28 × XZR2-2-4 for mass at untying.

DEDICATION

This dissertation is dedicated to my wife, Valodia and children, Nicole, Neville and Nicolette with love.

ACKNOWLEDGEMENTS

Firstly, I would like to thank the Tobacco Research Board of Zimbabwe (TRB), for allowing me to enroll for a D. Phil program and providing time and material resources for the study. In particular, my acknowledgements go to the General Manager of the Tobacco Research Board of Zimbabwe, Dr. Dahlia Garwe, for all the financial approvals she did in support of this study. I would also like to thank the TRB Assistant General Manager for Research and Extension, Dr. Susan Dimbi for practically supporting this research effort. Dr. Dimbi was always there to render assistance whenever it was needed especially in the area of monitoring off-station trials.

Let me also take this opportunity to thank my supervisors, Dr. S. Kageler (University of Zimbabwe – Crop Science Department), Dr. S. Dari (University of Zimbabwe – Crop Science Department), Dr. F. Magama (Tobacco Research Board – Plant Breeding Division) and Dr. D. Rukuni (Tobacco Research Board – Crop Productivity Services Division) for accepting to be my mentors and guiding me through the research process, compilation of data and proof reading this thesis. I shall always value your support and guidance in my life. You are leaders, mentors and friends to me. Thank you so much!

My acknowledgements also go to my subordinates in the plant breeding division at the TRB, Mr. G. Magura, Ms L. Vheremu, Mr. M. Shamudzarira, Mr. G. Makombe and Ms E. Mpofu for helping me with trial supervision and management as well as data collection. The process was hectic and tiresome but you toiled for success. The evidence of your success is this document.

This dissertation would not have been a success without the support of Molecular Biology specialists, Dr. N. Muzhinji, Ms F. Mufunda, Mr. T. Sagonda, and Ms. M. Magombo all of TRB's Molecular Services Department of the TRB. You went out of busy schedules rendering me training in the various skills in molecular biology. I would also like to give special mention of Dr. Muzhinji for giving advice now and again and helping me with molecular data analysis.

This piece of work would not have been a success without the input of Dr. E. Gasura in the creation of the proposal and support in genotype × environment, Hayman and Griffings Method IV data analysis. Thank you for all the critiques and adjustments to my original thoughts. To Mr. Stanford Banana, the TRB Biometrician, "Thank you for all the data analysis you helped me do." I also wish to thank the farmers on whose farms I set my multi-location trials for the two seasons of my study namely Mr. Seun van Rensburg, Mr. John Dollar, Mr. Crispen Isah, Mr. Jan Kageler, Mr. Nick Mostart and Mr. Denford Mutwiwa.

Special acknowledgement goes to Mr. Charles Mushore, the Head of Human Resources at TRB for the timely processing of fees and rendering me moral support whenever it was needed.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

AFLP	Amplified Fragment Length Polymorphism
CIMMYT	International Maize and Wheat Improvement Centre
DNA	Deoxyribonucleic Acid
GCA	General Combining Ability
GED	Genotype × Environment Data
PCR	Polymerase Chain Reaction
PCA	Principal Component Analysis
QTL	Quantitative Trait Loci
RAMP	Randomly Amplified Microsatellite Polymophism
RAPDs	Randomly Amplified Polymorphic DNA markers
RFLP	Restriction Fragment Length Polymorphism
SCA	Specific Combining Ability
SCAR	Sequence Characterised Amplified Region
SPAR	Sequence Primer Amplification Reaction
SRAP	Sequence Related Amplified Polymorphism
SSRs	Simple Sequence Repeats
STMs	Sequence Tagged Microsatellites
SVD	Singular Value Decomposition
TIMB	Tobacco Industry and Marketing Board
TMV	Tobacco Mosaic Virus
TRAP	Target Region Amplification Polymorphism
TRB	Tobacco Research Board

UPOV	International Union for the Protection of New Varieties of Plants
USA	United States of America

STRUCTURE OF DISSERTATION

1. Chapter 1: General Introduction
2. Chapter 2: Literature Review
3. Chapter 3: Genetic diversity analysis of Zimbabwean flue-cured tobacco (*Nicotiana tabacum* L.) breeding lines and commercial cultivars using SSR markers
4. Morphological diversity assessment of Zimbabwean flue-cured tobacco (*Nicotiana tabacum* L.) by phenotypic characterization.
5. Genetic analysis for yield, quality and agronomic traits of Zimbabwean flue-cured tobacco (*Nicotiana tabacum* L.) under normal growing conditions.
6. Genotype x environment interaction analysis for yield and quality in elite flue-cured tobacco (*Nicotiana tabacum* L.) varieties in Zimbabwe.
7. General overview of research findings.

CHAPTER ONE

General Introduction

1.1 General background

Tobacco (*Nicotiana tabaccum* L.) is a perennial shrub grown as an annual in almost over 100 countries of the world (Maryan *et al*., 2012; Denduangboripant *et al*., 2010). It originated from the Antilles Islands and the coastal areas of central America (Maryan *et al*., 2012). Although the native people of these areas had started using tobacco in their religious rituals, it was Christopher Columbus' voyage that initiated the popularization of tobacco to the rest of the world beginning with Europe (Tanser, 1991). There are basically four types of tobacco in the world and the classification in based on the method ocuring for each type. The most popular tobacco is Virginia tobacco which is commonly known as flue-cured tobacco because it is cured in barns with heat supplied from the source through flue pipes. Virginia tobacco is used as flavor tobacco in cigarates. The other type of tobacco is Burley tobacco which is cured in open barns using mainly a balance of humidity and natural air temperature. Burley tobacco is mainly used as filler tobacco in cigarate manufacturing. Dark tobacco is cured using the same method as burley tobacco when it is refired as dark aired cured tobacco. When dark tobacco is cured using smoke from selected tree species, it is called dark fire cured tobacco. Dark tobacco is used in the manufacturing of cigars or as cigar wrappers. The last type of tobacco is called oriental tobacco or turkish. Oriental tobacco is a short statured tobacco with very small leaves. It is used as filler and flavourant in cigerate manufacturing. Figure 1.1 shows the vegetative growth stages of the different types of tobacco grown commercially in the world.

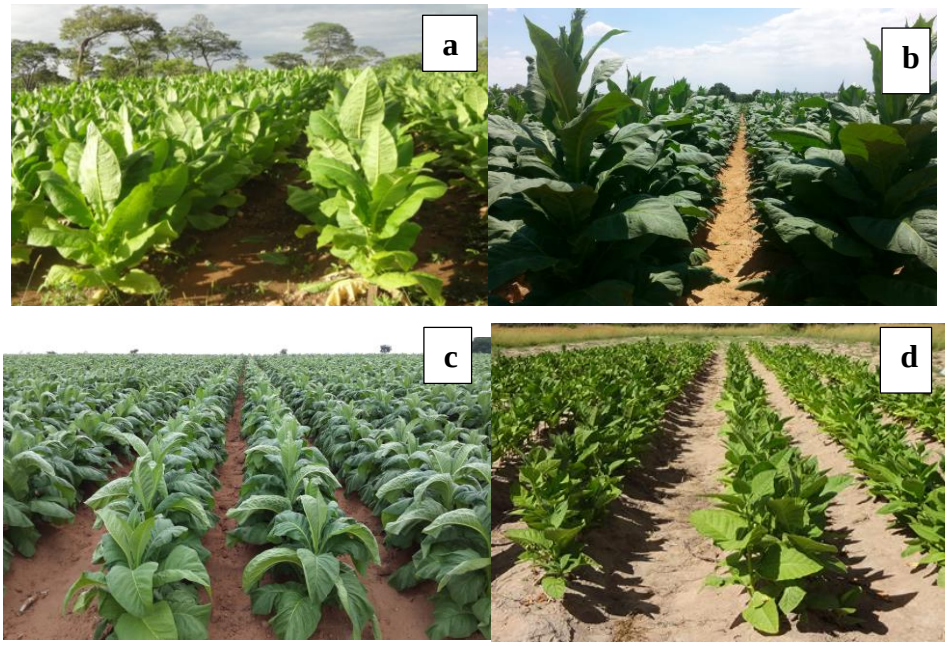


Figure 1.1: The different types of tobacco. a) Burley tobacco; b) Dark tobacco; c) Flue-cured/Virginia tobacco; d) Oriental/Turkish tobacco

The exact dates of the introduction of tobacco to Africa are unknown but it is believed that the crop was brought to the continent through South Africa by the Portuguese explorers and Arab traders and later spread into the interior of Africa (Tanser, 1991). The spread saw the local Africans also growing the crop for their own smoking or use as snuff. In Zimbabwe, commercial tobacco farming started in the early 1890s in Mutare district by Lionel Cripps but it was father Richartz who exhibited the crop for the first time at the 1897 Agricultural Show in Harare (Tanser, 1991; Mbanga, 1991). Since then, the importance of tobacco in Zimbabwe has grown through promotion by both the farmers and the government (Tanser, 1991; Mbanga, 1991). Currently, tobacco has become the major field crop foreign currency earner in the country with all the grown varieties being locally bred.

Early tobacco farming in Zimbabwe relied on variety introductions from the United States of America such as *Sumatra, Virginia Prior, Brazilian Spanish, Bonanza* and *Kentucky Yellow* (Mbanga, 1991). These varieties had the limitation that they lacked adaptation to local growing conditions; hence, they produced poor yields, quality and succumbed to many locally occurring diseases (Tanser, 1991). This resulted in the need to develop local varieties which are adapted to the growing conditions that prevailed in the country.

Breeding programmes were initiated to come up with varieties adapted to local conditions and at the same time possessing multiple disease resistance. The first promising tobacco variety to

be developed locally was *Delcrest* (Tanser, 1991). This variety was not grown for long before it was used in one of the early crosses that resulted in the popular variety *Kutsaga 51* with better adaptation to local growing conditions. Although *Kutsaga 51* was developed, there was need to continue developing even better varieties since it had the limitation that it was susceptible to major diseases and pests of tobacco especially root-knot nematodes (*Melodogyne javanica* L.). Incorporating disease and pest resistance into varieties of a crop creates a genetic barrier against loss of yields and quality and removes the costs associated with chemical control.

1.2 Morphological and genetic diversity in flue-cured tobacco germplasm

Successful crop improvement programmes rely on the availability of genetic variation on the parent populations from which new selections are done (Singh, 2006; Eshghi *et al*., 2012). One of the major hindrances to achieving quick genetic gains in most crop improvement programmes is lack of characterized genetic diversity within a crop to be improved (Darvishzadeh *et al*., 2013; Maryan *et al*., 2012). Lack of heritable variation is the reason why there are seemingly insurmountable plateaus for many traits in different crops. Genetic diversity within a crop species forms a pool from which individual plants with desired characteristics are selected. Broad crop diversity increases the chances of making genetic gains for any trait under improvement. Heritable genetic variation can be inherently present in a population to be improved like in natural populations or it can be introduced (Lahiji *et al.*, 2013). However, because of continuous intense selection driven by stringent crop trait preferences by consumers, variation has diminished in many crops of economic importance to man (Moon *et al*., 2009a).

For the breeders to appreciate the extent of variation within the populations they wish to improve for a specific trait or traits, the material constituting the populations have to be evaluated for morphological traits. This is because phenotypic selection is the first step before molecular markers are employed since specific markers may not exist for some traits like plant architecture and plant canopy shape as well as leaf texture in tobacco. Where variation is limiting in a population under improvement, it is created through exotic line introductions, mutation induction, inter and intra-specific hybridization and possibly through genetic engineering (Singh, 2006). All the other approaches have the major advantage that the variation created would be adapted to local growing conditions but with plant introductions, adaptation to local conditions can be limiting unless a variety is selected from the originally introduced variety or line. However, all these approaches have to be explored for variation to be created

so that selection can be effective in the effort to create new and better varieties of crops. In Zimbabwe, there has been continuous tobacco germplasm collection and use for over a century without any effort to understand its diversity by way of characterization using either, morphological or molecular methods. This has been the situation despite the fact that knowledge of such information is important for current and future breeding efforts as it guides selection and mating patterns in breeding programmes.

1.3 Understanding of the inheritance of key traits in flue-cured tobacco

Early studies have proved that major traits of economic importance to the farmer like grain yield in cereals, leaf yield in tobacco, disease resistance in most crops and quality are controlled by many genes, each of which exerts a small influence to the final phenotypic performance of the crop (Dabholkar, 1999; Singh and Chaudhary, 2004; Fricano *et al*., 2012). This presents a problem to the plant breeder because it is very difficult to achieve high genetic gains for these traits in variety improvement programmes. There is also evidence that some disease resistances are controlled by single major genes like tobacco mosaic tabamo-virus (TMV) and white mould in tobacco. The number of genes controlling the trait's phenotypic expression, the level of linkage of the genes involved and the extent to which the environment has an influence on the trait's phenotypic expression determines heritability of trait (Singh and Chaudhary, 2004; Patel *et al*., 2012). Some traits are dominant while others exhibit additive or recessive tendencies. In some traits in cereals, maternal effects have been shown to play a key role in the behaviour of offspring produced. Plant breeders have to know the inheritance of target characters in crops they work with in order for them to decide on the chances of success in their crop improvement initiatives. An understanding of the inheritance of characters in crops also allows the plant breeder to carryout informed decisions about what crosses to make to get the desired results expeditiously.

The early tobacco breeding work in Zimbabwe did not rely on the basic understanding of the inheritance of traits under investigation. Rather, it relied on making crosses and conducting field evaluations to assess the presence or absence of traits of interests (TRB Annual Report, 1961). However, in plant breeding work, understanding the mode of inheritance and the heritability of traits of interest helps in determining the ease with which traits can be passed on from parents to offspring (Butorac *et al*., 2004; Patel *et al*., 2012). In addition, such an understanding gives the breeder a clue on the appropriate methods of breeding to use in order to quickly achieve the desired results (Patel *et al*., 2012). Knowledge of the genetics and expression of traits under different conditions helps the plant breeder separate the effects of the

environment on the behavior of genes from the true expression of genes under study, hence, improving the relative progress and efficiency of the selection process. Traits that are non-genetic are not heritable and any selection efforts on such characters yield no genetic gains (Farshadfar *et al*., 2012). Genetic studies on tobacco were initiated in the early 1960s by Schwappenhauser who investigated resistance of tobacco to root-knot nematodes (TRB Annual Report, 1961). Through the exploitation of accumulated variation from introductions, selections with resistance to root-knot nematodes were developed and it was also established that the resistance to the root-knot nematodes exhibited partial dominance wherever it occurred. Relying on the foundation of the work of Schwappenhauser, tobacco varieties with resistance to root-knot nematodes were made available to growers. Other disease resistances were later bred for and these achievements are mainly attributed to the gains in yield and quality that the tobacco industry in Zimbabwe has recorded to date.

1.4 Genotype × environment interaction in flue-cured tobacco and cultivar adaptation

The major and direct beneficiaries of crop varieties are the farmers; therefore, varieties must meet the farmers' expectations in terms of yield performance and agronomic management requirements (Allard, 1999; Singh, 2006). Farmers grow crop varieties in different climatic, biotic and edaphic conditions, which exert different negative pressures on their crops. The intensities and durations of these pressures are neither static nor predictable because of seasonal differences and changes in fertility regimes as determined by individual farmer management practices (Gauch and Zobel, 1988). In different environments, varieties behave differently resulting in high performers in one area and season being poor performers in other areas and seasons (Gauch and Zobel, 1997). These genotype × environment interactions as they are known in plant breeding have to be anticipated and factored in during variety development. Many studies have shown that crop varieties behave differently under different growing conditions (Comstock and Moll, 1963; Eberhart and Russel, 1966; Lin and Binns, 1988; Ciccarelli, 1989; Yan and Kang, 2003; Gul *et al*., 2014) resulting in the recommendation that crop growing environments be subdivided according to their similarities in ability to support cropping systems (Gauch and Zobel, 1997). Environments with similar abilities to support different crops have been found to be transcontinental in some cases hence the coining of the term "mega-environments" by breeders at the International Maize and Wheat Improvement Centre (CIMMYT) (Gauch and Zobel, 1997).

In Zimbabwe, land use planning and agriculture specialists also classified the country according to farming potential based on soil and vegetation characteristics into agro-ecological

zones where different farming ventures are recommended (Rukuni and Eicher, 2006). Within the traditional agro-ecological regions recognized in Zimbabwe, tobacco has its own recognized growing belt, which is divided into three regions: the fast, the medium, and the slow growing regions (Tobacco Research Board, 1999). Classification in the tobacco-growing belt is based on altitude and temperature regimes experienced in the areas since soil characteristics and fertility regimes are similar. However, close analysis has indicated that even within the broad demarcations of growing regions observed in the tobacco growing regions, microenvironments do occur which require special treatments by growers for them to get the best tobacco yields (Sadeghi *et al*., 2012). Plant breeders recommend that mega-environments be delineated mainly based of their discriminating ability on varieties since biotic, edaphic and climatic conditions alone may not be sufficient pointers of the ability of an environment's ability to support different crop varieties (Gauch, 2006). So, varieties have to be physically grown in specific environments over more than one season in order for their behavior to be assessed and appropriate recommendations to be made about the environments in which they perform best. If new mega-environments are identified within the common growing belts of flue-cured tobacco in Zimbabwe, environmental delineation will increase overall yield of tobacco.

1.5 The Main Objective of the Study

The main objective of this research was to perform genetic characterization and analysis for yield and quality related traits in the Zimbabwean flue-cured tobacco germplasm.

The specific objectives were to:

1. Establish the genetic relatedness among elite Zimbabwean tobacco lines using Simple Sequence Repeats markers,
2. Determine the genetic diversity of Zimbabwean flue-cured tobacco germplasm by phenotypic characterization,
3. Investigate the mode of inheritance of growth rate, ripening rate and dark cured leaf colour in the Zimbabwean flue-cured tobacco germplasm, and
4. Determine yield performance and stability of tobacco hybrid varieties across the three growing environments in Zimbabwe.

The hypotheses of the study are:

1. There are differences among the elite Zimbabwean flue-cured tobacco germplasm at molecular level.
2. There is narrow agro-morphological diversity among the Zimbabwean flue-cured tobacco germplasm.
3. There is no known mode of inheritance of growth rate, ripening rate and dark cured leaf colour in the Zimbabwean flue-cured tobacco germplasm.
4. There is no stability for yield performance and stability among flue-cured tobacco hybrid varieties across the three growing environments in Zimbabwe.

CHAPTER TWO

Literature Review

2.1 The origin of tobacco and its implications on global tobacco germplasm variability

The origin, domestication and early cultivation of all tobacco are traced back to North and South America (Raju *et al*., 2009; Moon *et al*., 2009a). Tobacco originates from a cross of *N. tomentisiformis* (2n = 24) and *N. Sylvesris* (2n = 24) (Davalieva *et al*., 2010). It is, therefore, an amphidiploid ($2n = 4x = 48$) that is naturally self-pollinating but able to be hybridized through emasculation and the use of male sterility. Variations in tobacco types currently available have come about through mutations and selection by plant breeders. Most of the currently grown tobacco cultivars have their origins in the United States of America where the tobacco farming business was popularised earlier than other parts of the world. Because tobacco has one centre of diversity and that national breeding programmes share germplasm, there is limited variation in the current elite tobacco breeding programmes. Most tobacco varieties originate from crosses that involve more than one line, which can be traced back to the USA (TRB Annual Report, 2013; Moon *et al*., 2009b). A case in point is the line K 326 whose variants are used in countries like Zimbabwe, China, Brazil, South Africa and many other tobacco growing countries but having originated from America. Because of the use of common parents across national breeding programmes, most varieties are now similar across the whole world. Many early researchers have pointed out that there is limited variation in flue cured tobacco and attribute the lack of variability to continuous selection within a narrow genetic base being done by tobacco breeders (Moon *et al*., 2009a).

2.2 Tobacco germplasm diversity and its utility in tobacco variety development

Since the rediscovery of Gregor Mendel's findings on inheritance and the publication and acceptance of the principles of independent assortment and recombination, it is agreed that crop variety improvement depends on the presence of variability among breeding populations (Moon *et al*., 2009a; Ahmed *et al*., 2014). Wide genetic differences between any parental materials used in complex crosses give high chances of creating new useful recombinants for selection to act upon. In modern plant breeding, breeders use complex crosses and introductions to create variable populations from which they select promising material for further improvement. The tobacco germplasm is regarded as not widely variable because its origin is traceable to one centre of diversity (Liu *et al*., 2009). In addition, the major drivers of the tobacco industry in the world are few and they dictate the crop characteristics that suit their

specific needs (Moon *et al*., 2009a). As a result, tobacco national breeding programmes globally have sought to improve a few traits chief among them being cured leaf yield, quality and nicotine content. Although, different tobacco breeding programmes across the world have sought to develop varieties adapted to local farming conditions, their objectives are similar and centred on preserving specific traits in the crop. For instance, keeping leaf chemistry levels within specific acceptable limits has made tobacco breeders work with germplasm of limited diversity (Moon *et al*., 2009a) and this has led to breeding for other traits of economic importance in tobacco that are closely linked to nicotine level for example. The fear to dissipate the already recorded gains in breeding for nicotine content means that most tobacco breeders across the world work with tobacco germplasm of narrow diversity (Moon *et al*., 2009a). Such a scenario in tobacco farming has resulted in the present yield plateaus, which lie at 4.5 t/ha in Zimbabwe and 3 t/ha in most tobacco producing countries (TRB Annual Report, 2013). To surpass the plateaus, tobacco breeders have to exploit the potential genetic variability present in the tobacco germplasm worldwide.

2.3 Generating useful variability information in plant breeding

Most studies on diversity in plants and other living organisms seek to get general information about the extent of differences within the populations under study. That is, whether the population is highly heterogeneous or not (Maleki *et al*., 2011; Fricano *et al*., 2012; Lahiji *et al*., 2013). Plant breeding exploits variation that is present for specific traits in the improvement of crop varieties (Liu *et al*., 2009; Thakur *et al*., 2013). Most variation that is expressed phenotypically in most plant species is not useful variation for crop improvement as it is not heritable. The majority of genetic and phenotypic diversity that exists in tobacco is not useful in variety development because it does not bring about noticeable breeding value into elite breeding material. Most molecular studies on diversity in breeding crop species end by stating that either the lines or populations under study are highly variable or the opposite without recommending the lines to combine if a breeder is to obtain trait improvement in the crop species (Davalieva *et al*., 2010). Studies on diversity and variation in plants must bring out information that help the breeder choose line combinations that result in better variety development. Such information could be in the area of disease resistance, harvest index, phenology, plant architecture, or any other traits that can easily be selected for by the breeder. It is, therefore, important that when experiments to study diversity are set out, breeders give thought on the final use to which the generated information will be put.

2.4 An overview of the possible approaches to studying diversity in crop plants

Diversity studies in plants can be traced back to the work of the Russian botanist Nicoli Vavilov who observed and documented extensively phenotypic differences in plant species across continents. Variation is important for survival of a species as it confers fitness across a wide range of selection pressure. Wide diversity among crop species forms a broad base on which artificial selection by the breeder can be done (Darvishzadeh *et al*., 2013). Since the time of Vavilov, variation among plants has been studied extensively especially in economically important plants for ecological reasons mainly (Allard, 1999; Maryan *et al*., 2012). The traditional methods of studying diversity relied on morphological differences among the studied individual plants. The approach is premised on the understanding that genes determine the phenotype that is used in plant morphological characterisation. It fails, however, to factor in the influence of the environment on the phenotype. The method is also laborious and expensive as it requires extensive field testing and evaluation of plants (Lahiji *et al*., 2013). Nevertheless, this approach is still useful even to date especially in practical conventional plant breeding, which is hinged on field evaluation of germplasm before crossing. The major strength of this method is that it gives the breeder a chance to assess the traits of interest *in situ* factoring in potentially undesirable traits.

Modern methods of assessing diversity in living organisms are now hinged on a relatively new dimension of molecular genetics called genomics and uses molecular markers to distinguish genotypes (Moon *et al*., 2009b). Many molecular markers have been developed and are very sensitive being able to distinguish between different genotypes at DNA level like the single nucleotide polymorphisms (SNPs). More precise and accurate molecular markers now rely on whole genome sequencing to identify specific gene locations to determine the genetic basis of the phenotypic variation that exists among organisms of the same species. Although molecular markers are accurate they only help to the point of showing the broadness of differences among genotypes under study. How usable the material is depends on the phenotypic traits of interest, which are usually evaluated in the field. Molecular markers are very precise, and sensitive, distinguishing between genotypes at DNA level but some of the differences picked may not be exploitable by the plant breeder at phenotypic level as they can be easily confounded by the environmental factors. Because of this weakness, it is important for the breeder to use morphological approaches of distinguishing genotypes as well since they can reveal differences, which can easily and quickly add value to the farmer. More effort, however, has

to be directed at standardising the testing environment to reduce its non-heritable influence on the traits of interest.

2.5 Use of molecular markers in crop diversity studies

Since the popularization of the use of molecular markers in the early 1980s studies aimed at improving their efficacy have greatly increased. The results of the improved studies have led to an increase in the application of molecular techniques in solving plant improvement challenges. As a result of constant modification and upgrade, molecular marker technology is also now being used in genome analysis, gene tagging, phylogenetic analysis and selection of desirable genotypes for use in crop improvement programmes (Jonah *et al*., 2011). Before the introduction of molecular markers in diversity studies, plant scientists relied on phenotypic and morphological traits to distinguish genotypes. However, basing diversity studies on phenotypic and morphological traits is unreliable as these traits vary with environments (Jonah *et al*., 2011). In addition, plants have to grow to full maturity before they can be identified.

Molecular markers offer many advantages over morphological markers in crop improvement. The major advantages associated with molecular markers are to do with their objectivity owing to their environmental insensitivity (Jonah *et al*., 2011). Moreover, molecular markers shorten the time required to develop varieties since varieties, lines or genotypes can be distinguished even at the earliest stages of development. This feature of molecular markers also reduces the cost of developing varieties even though the initial cost of establishing molecular technology could be high.

2.6 Important properties of molecular markers applicable in diversity studies

Put simply, genes are arranged in a linear pattern on chromosomes. Closely linked genes are inherited together since they occur on the same chromosome. Some genes have been found to flank specific regions of a chromosome responsible for the inheritance of traits of interest. These individual genes, which have been found to occur at close intervals, are the ones known as molecular DNA markers. According to Piyasatian and Kinghorn (2003) molecular markers are identifiable DNA sequences found at specific locations of the genome and associated with the inheritance of a specific trait. However, Thottapilly *et al*. (2009) have broadly described molecular markers as naturally occurring polymorphisms, which take into consideration proteins and nucleic acids that are distinguishably different. Molecular DNA markers have, however, been used extensively in plant breeding to identify genes of interest in segregating populations in order to speed up breeding programmes. One very important property of markers

is that they must exist in different forms, which allow for a distinction between the normal and the mutant forms of the same gene. Through what is now commonly known as DNA fingerprinting many single locus detection systems are being used as versatile tools for studying many aspects of plant genomes which include genetic variability, genome fingerprinting, genome mapping and genome evolution among other disciplines (Jonah *et al*., 2011). In order for markers to be efficacious, the major properties they must exhibit include high polymorphism, co-dominance inheritance, frequency of occurrence in the genome of interest, selective neutrality to environmental background, easy accessibility, high reproducibility, easy and fast assay as well as easy exchange of generated data between laboratories (Joshi *et al*., 2010). It is, however, very difficult to get a marker that has all these properties. Choice of molecular markers for use must therefore be based on the ability of the marker to meet more of the properties.

2.7 Classification of DNA based markers

The discovery of the polymerase chain reaction (PCR) by K. Mullis in 1982 greatly improved the usefulness of in crop variability studies (Joshi *et al*., 2011). The discovery led to the separation of molecular markers as either PCR based or non-PCR based. Non-PCR based molecular markers include the Restriction Fragment Length Polymorphism (RFLP) (Joshi *et al*., 2011). RFLPs are simply inherited and naturally occurring. Variations occurring in the DNA sequences on the chromosomes are the basis of genetic diversity within species (Jonah *et al*., 2011). The variations are a result of simple mutations in the form of inversions, translocations, transpositions or deletions occurring on chromosomes resulting in loss or gain of fragments of different lengths (Jonah *et al*., 2011). PCR involves the amplification of DNA molecules in vitro. DNA fragments from different individual plants produce different amplifications so; specific DNA fragments generated from one plant but not from others represent DNA polymorphism can be used as genetic markers.

Joshi *et al*. (2011) stated that PCR based markers include the Random amplified polymorphic DNA markers (RAPD), which detects nucleotide polymorphism in DNA using a single primer of arbitrary nucleotide sequence; Amplified Fragment Length Polymorphism (AFLP), which combines the characteristics of RAPDs and RFLPs, and the simple sequence repeats (SSRs) or short tandem repeats or micro-satellites. SSRs have for long been regarded as ideal genetic markers for detecting differences occurring between and among species of all eukaryotes (Jonah *et al*., 2011). It consists of tandemly repeated 2-7 base pair units arranged in repeats of

1, 2, 3, 4, or 5 nucleotides with different sizes of repeat sequences. The repeats have been found to be commonly occurring throughout the plant genomes that exhibit high levels of variation based on the differences in the number of tandemly repeating units of a locus (Jonah *et al*., 2011). The variation in the number of tandemly repeated units result in highly polymorphic banding pattern, which are detected by PCR using locus specific flanking region primers where they are known. SSRs are co-dominant sequence tagged microsatellites (STMs) markers (Joshi *et al*., 2011). They are highly reproducible such that they can be used efficiently by different research laboratories to produce consistent data (Jonah *et al*., 2011). The other microsatellite based markers are the Randomly Amplified Microsatellite Polymophism (RAMP) (Agarwal *et al*., 2008), Sequence Characterised Amplified Region (SCAR), Sequence Primer Amplification Reaction (SPAR), Sequence Related Amplified Polymorphism (SRAP) and the Target Region Amplification Polymorphism (TRAP) (Agarwal *et al*., 2008).

2.8 Yield and yield determinants in tobacco

Tobacco is a C_3 plant whose organ of economic importance is the leaf. In tobacco farming, the common practice is to nip the apical bud at some point to prevent the plants from setting flowers, a practice known as "topping". The stage at which topping is done varies from country to country but in Zimbabwe, tobacco topping is done at 18 leaves (Flue Cured Tobacco Recommendations, 2010). Topping promotes enhanced leaf expansion and dry matter accumulation (Davis and Nielsen, 1999). Growers prefer varieties that exhibit pronounced leaf expansion and dry matter accumulation. Tobacco cultivars differ in their ability to add on weight after topping although management practice and environmental factors such as temperature and humidity also play a part (Ya-Nan *et al*., 2012). Parajuli *et al*. (2015) reported that, holding environmental factors and management practice constant, the yield of tobacco is determined by the number of plants per hectare, number of leaves per plant and the mass of each leaf per plant. Number of plants per hectare (15 000) and the number of leaves per plant (topped to leave 18 leaves) are constant in Zimbabwe. The only variable becomes dry matter accumulation per leaf per plant. Not ruling out the contributions of good standard agronomic practices and improved pest and disease management, selection for dry matter accumulation has contributed over 700% yield gain in Zimbabwe over the past six decades of tobacco research (TRB, 2010). Through the exploitation of parameters such as leaf duration, high cured leaf yields have been recorded with varieties like K E1 coming on board for adoption by farmers in the early 1960s (TRB Annual Report, 1962). There is potential variation for dry

matter accumulation in the existing flue-cured tobacco that can be exploited in the development of even better yielding varieties than those currently available on the market.

2.9 Breeding for improved cured leaf yield in flue-cured tobacco

Selection for yield in tobacco started with the beginning of tobacco cultivation (Maleki *et al*., 2011). In Zimbabwe selection for improved yield started with the introduction of the crop in the country (Tanser, 1991). Early varieties had limitations of possessing low leaf numbers, reduced leaf duration and leaf size. Parajuli *et al*. (2015) found number of leaves to be positively correlated to leaf dry mass. The parameter leaf expansion was used to measure leaf size of early varieties. Because leaf yield is a quantitative trait, breeding for this trait is slow. Djulgerski (2010) reported that leaf number is mono to olygogenic, controlled by 1-12 genes. Parajuli *et al*. (2015) also reported that this trait is inherited overdorminantly in favour of the parent producing the higher number of leaves with a high heritability of over 84%. In studies by Djulgerski (2010) and Maleki *et al*. (2011) it was also pointed out that leaf size is governed by additive genes. On the other hand, Sayyed-Nazari *et al*. (2016) reported additive gene action for fresh and cured leaf yield, number of leaves per plant, and the length and width of the leaves. Not much work has been done to date on the inheritance of leaf duration in flue-cured tobacco but it is hypothesised that it has low heritability since it is a polygenic trait (Maleki *et al*., 2011). When a trait has low heritability, its response to selection is low and unpredictable. Cured leaf dry mass is determined by more than one trait, each controlled by many genes (Maleki *et al*., 2011). This makes selection for it very difficult thus; realising yield improvement for dry leaf yield is difficult unless there is thorough control of the environment. It therefore means that to gain on yield, the farmer has to intensively mane a tobacco crop.

2.10 Cured leaf quality in flue-cured tobacco

Quality is simply defined as suitability for the intended purpose of a product. It determines the usefulness of a product to the end user. In tobacco farming, quality is a very important parameter as it is a major determinant of the value of tobacco. Tobacco quality is determined by the leaf nature and leaf finish (TIMB, 2010). Leaf nature refers to the oil content and the pliability of a flue-cured leaf while leaf finish is the extent of any blemishes in the form of leaf breakages, chaff or waste. Leaf colour also contributes to the quality of the tobacco and the major colour grades are lemon, orange and mahogany. Quality is given a score, which ranges from 1(the best quality) to 5 (the worst quality) in flue-cured tobacco. Cured leaf quality is a qualitative trait and is difficult to study in tobacco since it incorporates more than one property of the leaf. A grade index that takes into consideration leaf nature and leaf finish was developed

in the USA and also adopted in Zimbabwe to determine the acceptability of flue-cured tobacco leaf to the merchants (Tobacco Research Board, 1999). Although some tobacco breeders believe that acceptability score and presence or absence of blemishes is associated with genotypes, it is generally believed that these components of tobacco quality are to do with handling of the crop at and after harvesting. This would mean that they are purely environmental and management style dependent (Ya-Na *et al*., 2012). The only component of flue-cured tobacco quality that can be connected to the genotype is cured leaf colour/style (Figure 2.1). Variation among genotypes in their ability to produce different levels of a specific colour has been claimed in flue-cured tobacco breeding in Zimbabwe (TRB Annual Report, 1996). It was generally found that highly green pigmented, heavy bodied lines give orange and mahogany cured leaf styles. On the other hand, pale and thin leafed breeding lines easily give lemon cures (TRB Annual Report, 1996). Through the influence of the environment and field management, modifications in cured leaf styles cannot be ruled out. Little has, however, been done to study the inheritance of flue-cured leaf colour among the lines used in tobacco breeding programmes. Most of the hypotheses made in this area have always been based on field observations without scientific confirmations.

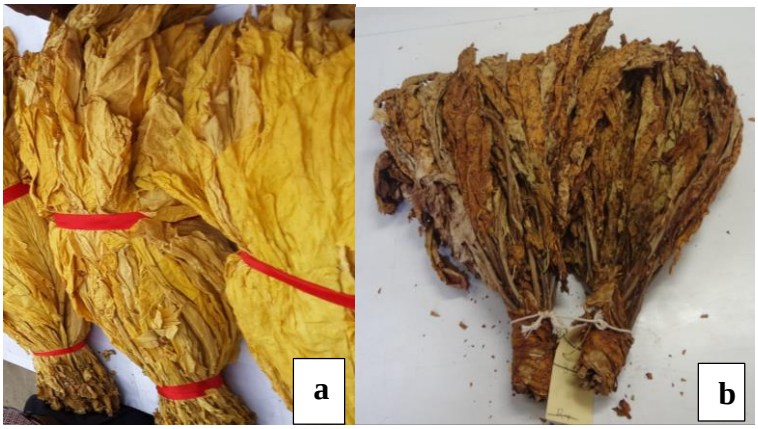


Figure 2.1: Lemon (a) and dark cured (b) styles in flue-cured tobacco

2.11 Breeding for cured leaf quality and colour in flue-cured tobacco

Early studies on cured leaf colour in flue-cured tobacco showed that it is controlled by oligogenes (Ahmed, 2014). These genes determine the chlorophyll content of the leaf, which in turn determines the flue-cured leaf colour produced after curing. Since leaf chlorophyll content is oligogenic, its breeding is relatively simple. Little work has, however, been done to study the inheritance of cured flue-cured tobacco leaf colour.

Breeding for quality is also very important since leaf quality and colour determine the price of flue-cured tobacco leaf on the market (Ya-Nan *et al*., 2012; Ahmed, 2014). Ya-Nan *et al*. (2012) reported that flue-cured tobacco quality is determined by the sugar-nitrogen balance and they observed that epistatic gene action play a large part in the control of this trait. In Zimbabwe, recommended sugar content of tobacco varieties is +/- 15% of total dry matter content while total nitrogen should be +/- 10%. The environment has also been shown to exert its influence as well so breeders have to take genotype × environment interaction into consideration when tobacco quality experiments are set out.

2.12 Leaf ripening and its implications on flue-cured tobacco yield and quality

Leaf ripening is a very important trait in flue-cured tobacco farming. When flue -cured tobacco matures, its leaves change colour from dark green to pale green. As the pale green colour turns yellow, tobacco is regarded as ripe. Tobacco varieties are designed to ripen at a rate of two leaves per week once ripening has started (Ya-Nan *et al*., 2012). Varieties that have more than two leaves ripening in a week are said to be fast ripening while those that have less than two leaves ripen in a week are classified as slow ripening. Fast ripening varieties are low yielders because they have less time for dry matter accumulation in the field (TRB Annual Report, 1996). These varieties are usually characterised by thin leaves which easily breakdown into trash during handling. Fast ripening varieties generally give lemon cures. On the other hand, slow ripening varieties give high cured leaf yield. It is hypothesised that this is because slow ripening varieties have more time to use nutrients available to build on weight. These varieties are usually thick bodied and dark green in colour at vegetative development. Such varieties do not easily break and have been found to give orange to mahogany cured leaf styles. Because they are easy to handle in the lands and after curing with less breakage and easily give high proportions of orange cures, growers prefer slow ripening varieties. In addition, slow ripening varieties are easy to handle at curing since they give farmers enough reaping interval which reduces barn capacity pressure (TRB Annual Report, 1996). The flue-cured leaf tobacco merchants in Zimbabwe have also indicated that they prefer relatively dark cured flavor style cured leaf tobacco which comes from the slow ripening varieties as they fetch relatively high prices on the market. Fast ripening varieties give farmers reaping pressure as curing systems need at least seven days to fully complete a curing process (Ahmed, 2014). If tobacco ripens at intervals of more than two leaves in a week, the curing space becomes limiting and yield and quality is lost due to over ripeness.

2.12.1 Breeding for ripening in flue-cured tobacco

Ripening in tobacco is a polygenic trait (Ahmed, 2014) and selection for ripening in flue-cured tobacco is practiced in many countries although there is not much work published in this area. In Zimbabwe, germplasm used in developing varieties of flue-cured tobacco is already grouped into slow ripening, fast ripening and intermediate ripening (TRB Annual Report, 2012; TRB Annual Report, 2013). This suggests that variation for ripening exists within the flue-cured tobacco germplasm. Through selection and evaluation, it is possible to develop slow ripening high yielding varieties giving orange cures and fast ripening low to medium yielding varieties giving lemon cures.

2.13 Growth rate and its implications on flue-cured tobacco yield and quality

Growth rate in tobacco farming refers to the time it takes for a tobacco plant to fully attain 18 leaves from the day of planting. Attainment of 18 leaves means that the variety is ready for the breaking of the apical bud through a process called topping (TRB Annual Report, 2012). The time taken to attain the topping height is correlated with biomass build-up. Slow growing varieties have been reported to give relatively high cured leaf yield than fast growing varieties. They have also been shown to give orange and mahogany cured leaf styles easily (Sayyed-Nazari *et al*., 2016). On the other hand, fast growing varieties have been found to give relatively thin leaf which give low yields of mainly lemon style cures. Growth rate is important in tobacco farming especially when a grower wishes to grow an irrigated and a dryland crop in the same season. In Zimbabwe, fast growing varieties are planted during early September as irrigated crops for them to grow fast and be taken out of the field and barns early to pave way for the slow growing and ripening and late maturing dry land varieties (TRB Annual Report, 2013).

2.13.1 The possibility of breeding for growth rate in flue-cured tobacco

Breeding for growth rate in flue cured tobacco is very important as it has the potential to solve the challenge of choice of varieties for planting early in September as irrigated crops (Ahmed, 2014). Very little documentation on the growth rate of tobacco has been recorded except blanket classification of varieties and breeding material either as fast, intermediate or slow growing. From these informal classifications by tobacco breeders and growers, it can be hypothesised that there is exploitable variation for growth rate in flue-cured tobacco (Parajuli *et al*., 2015). In Zimbabwe, germplasm is also grouped according to the above classifications and described and promoted for adoption by farmers using the same criteria. However, there is no tangible evidence of deliberate effort to investigate genetic properties of this trait yet

knowledge of traits like heritability, maternal influence, and nature of gene action would be useful to the plant breeders who want to improve tobacco germplasm for the trait (Sayyed-Nazari *et al*., 2016).

2.14 Heritability and mode of gene action controlling growth and ripening rate in flue-cured tobacco

Heritability is a measure of the extent to which a trait can be passed on from parents to offspring (Shah *et al*., 2015). It is influenced by the number of genes and nature of gene action involved in the control of the trait. Knowledge of heritability of yield and its components in crop plants is vital since it gives an indicator of the influence of the environment on a trait of interest (Parajuli *et al*., 2015). This in turn gives an idea of the extent to which selection would be able to improve a character. In flue-cured tobacco, varieties that exhibit either slow or fast growth and ripening rate have been developed consciously in Zimbabwe indicating that the traits are heritable. Preliminary evidence, however, suggests that the traits could be polygenic (Ahmed, 2014). In general, polygenic traits give low heritability in many crop species and Parajuli *et al*. (2015) reported that plant height and leaf colour is polygenic hence they have low heritability. However, Djulgerski (2010) who studied the gene behaviour and inheritance of leaf number suggested that dominance gene action in favour of the parent with high leaf number was at play. Patel *et al*. (2012) also reported that leaf number has high heritability in tobacco while Liu *et al*. (2011) found that dominance and additive effects play a major role in the control of this trait.

2.15 The existence and possibility of exploitation of heterosis in flue-cured tobacco

Singh (2006) defined heterosis as the superiority of the F_1 hybrid over either the mid-parent or the better parent in a cross. Heterosis varies for different traits in different cross combinations (Patel *et al*., 2012). Patel *et al*. (2012) also reported positive heterosis for leaf number, internode length, and plant height in flue cured tobacco. On the other hand Gixhari and Sulovari (2010) found heterosis of between 2% and 14% in flue-cured tobacco for yield. The main reason for creating hybrids in plant breeding is to exploit heterosis. Heterosis has been widely exploited in cereals like *Zea mays* L and *Sorghum bicolor* L. In crops whose economic parts are not grain but vegetative like tobacco and some brassicas, it has been reported that heterosis is not important (Ahmed *et al*., 2016). However, recent work on flue-cured tobacco suggests that heterosis exists for flue-cured tobacco cured leaf yield (Ding *et al*., 2008). In the past the main reason for which flue-cured tobacco hybrids were constituted was the pyramiding of disease

resistances, which would be contributed independently by each of the constituent parents. If heterosis for yield and other traits could be proved in flue cured tobacco, it means that better yielding varieties can be constituted to satisfy farmer needs. Ding *et al*. (2008) found good combining ability for yield and quality in flue cured-tobacco. Liu *et al*. (2011) also reported both general combining ability (GCA) and specific combining ability (SCA) effects for total sugar and reducing sugar content in flue-cured tobacco. All these reports suggest the presence of heterosis for other traits in tobacco. Where there is heterosis it means that combining ability studies can be done and best line combinations that result in superior varieties determined.

2.16 Genotype × environment interaction in flue-cured tobacco production

Genotype × environment interaction in crop production is a situation where the performance of a variety varies depending on the growing conditions to which the crop has been subjected. Numerous studies on genotype × environment interaction have been carried out in cereals and plantation crops (Salmon, 1951; Heslot *et al*., 2014; Saddoud *et al*., 2012). On the other hand, very little or no interaction studies have been done in crops whose harvestable parts are vegetative (Ahmed, 2014). The few early studies carried out on tobacco reported the absence of genotype × environment interaction in the crop. Studies by Ahmed (2014) have also reported no changes in rankings among tobacco varieties tested across different sites in Pakistan. However, significant genotype × environment interactions was found for fresh leaf weight. Many recent studies in Asia on different types of tobacco have also indicated the significance of genotype × environment in the crop. For instance, Ya-Nan *et al*. (2012) found significant genotype × environment interaction for total and reducing sugars in flue-cured tobacco. Ahmed (2014) also reported significant genotype × environment interaction effects for flue-cured tobacco quality and non-significant genotype × environment interaction for grade index. This was despite the fact that tobacco breeders recommended specific varieties for specific environments and also that growers preferred certain varieties against others in their respective growing areas citing poor performance (TRB Annual Report, 2009). In Zimbabwe, tobacco-growing areas are classified into slow, medium and fast growing environments and varieties are developed targeting each of these environments (TRB Annual Report, 2009; 2013). This acknowledges the presence of genotype × environment interaction and the need to scientifically study its implications on the major traits of the crop for better exploitation of it.

2.16.1 Exploiting genotype × environment interaction in tobacco agriculture

Genotype × environment interaction information is important to both farmers and plant breeders. Growers need to know the best varieties for their locations and farming programmes while plant breeders require such information for designing successful variety improvement programmes (Gauch and Zobel, 1997). Plant breeders also use genotype × environment interaction information for identifying ideal testing and breeding sites for varieties (Mohamed, 2013). However, for the information on genotype × environment interaction to be useful to either the plant breeders or the farmers who use developed varieties in different cropping programmes, it has to be presented in an unambiguous manner (Yan *et al*., 2000). For the farmer, a "which won where" graph is enough to help determine the best variety for their locations in a specific season for them to exploit specific adaptation of varieties. Gauch (1988) pointed out that broadly adapted varieties are not useful where the farmers aim at maximising returns since most broadly adapted varieties are average yielders. However, broad adaptation acts as a buffer against unpredictable weather conditions. Stability is another trait that is very important to farmers as it can be used as a measure of the resilience of a variety across a wide array of seasons (Yan *et al*., 2000). On the other hand, plant breeders want to know information about the existence of mega-environments and the possibility of targeting mega-environments in their breeding programmes.

2.16.2 The development of better understanding of genotype × environment interaction in crops

Even before the discovery of genes in the control of heredity at the turn of the 20th century, early farmers had already realised that varieties performed differently in specific growing environments (Heslot *et al*., 2014; Vargas *et al*., 2006). Both farmers and researchers had the challenge of choosing the best varieties for specific growing conditions and those varieties, which would give the best average performance across all the growing conditions (broad adaptation). Although genotype × environment interaction was evident, the problem was on how to scientifically detect it and its magnitude in different crops (Haun, 1982). Location and year/season variations are evident when different crops are grown in different and highly variable areas across more than one season. Early researchers in plant sciences therefore put across many possible ways of studying genotype × environment interaction in an effort to exploit the best performance of specific varieties in specific environments or across environments and years (Haun, 1982; van Eeuwijk, 1996; Ramburan *et al*., 2011).

The early efforts to understand the concept of genotype × environment interaction and its influence on crop varieties were made by Sprague and Federer in 1951 (Carvalho *et al*., 2014). Sprague and Federer showed how variance components might be used to separate the genotype main effects, environment main effects and interaction effects using the analysis of variance (ANOVA). Later on Plasted and Peterson (1959) proposed the use of the combined analysis of variance in the study of the influence of the genotype × environment interaction on genotype performance. This is an analysis that takes into consideration all the varieties under study at all the locations tested in a specific season. In the period between 1951 and 1959, many workers studied effects of genotype × environment interaction and genotype × year interaction on performance of crop varieties (Salmon, 1951; Horner and Frey, 1957; Sandison and Barlett, 1958). Their approaches were however shown to be of low precision when many environments and genotypes are involved. Other studies aimed at understanding phenotypic stability in crops were also carried out independently in the same period (Lewis, 1954; Dobzhansky and Levene, 1955).

The most popular methods of studying adaptability and stability of genotypes in multi-location trials were developed in the 1960s. Most of these methods are still applicable even today. Some of the methods developed in the 1960s but still being used by some current researchers are based on simple linear regression (Yates and Cochran, 1938; Finlay and Wilkinson, 1963; Eberhart and Russell, 1966).

The periods 1970s, 1980s, and 1990s saw the popularization of the use of multivariate analysis in the study of multi-environment trials (Mandel, 1971; Kempton, 1984; Gauch, 1988; Zobel *et al*., 1988; Crossa, 1990; Yan *et al*., 2000). Mixed models also came into use through the research efforts of Piepho in 1997 and Resende and Thompson in 2004 (Carvalho *et al*., 2014). Each of the methods developed by different researchers has its benefits and limitations.

2.16.3 Genotype-Genotype × environment biplot analysis

Since the beginning of formal agriculture, crop varieties have been known to give different levels of performance depending on either the physical environment or management approaches on them or both (Yan and Tinker, 2006). The initial approach taken to handle this phenomenon was to ignore it but with the emergence of research in agriculture two main groups of agriculturists have shown different interests in these differences in performance across environments which have been named genotype × environment interaction.

Genotype × environment data (GED) includes three main aspects which are mega-environment analysis, test environment evaluation and genotype evaluation (Yan and Kang, 2003). An ideal genotype in a particular mega-environment should have both high mean performance and stability across environments while an appropriate test environment should be both discriminating of genotypes and the mega-environments (Yan, 2001).

Biometricians have managed to quantify genotype × environment interaction mainly while plant breeders and agronomists have great interests in matching genotypes to environments in which they perform best (Yan and Tinker, 2006). The interest of all these groups have been met, mainly due to the advent of a technique in biometrics called biplot analysis.

Yan and Tinker (2006) described a biplot as a scatter plot that approximates and displays graphically two way tables by both row and columns in such a way that relationships among row, column factors and the associated interactions between the row and column factors can be visualized at the same time. Biplots were first reported by Gabriel (1971) and have since been used in visual data analysis in many scientific disciplines including medicine, economics, sociology, genetics and agronomy among others (Yan and Kang, 2003). Today most of the available statistical packages are designed to produce biplots of different information of presentation capabilities depending on the discipline to which the package is used. Just like principal component analysis (PCA), biplot analysis use singular value decomposition (SVD) developed by Pearson (1901) as a mathematical technique but biplot analysis is not equivalent to PCA.

Bradu and Gabriel (1978) were the first to apply biplots to agricultural data analysis when they used data from a cotton performance trial to show the diagnostic capability of biplots in model selection. The other researchers who worked on genotype × environment interaction include Kempton (1984), Gauch (1992), Cooper and DeLacy (1994), and Kroonenburg (1995). The turn of the 21st Century saw the introduction of the term "GGE biplot" and many visualization methods were introduced to address many questions to do with genotype × environment interaction (Yan *et al*., 2000). GGE brought to the fore the understanding that G and GE are the two sources of variation in genotype evaluation and that they must be studied together when genotypes and the environments they are to be grown are evaluated (Yan and Tinker, 2006). Since its proposal, genotype-genotype×environment (GGE) biplot analysis has evolved into a system that answers a lot of genotype × environment interaction questions graphically and clearly (Yan and Kang, 2003; Yan and Tinker, 2005a). Even diallel cross table (Yan and Hunt,

2002) and quantitative trait loci (QTL) × environment data can now be presented clearly using biplots (Yan and Tinker, 2005b). Today thousands of papers have been published in agriculture on the subject of genotype × environment interaction and presented using GGE biplots illustrating the relationships among variety testing environments and the performance of genotypes among testing environments among other concepts (Yan and Tinker, 2006).

2.16.4 Genotype × environment data analysis using AMMI 1 Graphs

AMMI is, however, is a family model rather than a single model hence model diagnosis is reguired to determine the model that fits a specific date set and research objective. The AMMI model family are hybrid statistical models that incorporate both the additive and the multiplicative components of the two-way data structure. The additive portion of the variance is suggested from the multiplicative variance (interaction) by ANOVA. Principal Component Analysis (PCA) analysis is then applied to the interaction (residual) portion from the ANOVA to extract a new set of coordinate axes which account more effectively for the interaction patterns. Estimation of the PCA axes is achieved according to the least squares principle (Bradu and Gabriel, 1978).

The AMMI 1graphs used in GED was first proposed by Gauch and Zobel (1997) to address the which-won-where pattern. In the graph, the abscissa represents the environment scores for the interaction principal component (IPC1) and the ordinate represents the nominal parameter under study. Each genotype is represented by a straight line defined by that genotype's average performance and IPC1 (that is the regression on the environment IPC1 score). Because the AMMI is a family of models, successive IPCs are indicated as IPC1, IPC2, and so on but the number of the components is always one less than the minimum number of genotypes and the environments. On the other hand, the the member of the AMMI model family with zero components are shown as AMMI 1, AMMI 2 going up to the full model retaining all the components which is denoted as AMMI F. According to Ebdon and Gauch (2002) mega-environment classification based on this method is virtually the same as that based on a GGE biplot. Yan *et al*. (2007), however, pointed out that this may be true in some instances but not in others and went on to point out that the GGE biplot explains more G + GE than the AMMI1 graph making it a more accurate presentation of the GGE of the data. Also, unlike the GGE biplot, the AMMI1 biplot also represents the environment main effects of the test environment, which is irrelevant to cultivar and test environment evaluation (Yan and Kang, 2003). The AMMI 1 biplot also displays the test environments by their main effects and IPC1 scores but it provides no interaction on the environment's ability in identifying superior cultivars.

2.17 An overview of mating designs used to study gene action in plants

Plant breeders use different mating patterns among the parents under investigation to understand the gene actions involved in the control of traits of interest. A mating design is a pattern of sexually combining desired parents in a crossing block to generate progeny for field evaluation (Nduwumuremyi *et al*., 2013). Mating designs help plant breeders to get information about genetic control of targeted traits, generate breeding populations on which selection for traits of interest can be done, estimate genetic gain from breeding efforts and in the evaluation of parents of interest in breeding programmes (Acquaah, 2012). Information generated from mating designs answer questions to do with the significance of parent genetic variation, its heritability and the gene action affecting the traits under investigation.

Selection of appropriate parents and the choice of a mating design determine the appropriateness of the information that is generated in the study. The mating design used in a plant gene action study is influenced by the technical and economic factors at play during the study. Some of the main technical factors to consider include whether the crop under study is self or cross pollinated, whether artificial or natural crossing will be used, the nature of pollen dissemination occurring in the species under study, the presence or absence of male sterility in the parents being crossed, size of the population needed in the subsequent generations of study as well as the objectives of the study (Acquaah, 2012).

Mating designs require that some genetic assumptions be met in the species of interest before application for the generated information to be scientifically valid (Griffings, 1956). The basic assumptions require that the species under study exhibit diploid behaviour at meiosis, genes under study must not be linked but exhibit independent assortment during meiosis, and there must be no non-allelic interactions and multiple alleles at the locus of interest (Nduwumuremyi *et al*., 2013). Reciprocal differences among the offspring of the parents and the genotype × environment interaction for the trait of interest must be absent.

2.17.1 Commonly used mating designs in plant breeding

The use of mating designs to study gene action and inheritance of traits in crops began the first half of the 20th century (Hill *et al*., 1998). However, it was in the early 1950s that the application of these designs in variety improvement was popularised (Griffing, 1956). Many descriptions of the possible mating designs for use in plant breeding have been put across but to date only eight are popular and commonly used by modern plant breeders and the choice of a mating design to use, however, is dependent on the objectives of the study, available space, the cost of

the layout and the possible biological limitations associated with the design (Nduwumuremyi *et al*., 2013). The commonly used designs are selfing, backcrossing, bi-parental progenies (BIP), polycross, top-cross, North Carolina designs (I, II, III), diallel mating designs (I, II, III, IV) and the line by tester designs (Hill *et al*., 1998). A brief description of these designs is given and the possible advantages associated with their use.

2.17.1.1 Self pollination

Selfing in planbreeding refers to the union of the male and the female gamates from the same plant. Gametes can be from the same flower or different flowere on the same plant. Selfing is the quickest way to achieving homozygosity in conventional plant breeding. In plants which are obligate selfpolinating, or where pollination can be effectively controlled, at least eight cycles of selfing are needed to achieve over 99% homozygosity in the selfed line. Selfing makes it possible to develop purelines which are used in constituting hybrids.

2.17.1.2 Backcrossing

Backcrossing works with corss compatible species and it is when a selected line is repeatedly mated to aspecific parent. This is usually done for the purpose of transferring a gene of interest from one line to the other. It also needs at least ten cycles of crossing to achieve the transfer of a gene of intrest to another line and retaining the original genetic constitution of the line of interest.

2.17.1.3 Bi-parental mating design

This is considered to be the simplest mating design in plant breeding. It is called the paired crossing design (Mather and Jinks, 1982). In this design, *n* number of parent plants are selected at random and crossed in pairs to produce ½*n* full-sib families (Acquaah, 2012). The progeny produced is evaluated in replicated trials and the variation observed partitioned by analysis of variance (ANOVA) into between and within families' variation. The main advantage of the bi-parental design is that it is very simple to execute (Acquaah, 2012). However, it has the limitation that it does not provide all the information needed by its model to estimate the parameters of the model. Also, because of lack of control of mating, relatedness among the progeny is limiting in this design. Thus unjustifiable assumptions are made on estimation of valuable genetic and environmental components (Nduwumuremyi *et al*., 2013).

2.17.1.4 The polycross

This is a scheme designed for inter mating a group of cultivars or lines by natural cross-pollination in an isolated block (Nduwumuremyi *et al*., 2013). Acquaah (2012) suggested that

this design is suited for species that are obligate cross pollinators. In the block, each parent is given an equal chance to naturally cross with other lines or cultivars present in the crossing block provided a proper layout that does not favour some lines versus others is done. The polycross design has the advantage that it can be used in the production of synthetic cultivars and in the recombination of specific lines during recurrent selection procedures as well as in the evaluation of general combining ability of the parent genotypes (Acquaah, 2012). However, random mating may be limited in the crossing block if naturally self-pollinating species are under study. The other limitations of this design are that it produces insufficient statistics to estimate all the parameters needed by the model, variance components for the maternal half-sibs are the only ones estimated, information about the males is lost, and the expected genetic gains are reduced by half and that the mating is non-random due to lack of synchronization of flowering and unequal pollen production in the crossing block.

2.17.1.5 The topcross mating design

The topcross design was suggested by Jenkins and Brunan in 1932 and refers to the mating between a line, selection or clone and a common pollen parent. It was later renamed topcross by Tysdal and Grandall in 1948 (Hill *et al*., 1998). In the design, the selected plants are crossed to a common tester(s) of known performance. The main reason of using the topcross method is to have a preliminary understanding of the performance or value of newly developed breeding lines (Charcosset *et al*., 1993). However, topcross progenies only yield GCA not SCA information. Its advantage is that it is simple and efficient in screening inbred lines for combining ability before pairing them in single cross yield trials (Nduwumuremyi *et al*., 2013). Rapid preliminary screening of genetic stocks involving the minimum cross effort and simple statistical analysis can be done using the topcross method (Mosa, 2010). The two major weaknesses of this mating design are that the single tester variety used may not offer broad genetic background for testing large inbred line stocks and that the numbers of crosses may become very large if the test inbred lines are numerous.

2.17.1.6 North Carolina Mating Designs

Three variants of the North Carolina mating designs exist namely I, II and III. North Carolina mating designs were designed by Comstock and Robinson in 1952 with the aim of generating more information about combining ability with relatively less effort than when the diallel is used (Nduwumuremyi *et al*., 2013). North Carolina mating designs are nested designs in which a group of parents designated females are pollinated by different individual males to produce progeny and these progeny are evaluated in replicated trials. The North Carolina design I is a

hierarchical design with non-common parents nested in common parents (Acquaah, 2012) and produces offspring which is both half and full sibs. The North Carolina mating design I has the advantage that it gives a test of significance for the additive genetic variance (Nduwumuremyi *et al*., 2013). The North Carolina II (NC II) mating design is a factorial mating matrix in which a different member of a group of parents designated male is used to mate a group of other parents designated females (Acquaah, 2012). This design is used to evaluate inbred lines for both general and specific combining ability (Acquaah, 2012). In the North Carolina III mating design, a random sample of F_2 plants is backcrossed to the inbred line parents from where the F_2 population was generated. This method allows for testing for epistasis and provides for estimates of additive and dominance components of variation.

2.17.1.7 The diallel mating design

The diallel mating design is the most popularly used mating design in plant breeding (Hallauer *et al*., 2010) to obtain different genetic information about populations. Four forms of a diallel mating design exist. A full diallel mating design allows for a set of parents to be intercrossed in all possible combinations and includes selfs and reciprocals (Makumbi *et al*., 2018). The second type is called Method II and involves one set of parents and one set of F_1s but excludes reciprocal F_1s. In the diallel method III, one set of F_1 and the reciprocals are included and in Method IV only one set of F_1s are included. Method IV is the most commonly used method in plant breeding. Diallel mating designs have been used widely to develop breeding populations for recurrent selection. It was also reported that the diallel method helps to provide the maximum chance to manage crossing in breeding populations and maximise the selection differential (Johnson and King, 1998).

Diallel cross analysis generates genetic information about plant populations. It is useful in identifying potential crosses that give the best lines early in the selection process (Makumbi *et al*., 2018). Diallel analysis is important to plant breeders because it gives information about general combining ability (GCA) and specific combining ability (SCA) (Nduwumuremyi *et al*., 2013). Diallel mating designs can use the fixed or random effects model. Where the random effects model is used, GCA and SCA variance are estimated whereas means of GCA effect for each parent and the SCA effect for each cross is measured using the fixed effects model (Hallauer *et al*., 2010).

To conduct GCA and SCA analysis, breeders often use the Griffing's (1956) method whereas for studying the action of genes, genetic components and estimation of heritability, breeders

use Hayman's (1954) method. In many studies about genetic parameters, the Griffing's and the Hayman methods have been used complementarily for interpretation of data because of the different sets of data each method is capable of giving.

Hayman's diallel analysis is premised on the assumptions that (i) parents used in the cross are homozygous, (ii) there is normal diploid segregation, (iii) multiple alleles are absent, (iv) maternal effects are absent, (v) there is no linkage of genes affecting the trait under study, (vi) mating is random, and that (vii) epistasis is absent.

The diallel analysis proposed by Jinks and Hayman (1953) and described further by Hayman's (1954a; 1954b) gives more extensive genetic information about a population. It provides information on additive and dominance effects of genes, average degree of dominance, direction of dominance, distribution of genes, maternal and reciprocal effects, number of groups of genes that control a trait and exhibit dominance, ratio of dominant to recessive alleles in all the parents and broad sense (H^2) and narrow sense (h^2) heritability (Makumbi *et al*., 2018). Hayman's (1954a, 1954b) analysis has been used extensively to study the mode of inheritance of many traits in several crops that include cereals, legumes, and cotton but less extensively in tobacco.

Hayman's (1954b) method of analysing diallel crosses entails estimation of genetic parameters and also graphical and statistical analysis of array variances and covariances (Makumbi *et al*., 2018). An "array" refers to all the crosses involving a specific parent in the experiment. Basically, the steps needed to carry out a Hayman diallel analysis are (i) an anova to detect genetic variation among the genotypes, (ii) creation of the array of means in the diallel table, (iii) calculation of the expected variances (Vri) and parent-offspring covariances (Wri) of individual arrays, (iv) calculation of the mean variance ($\tilde{V}r$) and covariance ($\hat{W}r$) over all arrays, (v) calculation of the variance of the array means ($\tilde{V}r$), (vi) testing the validity of the additive-dominance model, (vii) generation of the Wr-Vr graph, and (viii) estimation of the genetic components (Makumbi *et al*., 2018).

2.17.1.8 The line × tester design

This mating design is an extension of the topcross design. The difference only lies in that the line x tester design use more than one tester while the top cross uses only one tester. The line × tester design was proposed by Kemptone in 1957 (Sharma, 2006). It involves hybridizing of lines and broad testers in a one to one pattern to give f × m = fm hybrids (Sharma, 2006). This

method is regarded as the simplest approach that gives both full sibs and half sibs at the same time. It gives the SCA and GCA information of crosses and lines as well as the testers.

2.18 Parameters used to study gene action and their importance in plant breeding

Crop improvement is based on the presence of heritable variation within a plant population. Plant breeding works by selection of individual plants with traits of interest from within broad and variable plant populations (Allard, 1999). Variations within breeding plant populations can be naturally present or artificially created through mating matrices. When mating designs are used, it is also possible to fashion them in such a way that other genetic parameters about the traits under investigation are estimated. The most important genetic parameters estimated by plant breeders for a crop population are combining abilities (both general and specific combining ability), gene action (additive, dominance or epistatic), heritability and the influence of maternal effects on the trait of interest (Griffings, 1956a). In plant breeding, the goodness of a line is judged by the superiority of its progeny. Good inbred lines are selected on the basis of their combining ability to form hybrids. The combining ability of a line for a trait is dependent on the genes controlling the trait. Combining ability in plant breeding is partitioned into general combining ability (GCA) and specific combining ability (SCA) (Nduwumuremyi *et al*., 2013). GCA is an indicator of the average performance of a line in hybrid combinations. GCA effects are thought to be due to additive gene action, which in turn is associated with average effects of genes controlling a trait. GCA can also be described as the deviation of progeny mean from the mean of the progeny of all lines in the experiment (Acquaah, 2012).

On the other hand, SCA refers to the relative performances of line combinations or crosses when compared to the average performance of all the lines in the trial. It is due to non-additive gene action (Acquaah, 2012). Mating designs that involve reciprocal crosses are ideal for estimating maternal effect in plant breeding since maternal effects is a study of the influence of maternal background on the performance of resultant progeny in crosses. Heritability is another parameter that is of interest to plant breeders as it increases the extent to which a trait can be passed on from parents of offspring. It is therefore a measure of genetic gain in crop improvement. Depending on the objectives of the study, breeders can choose any one among the six popular designs to study genetic behaviour of genes controlling traits of interest. In this study, the full diallel was chosen because it is relatively inexpensive but can provide all invormation about about the parents under study expecially information about reciprocal and maternal effects which are considered imprortant for managing quality in flue-cured tobacco hybrid constitution.

CHAPTER THREE

Genetic diversity analysis of Zimbabwean flue-cured tobacco (*Nicotiana tabacum* L.) breeding lines and commercial cultivars using SSR markers

3.1 Introduction

Tobacco (*Nicotiana tabacum* L.) is a complex allotetraploid species ($2n = 4x = 48$) that is believed to have originated from natural hybridisation of two diploid species, *N. tomentosiformis* ($2n = 24$) as a paternal donor and *N. sylvestris* ($2n = 24$) as a maternal donor (Leitch *et al*., 2008). Tobacco is one of the most economically important non-food crops in Zimbabwe and other tobacco growing countries like Malawi, Brazil and China (Hu and Lee, 2015). The increasing demand of high yielding tobacco cultivars with favourable characteristics, combined with an emphasis on releasing new cultivars quickly, have led to the widespread use of conventional tobacco breeding approaches.

In the past few decades, efforts have been made to improve tobacco yield, quality, disease resistance, and other agronomic traits resulting in great improvement in many traits (Tong *et al*., 2016). Very few selected genotypes with narrow genetic base were used as parental lines in the development of high yielding tobacco cultivars (Moon *et al.*, 2009a; Fricano *et al*., 2012). This has resulted in extremely limited genetic diversity in the flue-cured tobacco gene pool as revealed by several molecular marker-based studies on diversity (Moon *et al*., 2009b; Fricano *et al*., 2012; Tong *et al.*, 2016). By continuously using related material in developing new tobacco varieties, plant breeders fear to dissipate gains achieved in other attributes of the crop such as cured leaf quality, flavor and nicotine content (Moon *et al*., 2009a). For most plant breeding programmes on flue-cured tobacco, it has become very difficult to select unique lines using morphological traits due to their close relationships. More so, this has been worsened by the effects of the environment in which the crop is grown, which has also failed to induce the expression of unique traits in elite-flue cured tobacco germplasm.

To ensure sustainable tobacco breeding programmes, it is essential to investigate and exploit genetic variation in germplasm resources for efficient development of new cultivars. New variety development or existing cultivar improvement programmes exploit genetic variation present in the population of a crop species under improvement (Saddoud *et al*., 2012). The extent of genetic variation within species differs depending on the evolutionary path taken by the species (Bourguet *et al*., 2002; Naik *et al*., 2009). Some species exhibit extensive genetic

variation that is clearly visible while others hardly show discernible differences in morphological characteristics. Often, the environment plays a major role in influencing the extent to which important traits of a crop population are expressed in the phenotype (Bernardo, 2008; Naik *et al.*, 2009; Johnson, 2007; Sabaghpour *et al.*, 2012). Most differences which exist within species' populations are not expressed in the phenotype mainly because of the influence of the environmental background in which the plants have been subjected to. This phenomenon has always posed a challenge to plant breeders and other plant scientists relying on phenotypic differences when studying variation among plant populations (Sajjad and Khan, 2009).

Recently, molecular methods have become an indispensable tool in analysis of genetic diversity of germplasm accessions and for molecular breeding programmes. The adoption and use of molecular tools in plant breeding solved the problem of environmental masking of traits during the selection process by plant breeders (Moose and Mumm, 2008; Ali and Ali, 2010; Rahman *et al*., 2011). Environmental background does not affect the genetic construct of an organism at molecular level rendering the use of molecular markers reliable and amenable to use across all environments with high repeatability (Saddoud *et al.,* 2012; Gholizadeh *et al.,* 2012; Guddadamath *et al.,* 2011). Many molecular markers have been developed but the most reliable ones that have been developed for tobacco genetic diversity and mapping studies include the RFLPs (Yang *et al*., 2007), AFLPs (Liu *et al*., 2009), RAPDs (Rossi, *et al*. 2001), ISSRs (Yang *et al.,* 2007), and microsatellites or simple sequence repeats (SSRs) markers (Bindler *et al*., 2011; Davalieva *et al*., 2010; Tong *et al*., 2016). SSRs have proven to be the markers of choice for genetic mapping (Lan *et al*., 2012), germplasm identity (Ye *et al*., 2009; Jia *et al*., 2014), gene localization, molecular marker-assisted selection breeding and genetic diversity analysis in tobacco (Bindler *et al*., 2011) and other crops of economic importance (Chena *et al*., 2017; Bonierbale *et al*., 1988). SSRs have the advantage of high polymorphism, co-dominance, specificity, extensive distribution, low cost and follows the Mendelian inheritance patterns, making them ideal for genetic diversity studies (Nunome *et al*., 2009). SSRs have been effective and more informative in producing considerable higher polymorphic information content (PIC), and to date it is the most commonly used molecular marker in the tobacco community.

In this study, SSR markers were used to study genetic diversity and population structure among a collection of flue-cured tobacco breeding lines and cultivars. Little is currently known about the amount of genetic variation or about the genetic relationships among germplasm

collections at Tobacco Research Board in Zimbabwe. Knowledge of the genetic diversity and population structure could assist in the selection of genetically diverse materials for breeding or development of novel tobacco cultivars with increased yield performance. This study also was to determine whether the available SSR markers can be utilized to distinguish between the most popular hybrid cultivars currently grown in Zimbabwe.

3.2 Materials and Methods

3.2.1 Plant materials

A total of 93 tobacco accessions were used in the study. Their identification codes and countries of origin are shown on Table 3.1. The samples were supplied by the Plant Breeding Division of Tobacco Research Board (TRB) in Zimbabwe. The accessions represent 10 selected hybrid cultivars developed at Tobacco Research Board for a period spanning 55 years, 19 breeding lines developed by different breeding programmes at TRB, 25 breeding lines from USA, 19 accessions from other tobacco producing countries excluding USA (Six from South Africa, two each from Democratic Republic of Congo (DRC), Malawi, Canada and Australia, and one each from Germany, Japan, Belgium, Poland and China), and 20 breeding lines whose origin cannot be traced (Table 3.1). For the purpose of analysis, accessions were divided into five populations, that is, samples from other countries excluding USA (Other countries), accessions of unknown origin (unknown), USA, breeding lines at TRB (Zimbabwe lines) and cultivars produced in Zimbabwe (Zimbabwe cultivars). All the accessions are conserved at TRB, Zimbabwe and are assigned, sequentially, an accesion number which starts with a “V” as they are entered into the genebank catalogue for identification purposes. The accession number or “V” number is given as the entry code in this study. All the tobacco lines and cultivars were planted in the greenhouse for three weeks prior to leaf sample collection for genomic DNA extraction.

Table 3.1: List of tobacco accessions used in this study including their entry codes and countries of origin

Entry No.	Entry Code	Entry Name	Origin	Entry No	Entry Code	Entry Name	Origin
1	V255	Warne	Malawi	48	V792	Flue-cured Kentucky 14	USA
2	V256	Cash	Malawi	49	V262	Oxford 1-181	USA
3	V154	Special 400	DRC	50	V263	Virginia 21	USA
4	V261	Oxford 1	DRC	51	V174	Golden harvest	USA
5	V499	Delcrest 66	S. Africa	52	V229	Virginia Bright Leaf	USA
6	V657	TL (13-20)	S. Africa	53	V231	Golden wilt	USA
7	V215	Silver dollar	S. Africa	54	V200	Yellow special A	USA
8	V604	TL 33	S. Africa	55	V219	Ruffle 1	USA
9	V413	WK 39	S. Africa	56	V221	Ruffle 3	USA
10	V620	Coker 347	S. Africa	57	V222	Ruffle 4	USA
11	V39	Proctor's Special	Canada	58	V224	Golden cure	USA
12	V675	Jing chang	China	59	V225	White Gold	USA
13	V148	Hicks broad leaf	Canada	60	V226	Carolina bright	USA
14	V838	Wisana	Poland	61	V228	401	USA
15	V396	Bel 61 -10	Germany	62	V510	HICKS	USA
16	V860	Enshu	Japan	63	V556	TI 1138	USA
17	V853	XA 219-5 d.h.	Australia	64	V560	Speight G28	USA
18	V862	Sirone	Australia	65	V183	White stem Orinoco	TRB Zimbabwe
19	V237	Hickory Prior	Belgium	66	V185	Yellow prior	TRB Zimbabwe
20	V41	Duquesne Johnson	*Unknown	67	V252	Mammoth Jamaica Wrapper	TRB Zimbabwe
21	V53	Harrison's Special	Unknown	68	V614	RKL (115-1-1)	TRB Zimbabwe
22	V624	NC 79	Unknown	69	V629	TB 22	TRB Zimbabwe
23	V671	NEW DEL	Unknown	70	V634	WM (84-20-2)	TRB Zimbabwe

24	V672	NORDEL	Unknown	71	V643	PY (1-19-1)	TRB Zimbabwe
25	V673	DELGOLD	Unknown	72	V649	ETW 22	TRB Zimbabwe
26	V731	Victoria selection	Unknown	73	V653	RL (1-2-1)	TRB Zimbabwe
27	V744	WZ (1-7-1-1)	Unknown	74	V655	ST (1-2-8)	TRB Zimbabwe
28	V149	Harrison's special	Unknown	75	V724	SR 1	TRB Zimbabwe
29	V112	C7	Unknown	76	V315	Aurea Delcrest	
30	V141	Vamor 50	Unknown	77	V688	AW(22-10-1-1-65)	TRB Zimbabwe
31	V169	Oxford 26	Unknown	78	V699	STNCA (1-10-10-7-28-4)	TRB Zimbabwe
32	V143	Dixie Bright leaf 102	Unknown	79	V86	Broad leaf	TRB Zimbabwe
33	V66	Coker's Certified 401	Unknown	80	V93	Daines bright leaf	TRB Zimbabwe
34	V67	Coker's Certified 402	Unknown	81	V12	Jamaica Wrapper	TRB Zimbabwe
35	V31	Gold Dollar	Unknown	82	V15	Willow Leaf	TRB Zimbabwe
36	V73	Oxford 26	Unknown	83	V25	Yellow Mammoth	TRB Zimbabwe
37	V229	Virginia	Unknown	84	V617	KUTSAGA 110	TRB Zimbabwe
38	V1	Station Bonanza	Unknown	85	V595	KUTSAGA 51E	TRB Zimbabwe
39	V75	White Mammoth Norfolk	Unknown	86	V596	KUTSAGA MAMMOTH E	TRB Zimbabwe
40	V64	American 401	USA	87	Uncoded	K RK 29	TRB Zimbabwe
41	V81	Virginia Scheviez	USA	88	Uncoded	K RK 26	TRB Zimbabwe
42	V84	American Joiner	USA	89	Uncoded	K RK 66	TRB Zimbabwe
43	V677	K326	USA	90	Uncoded	K RK72	TRB Zimbabwe
44	V303	Hicks	USA	91	Uncoded	K RK71	TRB Zimbabwe
45	V304	PD 83	USA	92	Uncoded	T75	TRB Zimbabwe
46	V234	402	USA	93	Uncoded	T 76	TRB Zimbabwe
47	V260	Coker 187	USA				

*Origin not known.

3.2.2 Genomic DNA extraction

Fresh young leaves of tobacco seedlings (100 mg) were collected, and ground in liquid nitrogen. Genomic DNA was extracted from the leaves using a Direct Zol™ Plant DNA extraction kit (Zymo Research, USA) following the manufacturer's instructions (https://www.baseclear.com/lab-products/zymo-research/sample-kits/). Electrophoresis on 0.8% agarose gel stained with ethidium bromide allowed estimation of the integrity of extracted DNA samples. The purity and quantity of extracted DNA were measured with a NanoDrop 2000c UV-Vis spectrophotometer (Thermo Scientific, Pittsburgh, PA, USA) (https://www.thermalscientific.com/). The working DNA concentration was adjusted to a concentration of 20 ng/µl prior to genotyping.

3.2.3 Microsatellite analysis

Microsatellite amplifications were performed using 10 SSR markers (Table 3.2) described by Bindler *et al*. (2011). The 10 SSR markers were selected based on higher number of alleles per marker and higher polymorphism on tobacco (Bindler *et al*., 2011). The SSR markers were amplified using a three-primer system for indirect labelling PCR as described by Shuelke (2000). The forward primer was labelled with a M-13 tail at the 5' end labelled with a fluorescent dye (FAM, NED, PET or VIC). The PCR reaction consisted of a 25 µl final volume containing 12.5 µl of PCR master mix (Biolab), 0.25 µM tailed M13 primer, 4 µM reverse primer, 2 µM M13-flourescent labelled primer (Invitrogen, Carlsbad, CA, USA), 20 ng of DNA, and dH_2O up to the final volume. Amplifications were performed using a 9600 thermal cycler (Applied Biosystems, Carlsbad, CA, USA). The following thermal conditions were used: 5 min at 94^0C and 35 cycles at 94 ^{0}C for 30 s, 58 ^{0}C for 30 s and 72 ^{0}C for 45 s and a final extension at 72 ^{0}C for 10 min. After amplification, primers containing different fluorescent dyes were multiplexed into different panels with panel 1 consisting of (PT60543, PT54629, PT54303, and PT61596) and panel 2 (PT51076, PT53409, PT60151, PT52480, PT61101, and PT61368). The multiplexed amplicons (1µl) were mixed with 9 µl Hi-Di formamide and 0.1 µl GeneScan 500 LIZ size standards (Applied Biosystems). Products were denatured at 94°C for 5 min. Capillary electrophoresis was done using an ABI PRISM 3500XL Genetic Analyzer at Inqaba Biotechnology Industries in South Africa. The size of each peak was determined and scored using GeneMapper Software v4.0 (Applied Biosystems). Genotyping tables were exported as tab-delimited files and formatted in Microsoft Excel for further statistical analysis.

Table 3.2: List of SSR loci selected for use in this study

Primer name	Primer Sequence (5'-3')	Repeat	Fluorescent dye	#Loci	T^0C
PT52480	F: M13-AGGAGGAGGGATTCCATTTG	TA	NED	3	56
	R:ATTCTTCCACATGACCCACC				
PT53409	F:M13-TCTCAAGAAGTAACGAAAGGACC	TA	VIC	3	55
	R:CAATCACATCTTGCAGCCAC				
PT54303	F:M13-TCTTTGCTAGCACGGAGGAT	TA	PET	4	55
	R:M13-CAAATGGGTTTGAATTTGCC				
PT54629	F:M13-GTCCCATTACCCACTCCCTT	TA	VIC	4	53
	R:TCCGTTGCAAATTTCTCCTT				
PT51076	F:M13-AACGTCTTGTAGGAGTGTCCAAA	TA	FAM	1	56
	R:CAAGACAAAGCGTTGGATGA				
PT61101	F:M13-AATCTCAAATATCTCGCTTCAGA	TA	PET	4	53
	R:CATGATAAGCAATTCTCAGGTGAT				
PT60151	F:M13AGCCGATAAGTAATAGGAAATACTACA	TA	PET	1	57
	R:TTTAACCCGCTCCAGAAATTA				
PT61368	F:M13TCATGTTGCCACAAATATCCA	TA	PET	3	56
	R:TTGCAGGTGAATTCATGGAG				
PT61551	F:M13TGTTCACAGCAGGATGAGGA	TA	NED	3	55
	R:GCATTCAACCTCCAAACAAGA				
PT61596	F:M13GCGCATGGAACTTCAATCA	TA	FAM	3	55
	R-GCTTGCCCTTGATAAAGCTTACTA				

T^oC- Annealing temperature

3.3 Data analysis

The program iMEC, an Online Marker Efficiency Calculator (Amiryousefii *et al*., 2018) was used to estimate the total number of observed alleles (Na) and polymorphic information content (PIC) = $1 - \Sigma\, p\mathrm{i}^2 - \Sigma\,\Sigma\, p\mathrm{i}^2\, p\mathrm{j}^2$ where *p*i and *p*j are the population frequency of the *i*th and *j*th allele. The first summation is over the total number of alleles, whereas the two subsequent summations denote all the *i* and *j* where $i \neq j$. P*ij* is the frequency of the *i*th allele at the *j*th marker. To evaluate the diversity level of each SSR marker, GenAlEx v. 6.5 (Peakall and Smouse, 2012) was used to estimate the number of effective alleles (Ne), Shannon's information index of genetic diversity (I), observed (Ho) heterozygosity, expected (He) heterozygosity that is, Nei's unbiased gene diversity calculated according to Nei (1978), as $n/(n-1) \times (1 - \Sigma ip\mathrm{i}\ 2)$ where *p* is the observed frequency of the *i*th allele and *n* is the sample size. Analysis of molecular variance (AMOVA) was used to evaluate the partitioning of diversity within populations, between populations, and between groups of samples using Arlequin (Excoffier and Lischer, 2010). Groups were defined as tobacco lines whose origin

cannot be ascertained (unknown), from different countries (other countries), USA, Zimbabwe breeding lines (Zim L) and Zimbabwe cultivars (Zim C). F-statistics was calculated according to the variance of heterozygosity in the population at different levels of population structure (i.e., individual, subpopulation and population levels).

The genetic relationship between the 93 tobacco lines and varieties (Table 3.1) was investigated using principal component analysis (PCA) implemented in GenAlEX v. 6.5 (Peakall and Smouse, 2012). DARwin 6.0.1.2 (Perrier and Jacquemoud-Collet, 2006) was used to construct the dendrogram using the weighted Neighbor-Joining (NJ) method. In addition, the genetic structure of the 93 tobacco lines and varieties was modelled using a Bayesian clustering algorithm implemented in STRUCTURE v. 2.3.4 (Pritchard *et al*., 2000). Since no prior knowledge about the origin of the populations under study was available, the "admixture model" was used and then a correlated allele frequencies model was selected. Ten replicate simulations were conducted for each value of K, with the number of founding groups ranging from 2 to 10, using a burn-in of 50000 and a final run of 100000 Markov Chain Monte Carlo (MCMC) steps. The optimal value of K was calculated using the maximum value of K (Evanno *et al*., 2005) implemented in STRUCTURE HARVESTER 0.6.93 (Earl and von Holdt, 2012).

3.4 Results

3.4.1 Descriptive statistics

Table 3.3: Data on the number of alleles, allele size range, highest frequency allele and polymorphic information content (PIC) found among tobacco varieties using 10 SSR markers.

Primer Name	Na	Size range (bp)	Highest frequency allele		Ne	He	I	Ho	F	PIC value
			Allele	(%)						
PT52480	4	74-151	131	89	1.80	0.41	0.68	0.41	0.123	0.31
PT53409	5	108-195	108,216	56	1.91	0.48	0.55	0.73	-0.45	0.55
PT54303	5	213-293	250,293	85	2.44	0.59	1.00	0.97	-0.60	0.34
PT54629	4	213-225	213	90	1.98	0.40	0.81	0.59	-0.25	0.28
PT51076	4	79-125	125	70	1.42	0.29	0.46	0.36	-2.22	0.48
PT61101	4	203-219	205	36	2.50	0.59	0.96	0.83	-0.45	0.62
PT60151	2	208-216	153	73	1.67	0.37	0.60	0.04	0.88	0.54
PT61368	2	200-208	200,208	100	2.01	0.41	0.75	1.00	-0.75	0.00
PT61551	4	208-216	208,212,213,216	100	1.78	0.53	0.85	0.95	-0.75	0.00
PT61596	16	84-226	202	30	1.56	0.59	1.03	0.97	0.72	0.76
Total	50			729.00						3.09
Average	5			72.90	1.90	0.41		0.59		0.31

Na - Number of alleles

Ne - Effective number of alleles

He - Expected heterozygosity

I - Shannon's information index;

Ho - Observed heterozygosity;

F - Wright's fixation index.

PIC - Polymorphic information content

Descriptive statistics were as shown for all markers and accessions (Table 3.3). The total number of alleles amplified at the 10 SSR loci was 50. The mean number of alleles detected for each locus was 5, ranging from 2 alleles for markers PT60151 and PT61368 to 16 alleles for marker PT61596. The level of polymorphism of the studied microsatellite loci was evaluated by calculating PIC values for each of the 10 SSR loci. PIC values varied among loci and ranged from 0.28 (PT54629) to 0.76 (PT61696), with an average value of 0.31 per locus local. The effective number of alleles per locus (Ne) varied from 1.42 (for marker PT51076) to 2.50 (PT61101), with an average of 1.9. The expected heterozygosity (He) ranged from 0.29 (PT51076) to 0.59 (PT61596, PT61101, PT61596), with an average of 0.41. The observed

heterozygosity (Ho) varied from 0.04 (PT60151) to 1 (PT61368), with an average of 0.59 (Table 3.3).

Table 3.4: Summary statistics of genetic diversity for model-based groups.

Population	N	Na	Private alleles	Ho	He
Other countries	19	33	6	0.628	0.480
Unknown	20	27	4	0.636	0.479
USA	25	27	6	0.597	0.424
Zim lines	19	30	7	0.711	0.50
Zim cultivars	10	25	5	0.679	0.047

N - Number of samples

Na - Number of alleles

He - Expected heterozygosity

Ho - Observed heterozygosity

3.4.2 Genetic diversity and relationships

Table 3.5: Analysis of molecular variance (AMOVA) of tobacco accessions and cultivars based on 10 SSR markers.

Source	df	SS	MS	Est. Var.	% Var	F_{IS}	F-statistics
Among Pops	4	32.048	8.012	0.161	3%	-	F_{ST} (P<0.001
Among Individuals	88	188.506	2.142	0.000	0%	-	F_{IS} (P<0.001)
Within Individuals	93	419.000	4.505	4.505	97%	-	F_{IT} (p<0.001)
Total	185	639.554		4.666	100%	0.046	

DF = Degrees of freedom, SS = sum of squares, MS = mean squares, Est. var. = estimated variance, % Var. = percentage variation, F_{ST} = genetic differentiation among populations, F_{IS} = inbreeding coefficient (fixation index), F_{IT} = overall inbreeding coefficient of an individual relative to the total population (Individual within the Total population)

The AMOVA analysis revealed that 3% of the genetic variation (P <0.001) was distributed among the five populations, whereas 97% (P <0.001) was within the individuals (Table 3.5). These results indicated that the genetic variation within the populations contributed to most of the genetic diversity detected. The fixation index (F_{IS}) was 0.046. The variation among individuals contributed 0% to total genetic variability.

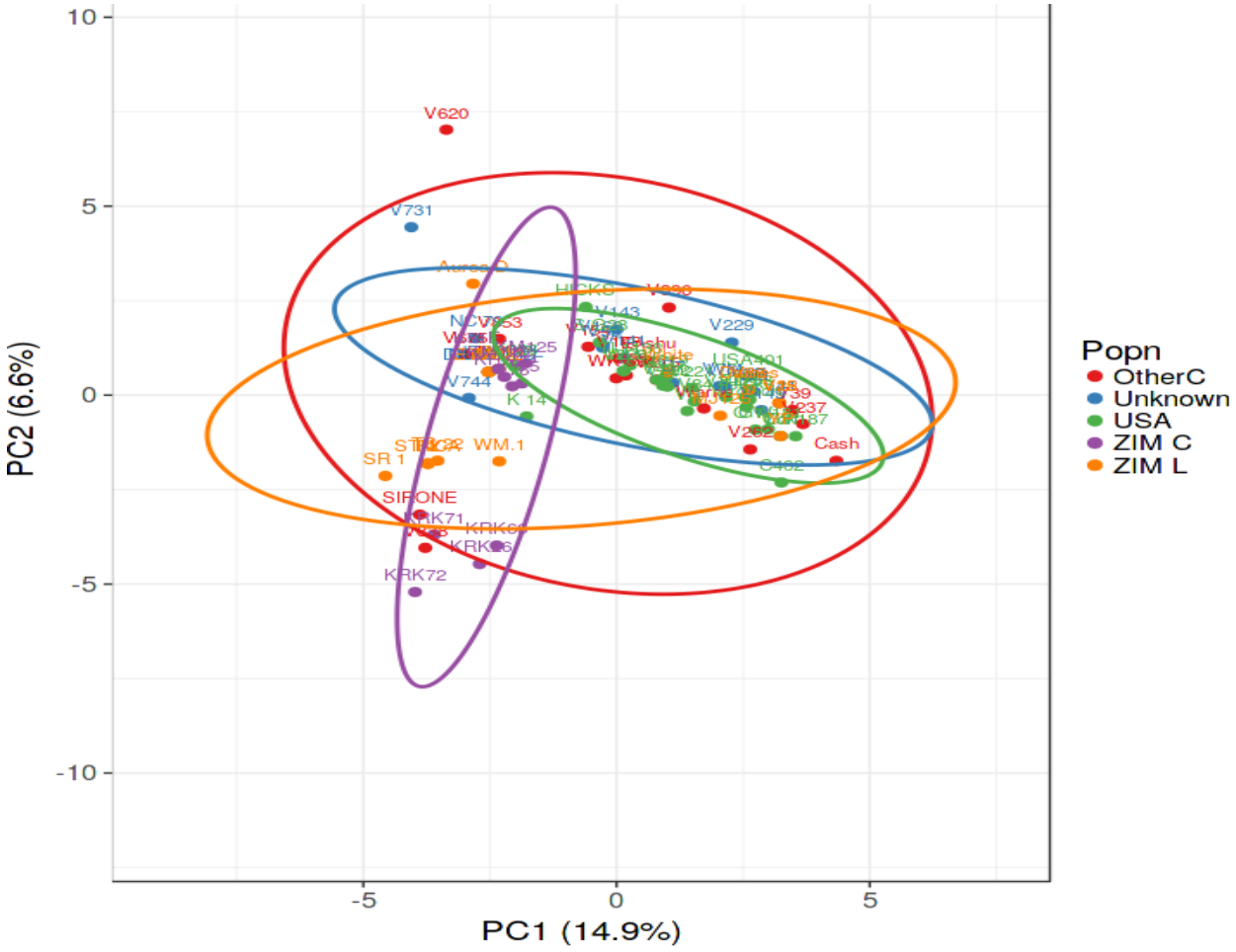


Figure 3.1: Principal coordinate analysis (PCA) plot showing the clustering of 93 tobacco accessions from Zimbabwe's germplasm.

The principal coordinate analysis (PCoA) plot of the 93 tobacco lines from five populations showed no distinct clustering pattern (Figure 3.1). The first two principal coordinates explained 21.5% of the total genetic variation within the accessions. Specifically, the first and second coordinates explained 14.9% and 6.6% of the total genetic variation, respectively. The ZIM C c(in purple) were the elite lines which were grown as cultivars by farmers in Zimbabwe (Figure 3.1). They intersect with the other lines from the USA and some old Zimbabwean lines (ZIM L) and those of undefined origin showing the presence of many common genes in the genotypes.

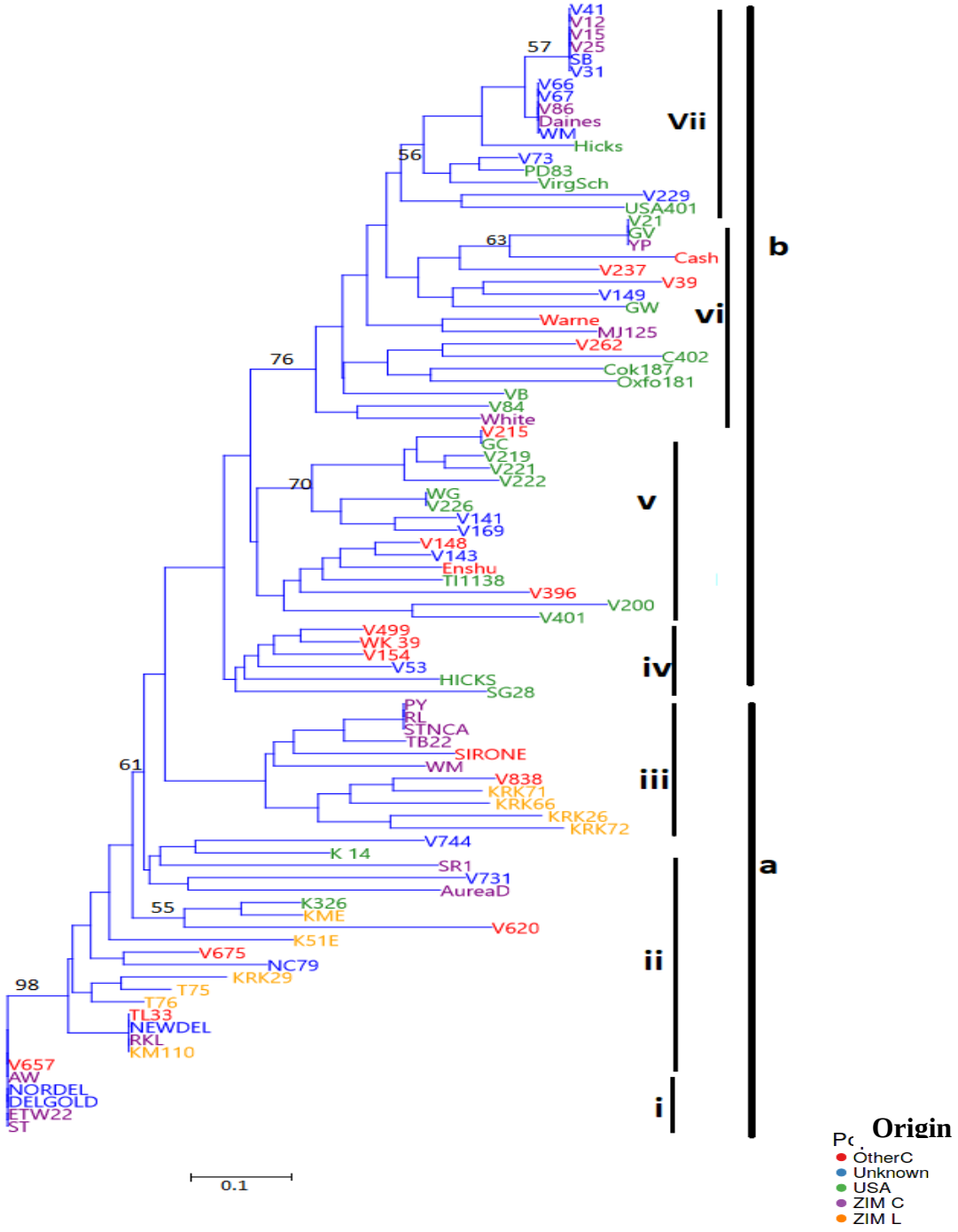


Figure 3.2: Unrooted Neighbor joining tree of 93 tobacco accessions estimates computed among pairwise comparisons of tobacco accession based on 10 microsatellite markers. The origins are color-coded as shown in the tree.

A dendrogram generated using the unweighted neighbour-joining tree resulted in the grouping of the 93 tobacco lines into two distinct clusters (Figure 3.2). Cluster (a) included 35 accessions from all five populations and was represented mainly by Zimbabwean accessions. All the tobacco cultivars bred in Zimbabwe were in this cluster including varieties Kutsaga Mammoth

E (KM E) and K51E that were developed in Zimbabwe in the 1960s. The cluster also include two lines from USA, the flue-cured Kentucky 14 (Ky14) and K326 a popular breeding line known for its curing quality and smoking qualities. Cluster (a) also has six lines from unknown sources and five from other countries. Cluster (a) may be further divided into three subclusters (I, II, III). Subcluster III consisted of hybrid varieties (K RK66, K RK72, K RK26 and K RK 71) together with breeding lines STNCA, WM, PY, RL as well as SIRONE from Australia and Wisana (Poland). Cluster (b) consisted of 58 accessions, 23 accessions from USA, 8 Zim lines, 14 other countries and 13 unknown. The cluster was represented mainly by USA accessions. An evidence of the use of common materials for obtaining cultivars can be visualized by the analysis of cluster (a) in the dendrogram (Figure 3.2). The hybrids K RK26 (K326 x RW), K RK71 (ONC x BAZ) and K RK66 (ONC x MG) as well as K RK72 (ONC x xm26) were grouped together (cluster iii) indicating narrow genetic base. XM26 parent to K RK72 is present in the genealogy of MG the parental line to K RK66. However, K RK71 was in different sub-cluster with K RK75 although they share the parent 'BAZ' (Figure 3.2). Similarly, K326 (cluster II) and KRK 26 (group III) were in different sub clusters despite the fact that K326 is a parent to K RK26. Bootstrap values were used to indicate the strength of relationship between branches in the dendrogram. For instance on the denrogram (Figure 3.2) the branch with the lines V657, AW, Nordel, Delgold, ETW and ST had a 98% relationship with the next cluster on the dendrogram and the relationship was significant at $p = 0.05$ level of significance. Only bootstrap values of 50 – 100 are indicated on the dendrogram because they were significant at $p = 0.05$ and those whose values are below 50 are not shown.

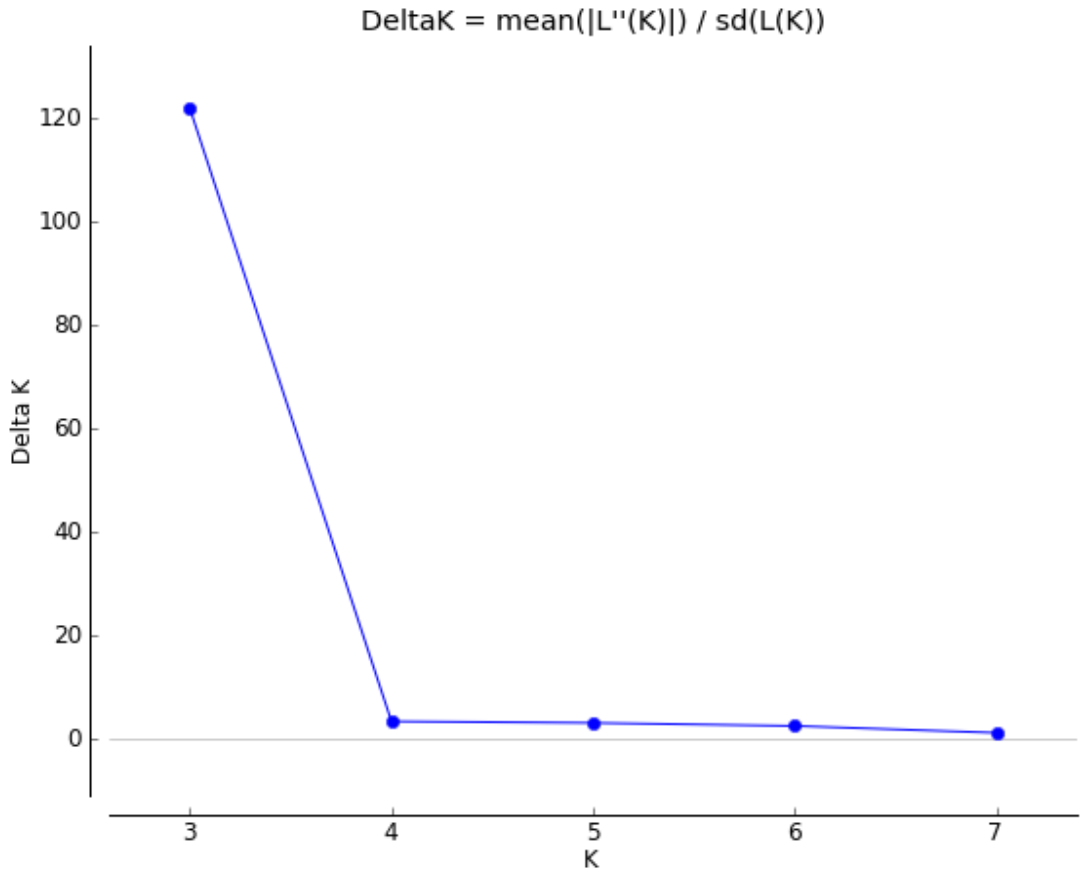


Figure 3.3: Magnitude of delta K calculated for each level of K, used to determine the most probable number of genetic clusters K=3 of tobacco accessions.

The bayesian clustering results, implemented in STRUCTURE software, were in agreement with the results of the PCoA and UPGMA analyses. In this analysis the best inferred number of clusters was *K* = 6, obtained by the Evanno's method (Figure 3.3). The majority of accessions of each group were grouped preferentially in one or the other cluster. At K = 3, three populations are apparent with minor admixture within groups. Overall, admixture and putative hybrids were observed in many accessions in this data set. These same general trends were observed at each K value assessed.

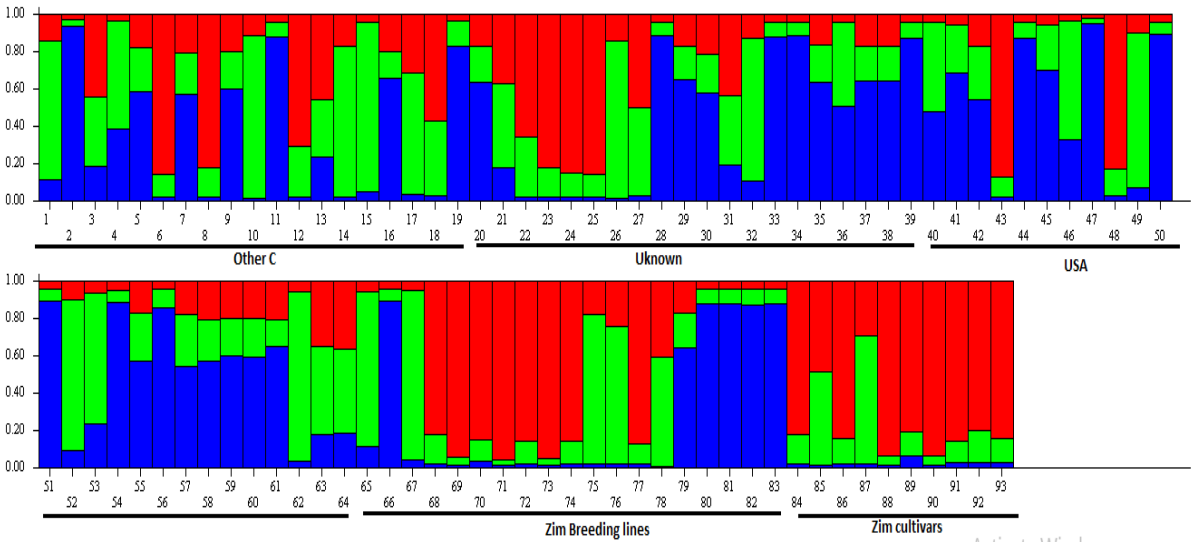


Figure 3.4: Population genetic structure of tobacco accessions in Zimbabwe as estimated by STRUCTURE using the SSR marker data set.

Each sample is represented by a vertical histogram partitioned into K = 3 colored segments that represents the estimated membership. The proportion of ancestry (%) is reported on the ordinate axis and the identification number of each accession follows that on Table 3.1 where the germplasm has been grouped into: other cultivars, unknown lines, USA derived lines, Zim breeding lines and Zim cultivars. STRUCTURE v2.3 (Pritchard *et al*., 2000) was used to investigate the genetic structure of the tobacco germplasm collection. The number of genetically homogenous clusters (ΔK) were identified at K = 3 (Figure 3.3) following the procedure of (Evanno *et al*., 2005) and was identified according to the rate of change of Ln *P*(D). From Figure 3.3 it can be seen that Structure analysis revealed significant admixture memberships (shown by single accessions with different colors) in each of the 93 accessions used in this study and no fixed clusters representing any particular designated population based on origin of the lines was observed. The overall pattern showed that the accessions shared a large proportion of genes making the whole population less diverse.

3.5 Discussion

3.5.1 Descriptive statistics

The genetic diversity revealed by the 93 accessions evaluated in this study was low, as reflected in the average allele number per locus (5). Darvishzedah *et al*. (2013), studying diversity in 100 oriental tobacco accessions reported average allele number ranging from 2 to 6 and reported low diversity in the accessions studied. The expected heterozygosity (He) of 0.41 found in this study also concurs with the findings of Darvishzedah *et al*. (2013), who also

reported an expected heterozygosity of 0.43 in their experiment. Polymorphic information content (PIC) is another measure of polymorphism within alleles. In this study, an average PIC of 0.31 was recorded further supporting the observation that there is limited diversity in this germplasm. Davalieva *et al*. (2010) assessed the level of polymorphism of the loci of Macedonian tobacco germplasm they studied by calculating the PIC values for each of the SSR loci. They reported that the PIC values ranged from 0.0 to 0.79 per locus which are comparable to the values in this study which ranged from 0.0 to 0.76. The mean number of alleles and PIC values detected were in line with other average number of alleles and PIC values observed in genetic diversity studies of tobacco (Fricano *et al*., 2012; Moon *et al*., 2009b). Similarly, the average expected and observed heterozygosity was in the range reported by Fricano *et al*. (2012); Gholizadeh *et al*. (2012); Moon *et al*. (2009b) and Davalieva *et al*. (2010) whose ranges were all from 0.0 to 0.52.

Analysis of the 93 accessions studied resulted in the creation of four distinct groups namely: other countries (other lines) , those from USA, lines developed in Zimbabwe during the early years of tobacco breeding (Zim lines.) and the currently cultivated varieties in Zimbabwe (Zim cultivars) (Table 3.4). The highest number of private alleles (7) was observed in Zimbabwe breeding lines (Zim lines) followed by breeding lines from USA and "other countries", with six private alleles. Accessions whose origin could not be ascertained had the least number of private alleles (4). Irrespective of populations, the observed heterozygosity was higher than expected heterozygosity indicating gene flow among the populations. This is consistent with the use of shared breeding lines in various tobacco breeding programmes in different countries. The analysis of the distribution of alleles in different populations is important for explaining the genetic diversity and population relationships. Private alleles are very important in plant breeding as they introduce genetic variations within a population.

3.5.2 Genetic diversity and relationships

Clustering was not tightly based on the geographical location indicating gene flow across the breeding programmes. Analyses of the accessions grouping suggest that there is an exchange of germplasm between the breeding programmes. Accessions from different countries are distributed across the clusters an indication of reciprocal exchange of material among breeding programmes. Overally, the results from the population structure analysis were consistent with the results from the genetic cluster analysis. The presence of admixtures may be due to the use

of common ancestors carrying important traits or germplasm exchange between different breeding programmes.

The demand for higher yielding, diseases resistant and better adapted tobacco varieties has increased the necessity to exploit the genetic variation of genebank collections. Evaluation of the genetic diversity of tobacco accessions can be an important source of information in order to discover and exploit novel alleles to be used in breeding programmes.

The observation herein, that tobacco has a narrow genetic base is in agreement with other studies (Del Piano *et al.,* 2000; Julio *et al.,* 2006; Rossi *et al.,* 2000). Previous studies have also shown that cultivated tobacco is a tetraploid species with a very large genome (Doganlar *et al.,* 2002; Ren and Timko, 2001) and that the polymorphism levels between tobacco cultivars (especially between those of the same type) are generally low (Del Piano *et al.,* 2000; Julio *et al.,* 2006; Rossi *et al.,* 2000). In this study, extremely limited genetic diversity was observed in flue-cured. Tobacco was domesticated from only a few wild progenitors, and the genetic diversity in tobacco has decreased in recent decades under strict breeding selection (Moon *et al.,* 2009a), resulting in extremely limited genetic diversity in the flue-cured tobacco gene pool as revealed by several molecular marker-based studies on diversity. This genetic diversity reduction may have occurred because of the selection of specific alleles that confer stability in Zimbabwe environment. However, the narrow genetic diversity identified among tobacco accessions collections at TRB necessitates the need of broadening genetic diversity by introducing more genetically diverse germplasm and the use of landraces to prevent inbreeding depression.

3.6 Conclusions and recommendations

3.6.1 Conclusions

1. The study showed that there is limited genetic diversity among the tobacco germplasm held at the Tobacco Research Board of Zimbabwe.
2. However, there are some genotypes like Coker 347, which are different from the rest of the genotypes and could contribute to variation in tobacco population.
3. The study also showed that the available SSR markers in Zimbabwe can be effectively used to study diversity in flue-cured tobacco.

3.6.2 Recommendations

1. The use of more diverse lines could be of interest to exploit novel alleles in crop breeding programmes without losing quality.

2. Genotypes with private alleles like Coker 347 (V620) should be considered for future use in breeding programmes in order to expand the tobacco genetic base.

CHAPTER FOUR

Morphological diversity assessment of Zimbabwean flue-cured tobacco (*Nicotiana tabacum* L.) by phenotypic characterization

4.1 Introduction

Conventional plant variety breeding and cultivar improvement depends mainly on the use of observable phenotypic differences among individual plants of a crop species (Ahmed *et al*., 2014). This means that the traits used in most conventional plant breeding programmes are major traits that can easily be expressed and observed in the phenotype of a species. Traits that are not expressed easily in the phenotype are not usually used as selection parameters in conventional plant breeding since they are not picked up by the naked eye or simple apparatus used in the selection process (Moose and Mumm, 2008). The conventional plant breeder, therefore, bases selection on observed phenotypic traits since such traits are the ones easily expressed in the population under improvement.

Phenotypic differences that exist among plants of the same species are attributed to genetic variation within the concerned crops and is commonly referred in plant breeding as genetic variation or genetic diversity (Lv *et al*., 2012; Lahiji *et al*., 2013). The presence of diversity in breeding material is important since diversity provides for sources of genes for pest and disease resistance and possible gene recombination resulting in new potentially beneficial crop varieties (Govindaraj *et al*., 2015; Yong-Bi Fu, 2015).

Even if molecular methods have been used to distinguish genotypes, morphological or phenotypic traits also need to be further explored to be able to identify those differences among plants that are expressed in the phenotype.

Many studies on genetic diversity in elite flue-cured tobacco germplasm have reported limited exploitable variation for possible future improvement of the crop (Zhang *et al*., 2008; Moon *et al*., 2009a; Moon *et al*., 2009b). The loss of variation is attributed to continuous selection among related breeding material and the strict quality requirements dictated by the consumers of tobacco (Moon *et al*., 2009a). The presence of genetic variation is a prerequisite for any improvement of a population for specific traits in plant breeding (Ulukan, 2011) and such variation can be obtained from wild relatives, obsolete lines, old varieties, plant introductions, landraces and possibly genetic engineering (Al-Khayri *et al*., 2015). A lot of effort has been made to gather and store securely germplasm of many different crop species across the world for future use by plant breeders in crop variety improvement programmes (Glaszmann *et al*.,

2010). These germplasm stores are rich reservoirs of potentially useful genetic variability for a wide range of plants for crop breeding since such genetic resources allow crop breeders to create novel plant gene combinations that exhibit superior performance phenotypically for traits of interest.

Although the accessions have been collected and safely stored in gene banks, most of the materials' characteristics are not known to potential users, the plant breeders. This scarcity of information about the genetic resources available to breeders is partly the reason why most of the stored accessions are not adopted in modern plant breeding schemes to increase the diversity of elite breeding material (Glaszmann *et al*., 2010). Field screening of such stored material and characterizing it will expose to plant breeders those accessions which are potentially useful for crop improvement and allow them to be used in modern plant breeding schemes (Miflin, 2000).

Zimbabwe has a large collection of accessions collected by TRB that are kept at Kutsaga over the years since the 1800s when tobacco farming began in the country. The collection included introductions, landraces, obsolete varieties and tobacco progenitors. Although the accessions have been catalogued, no effort has been made to characterise them, hence, little is known about their agro-morphological traits that can be useful in plant breeding programmes. This study seeks to agro-morphologically characterise the Zimbabwean flue-cured tobacco accessions *in situ* in order to expose their diversity and potential usefulness in new variety development programmes.

4.2 Materials and Methods

A total of 196 tobacco germplasm selections from the Tobacco Research Board tobacco gene bank were used in the study. These entries included all the early introductions into the country from different parts of the world as well as the early varieties developed in the country. The selection included the core germplasm used in the development and improvement of the modern tobacco varieties in Zimbabwe whose qualities are recognized globally but excluded modern multi-lines derived through intercrossing foundation germplasm in the country. This collection started in 1896 and ended around 1990. The list of entries used in the study is given in Appendix 1.

4.2.1 Experimental design

The 196 tobacco germplasm entries (Appendix 1) from the Tobacco Research Board's gene bank were planted and evaluated using a 14 × 14 α-lattice design with two replicates over two

seasons. The entries were evaluated for the thirty-five morphological traits identified and recommended in the International Union for the Protection of New Varieties of Plants (UPOV) guidelines (UPOV, 2002).

4.2.2 Agronomic practices

4.2.2.1 Land preparation

The land was deep ploughed to a depth of beyond 30 cm soon after the end of previous seasons' rains in April 2016 and 2017 to take advantage of residual moisture in the soil. This was done to conserve moisture in the soil and promote organic matter decomposition and release of nitrogen. Deep ploughing was earlier done using a mouldboard plough, and later in August 2016 and 2017, the land was disced to destroy weeds and create a fine tilth. Ridges were built using a tractor mounted ridger at a spacing of 1.2 m apart. At ridging, fumigation was done using using Ethylene dibromide $(CH_2Br)_2$ at a rate of 4 ml per planting station to control root-knot nematodes (*M. javanica* L).

4.2.2.2 Planting

Seedlings used in this experiment were raised using the float tray seedling production system (TRB, 2002). Seeds were sown during the second week of July 2016 and 2017 and seedlings were managed according to the Float-tray Seedbed Management Guidelines (TRB, 2002). When the seedlings were ready for transplanting, planting was done on the ridges at an in-row spacing of 0.56 m on the 25th of October of the years 2016 and 2017. The planting process involved opening up holes at a spacing of 56 cm and applying 5 litres of water to the hole. The seedlings were then planted in the water filled hole in a process called "Water Planting" in tobacco farming. When the water had dried in the planting hole with a planted seedling, dry soil was used to cover the planting hole thus creating soil mulch. Basal fertilizer (N: P: K; 6: 28: 23) was applied at a rate of 32 g at a distance of 10 cm away from each planting station. Ammonium Nitrate (35%N) was later applied after 4 weeks of planting at a rate of 8 g per planting station at a distance of 10 cm away from the planted seedlings. The plants were left to grow up to 50% flowering and measurements were taken.

4.2.2.3 Weed and pest management during crop growth

To control broad leaf, grasses and sedges during the early stages of development of the crop, a tank mix of the herbicides Clomazone® (Active ingredient: 2-(2-Chlorophenyl)methyl-4, 4-dimethyl-3-isoxazolidinone) and Authority ® (Active ingredient: Sulphentrazone WP)was applied within three days of planting. Later in the season, the experiment was kept weed free

by manual weeding whenever weeds appeared in the experiment land. A systemic aphicide, Acetamiprid was applied in the planting water at a rate of 5g/100L of water and applied using a 30ml cup. Cutworm control was done by applying Imidacloprid diluted at a rate of 13ml/100L of water at the base of each seedling soon after planting at a rate of 30ml per planting station. To control budworms during the development stages of the crop, Imidacloprid (Active ingredient: Imidaclorid) was applied whenever seven budworms were spotted on every 1000 plants.

4.3 Data collection, analysis and presentation

Collected data comprised the thirty-five morphological traits identified and recommended in the UPOV guidelines (UPOV, 2002). Data reduction was done using TwinSpan to produce 10 graphs. Each of the groups was analysed using hierarchical cluster analysis to produce dendrograms using Euclidean's distance with the UPGMA algorithm (Harding and Payne, 2011) using Genstat Version 17 (VSN International (2017) to show the relationships among the germplasm.

4.4 Results

4.4.1 Similarity of the Zimbabwean flue-cured tobacco accessions by group

Figure 4.1 to Figure 4.10 show 196 main flue-cured tobacco accessions from which all the other elite lines were derived through different breeding methods in Zimbabwe. The 196 accessions could not be presented together on one dendrogram since the output was so cluttered that it was difficult to interpret the output. The accessions were, therefore, put into ten groups of similar accessions each which were identified using TwinSpan analysis for data reduction.

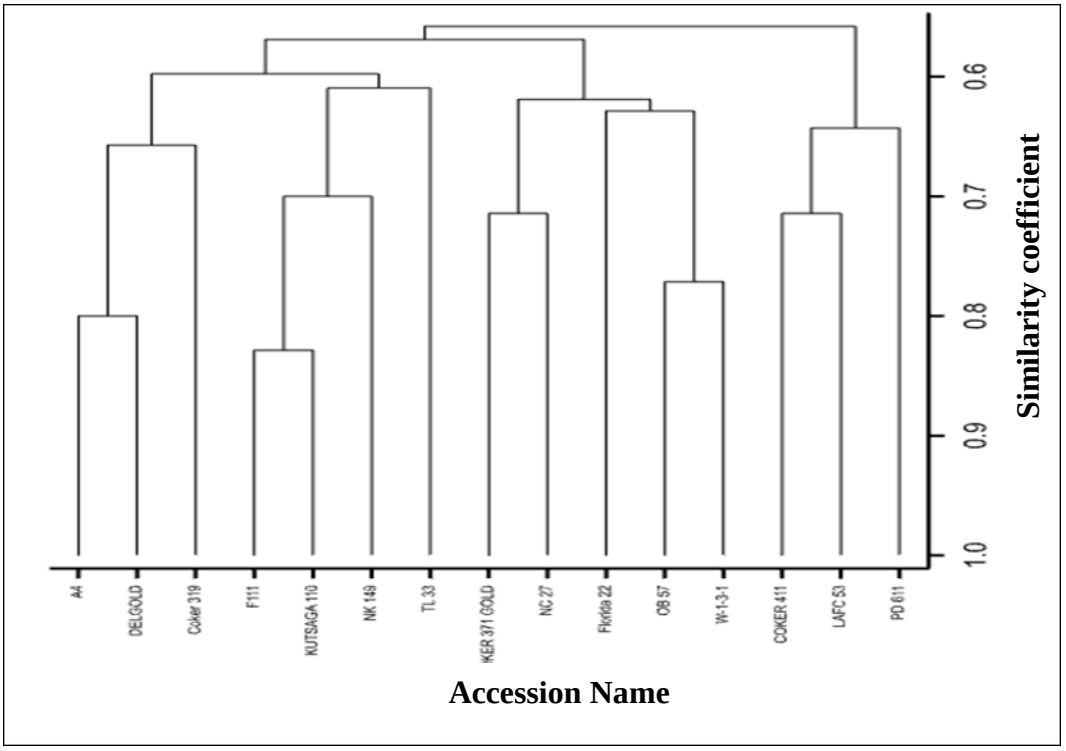


Figure 4.1: Average similarity cluster analysis for Group 1

Figure 4.1 is a cluster of 15 accessions which are closely related to one another in the thirty-five morphological traits studied. In this group, four distinct clusters emerged. Accessions A4, Delgold and Coker 39 had a similarity coefficient of 80%. The second major branch had four accessions namely F111, Kutsaga 110, NK 149 and TL30. Accessions Coker 371 Gold, NC 27, Florida 22, OB 57 and W-1-3-1 clustered with a similarity coefficient of 62%. The fourth branch had three varieties Coker 411, LAFC 53 and PD 611 and gave a similarity coefficient of 65%.

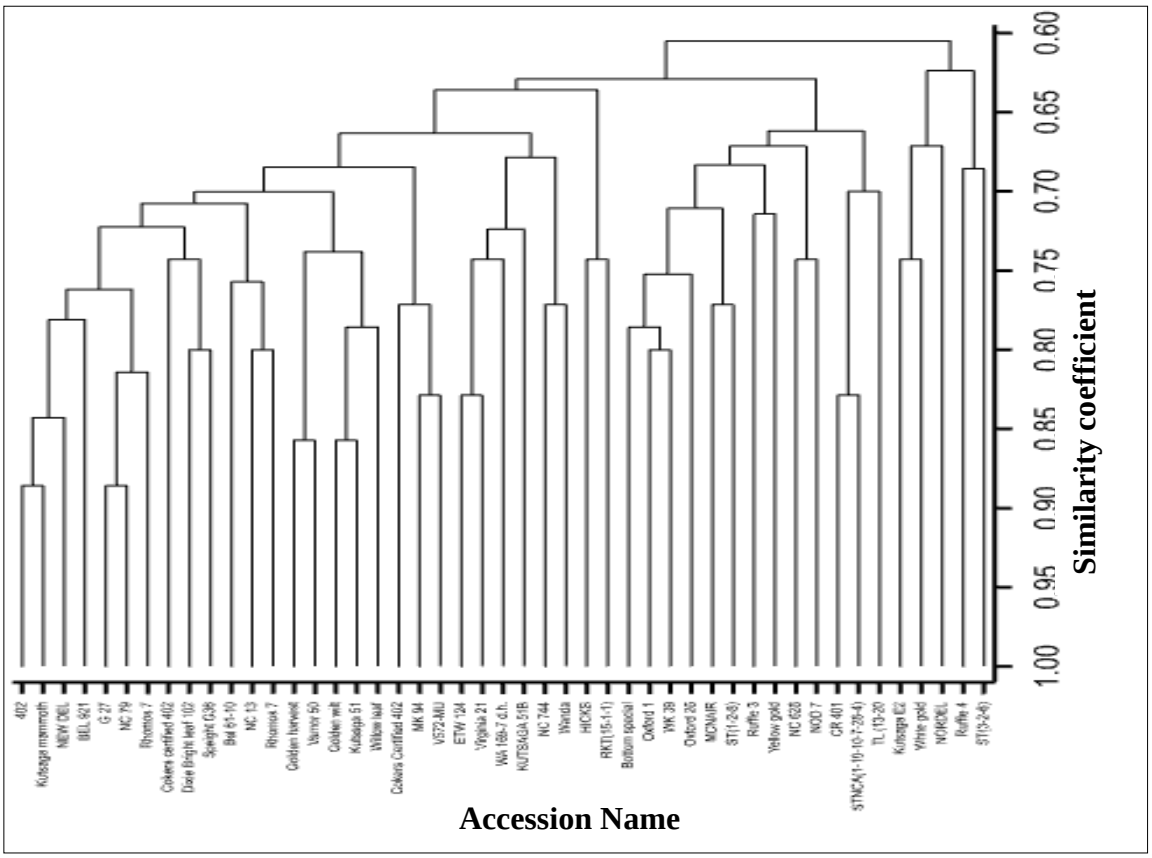


Figure 4.2: Average cluster analysis for Group 2

In Figure 4.2, 47 accessions clustered together in a dendrogram to give eight major clusters composed of different numbers of accessions. The first major cluster has seven accessions namely 402, Kutsaga Mammoth, New Del, Bel 921, G27, NC 79, and Bromos 2. Accessions Coker Delfield 402, Dixie Bright Leaf 107, and Speight G38 clustered together in another group with a similarity coefficient of 75%. Hicks and RKT(15-1-1) clusterd together in one of the smallest branches and the largest branch of the dendrogram has 13 accessions starting with Bottom special and ending with accession TL(1320) having similarity of 62%. The lines have relatively dark pigmentation on the leaves and stem and were bred for root-knot nematode resistance.

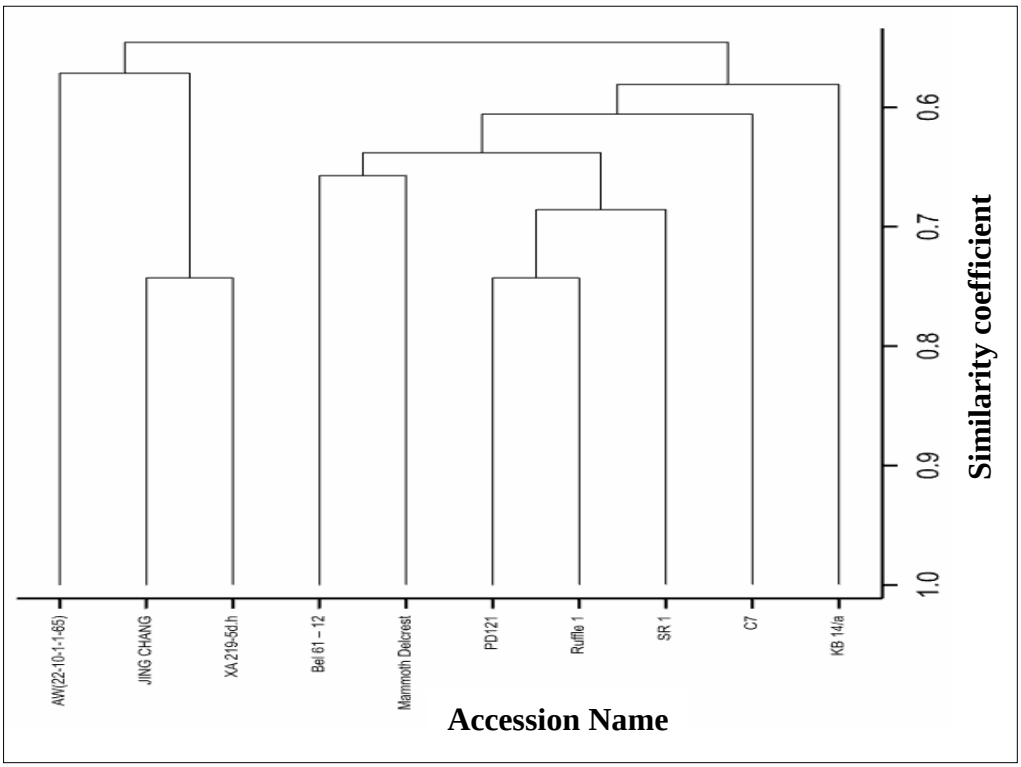


Figure 4.3: Average cluster analysis for Group 3

Cluster group 3 has ten accessions (Figure 4.3) which have formed five major branches of similar accessions. AW(22-10-1-1-65 ha its own branch like accessions C7 and KB14/a. The Chinese varieties Jing Chang and XA219-5dh formed another branch with 75% similarity. Bel61-12, Mammoth Delcrest, PD121, Ruffle 1, and SR1 formed another major branch with 65% similarity.

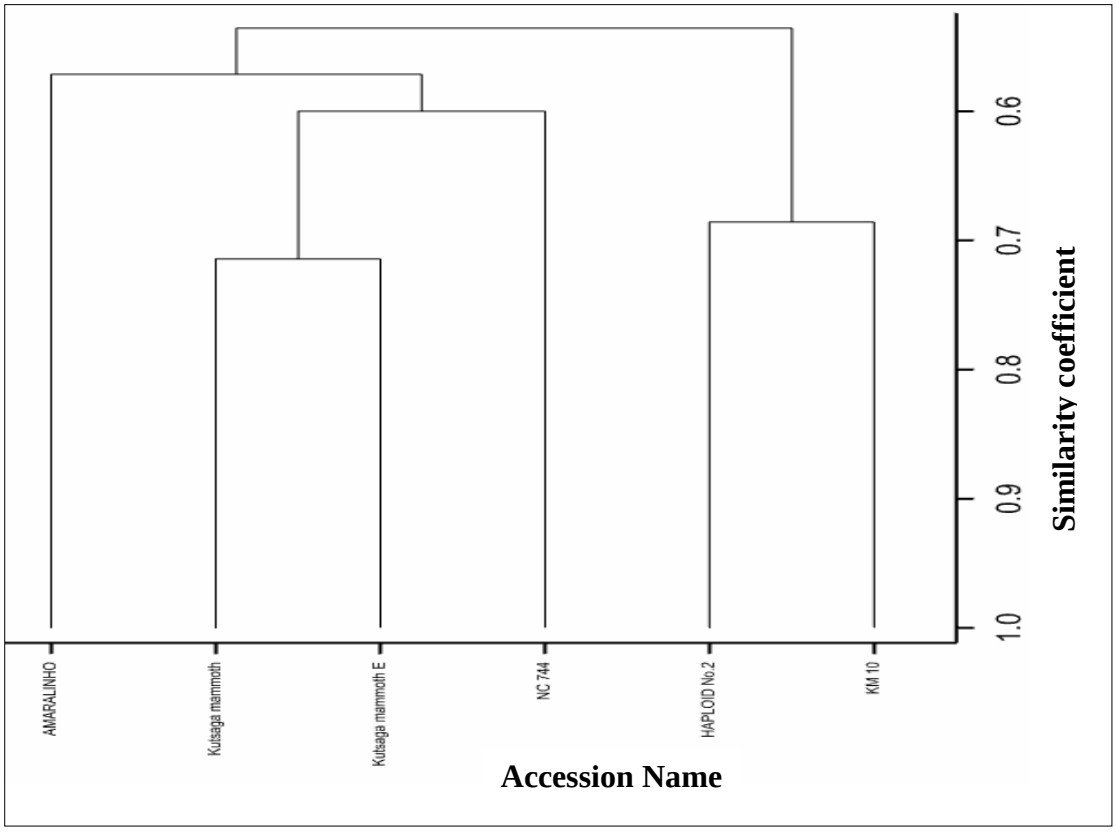


Figure 4.4: Average cluster analysis for Group 4

Group 4 is composed of a cluster of six accessions (Figure 4.4) with four major branches. Amaralinho and NC 744 each stood alone while Kutsaga Mammoth, and Kutsaga Mammoth E had a similarity of 70%. Haploid No. 2 and KM10 had a similarity of 69% and clustered together in their own group.

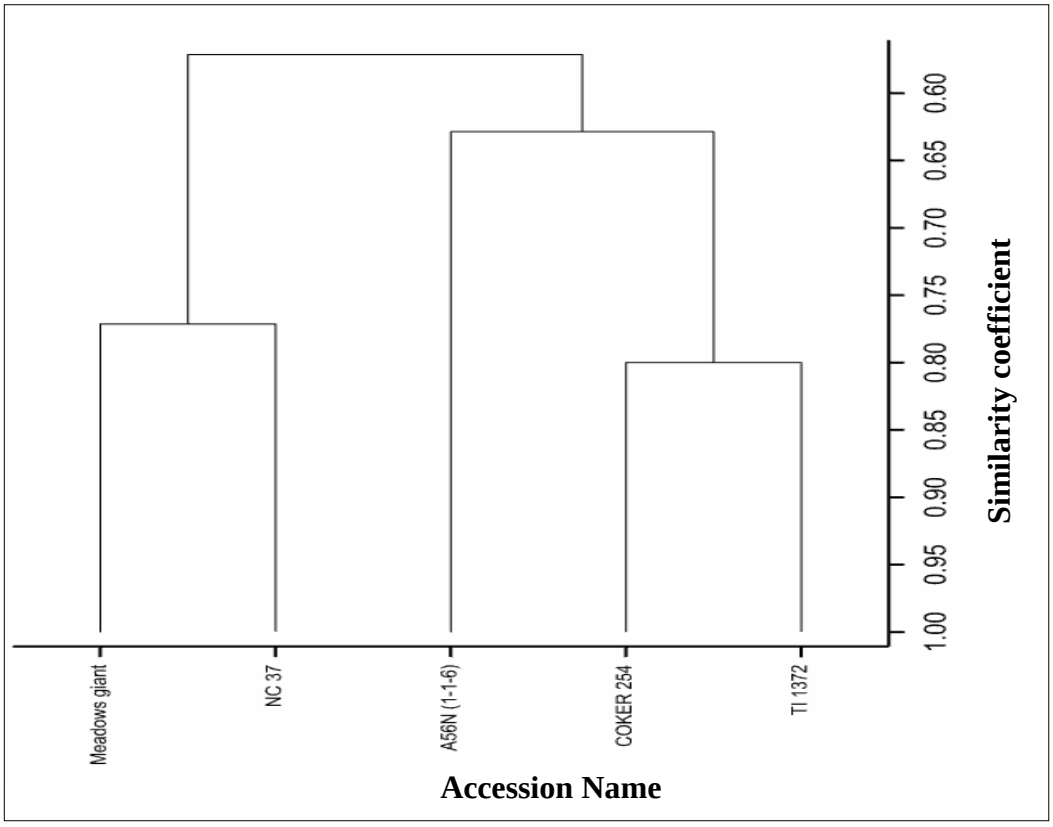


Figure 4.5: Average cluster analysis for Group 5

Group 5 has five accessions which have been clustered into three major branches. Meadows Giant and NC 37 formed the first branch with a similarity of 77.5%. Accession A56N(1-1-6) formed another branch alone while the accessions Coker 254 and TI 1372 formed another branch with a similarity of 80%.

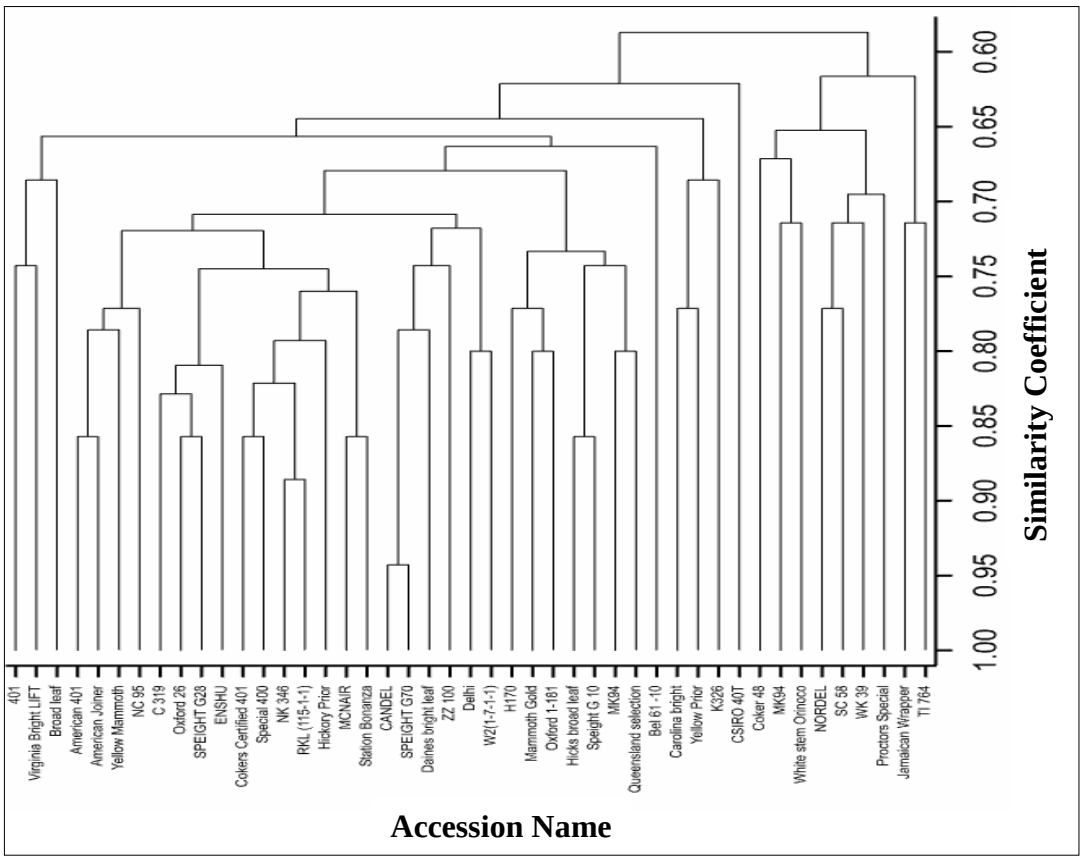


Figure 4.6: Average cluster analysis for Group 6

Group 6 has 45 accessions (Figure 4.6). The 45 accessions formed 10 major branches of similar accessions. Accession 401, Virgina Bright Lift and Broad leaf have a similarity of 69% and formed a branch together. The next branch with a similarity of 77% had accessions American 401, American Joiner, Yellow Mammoth and NC 95, which are all giant tobacco lines. The largest branch had C319, Oxford 26, Speight G28, Enshu, Coker Certified 401, Special 400, NK346, RKL(115-1-1), Hickory Prior, McNair and Station Bonanza with 75% similarity. In the same order the next group starts with Candel ending with WZ(1-7-1-1) having a similarity of 72.5%. Bel61-10 and Csiro 40T are the only accessions that clustered individually while Jamaican Wrapper and TI764 formed a branch together at a similarity of 72%. The accessions have moderate plant heights and have relatively highly pigmented stems and leaves.

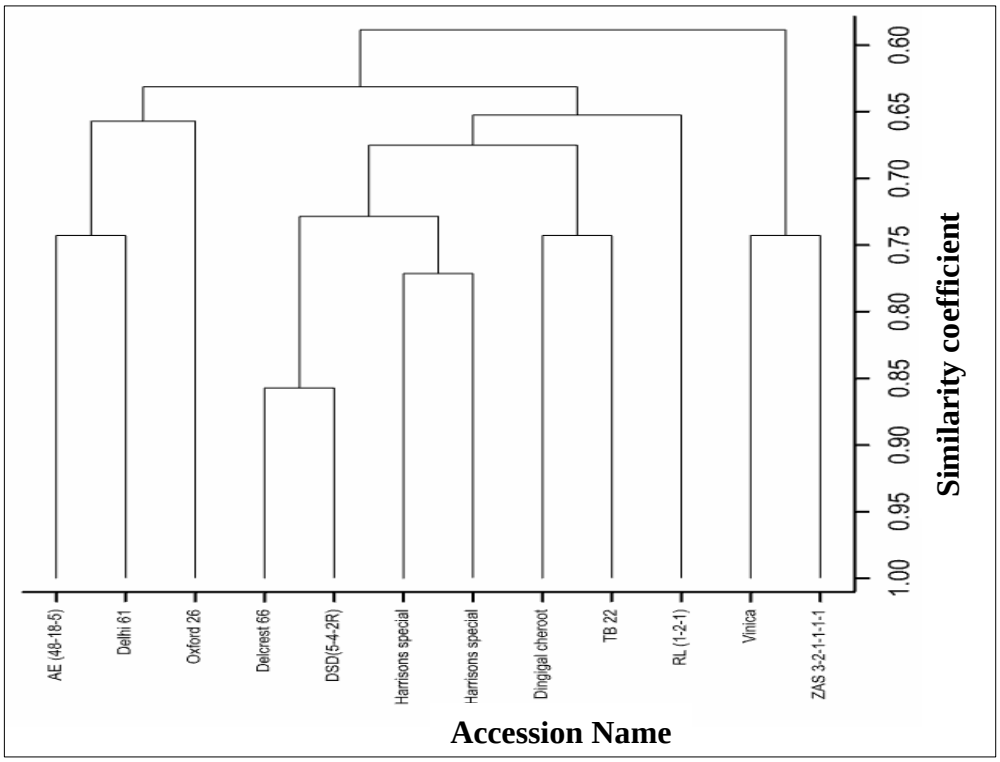


Figure 4.7: Average cluster analysis for Group 7

Group 7 is a cluster of 12 accessions, three of which are early breeding lines (AE(48-18-5), RL(1-2-1), and ZAS3-2-1-1-1-1) developed at Kutsaga Research Station. Four branches were outstanding composed of AE(48-18-5) and Delhi 61. The other cluster has Delcrest 66 to TB22. RL(1-2-1) formed its own branch alone while Vinica and ZAS3-2-1-1-1-1 clustered together with a similarity of 74%. Delcrest 66 and Delhi are Canadian introductions while the rest of the lines came from the USA. The accessions are typified by moderate plant heights and moderately green leaf and stem pigmentation.

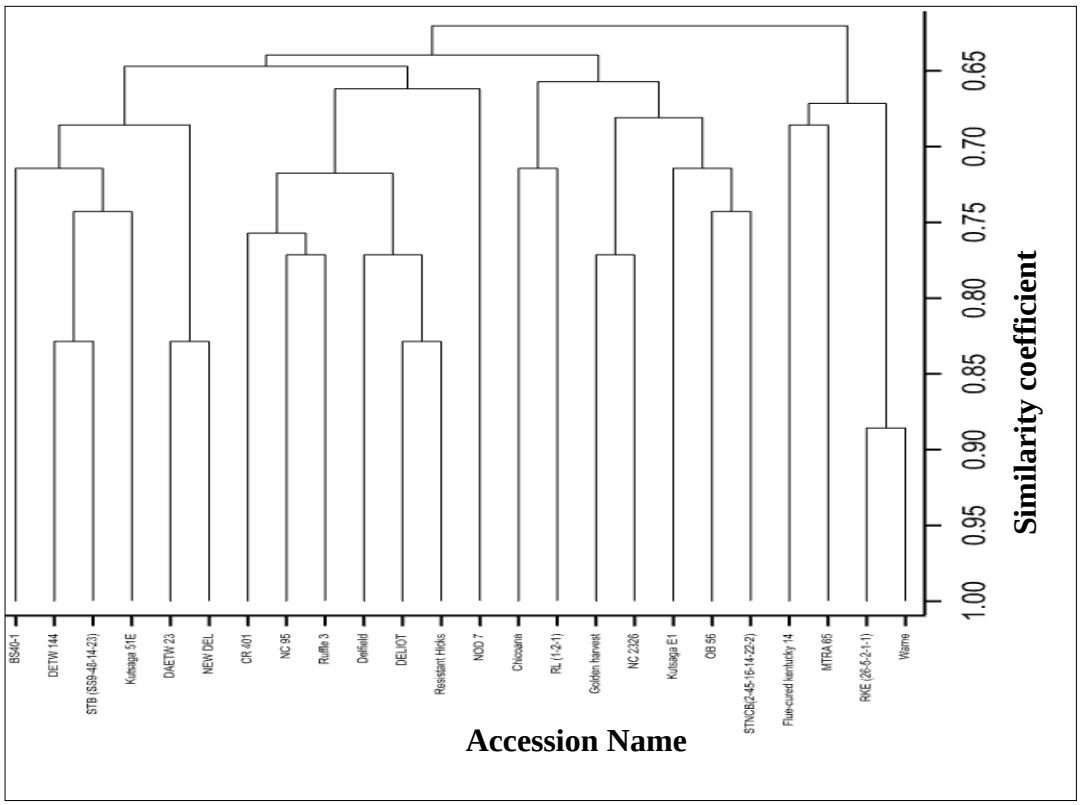


Figure 4.8: Average cluster analysis for Group 8

The dendrogram with Group 8 accessions formed 10 branches based on the similarity of the accessions. Accessions BS40-1, and NOD 7 stood each alone while DETW144, STB(598-45-14-23) and Kutsaga 51E are in one branch with a similarity of 73%. The branch with the least similarity of 69% contains accessions Flue-cured Kentucky 14 and MTRA 85 while that with the highest similarity of 88% in this group contain RKE(25-5-2-1-1) and Warna. Deliot and Resistant Hicks formed another branch at a similarity of 82%.

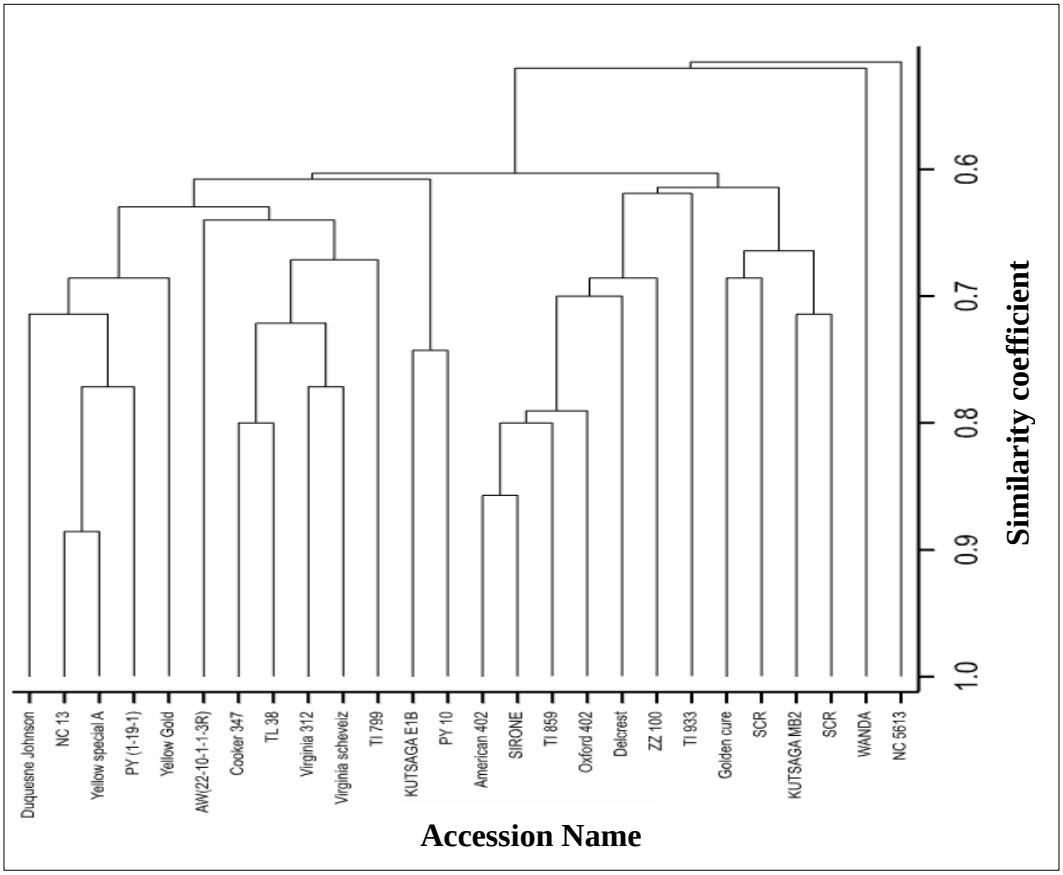


Figure 4.9: Average cluster analysis for Group 9

In this cluster of 26 accessions (Figure 4.9), four major branches of similar accessions were formed. The first branch has Doquesne Johnson, NC 13, Yellow Special A, PY(1-19-1), Yellow Gold, AW(22-10-1-1-3R), Coker 347, TL 38, Virginia 312, Virginia Schevels, TI799, Kutsaga E1B and PY10 and had a similarity of 61%. American 402, Sirone, TI 859, Oxford 402, Delcrest, ZZ100, TI 933, Golden cure, SCR, and Kutsaga MB2 formed the second branch also with a similarity of 62% among accessions. Wanda and NC 5613 stood alone in the dendrogram with similarity of below 50% from the rest of the accessions in the group. In general, the cluster had relatively short statured accessions with few, long and broad, pale leaves and long internodes.

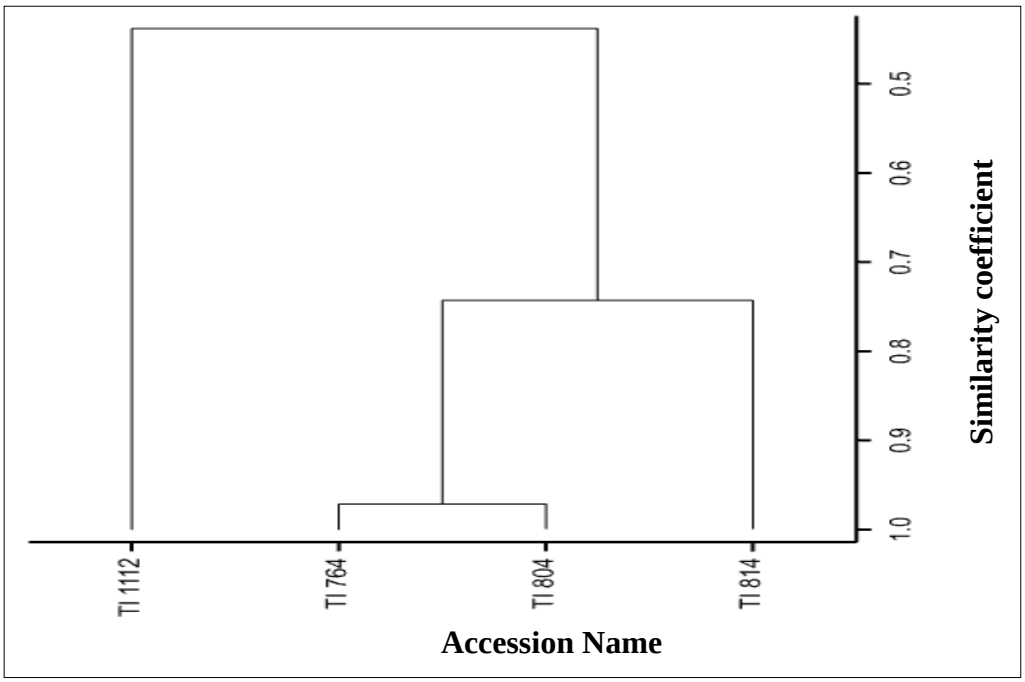


Figure 4.10: Average cluster analysis for Group 10

All the accessions in Group 10 are introductions from the USA (Figure 4.10) and formed three branches. TI 1112 and TI 814 clustered individually while accessions TI 764 and TI 804 had a similarity of 98% and formed a branch together. These accessions have flue-cured characteristics that have some wild type features. They have small leaves with pubescence. The accessions were mainly meant to be used as sources of resistance genes for root-knot nematodes during the early breeding programmes.

Table 4.1 below summarizes the characteristics of the germplasm in each of the ten (10) clusters above. Similarity coefficiencies were relatively high in each cluster. Group 2 had the highest number of germplasm while Group 10 had the least. Group 10 is made up of the oldest lines in tobacco breeding in Zimbabwe which were introduced from the USA while Group 2 is composed of segregants in the early breeding programmes in Zimbabwe at Kutsaga Research Station.

Table 4.1: General description of the characteristics of germplasm in each cluster.

Group	General characteristics of the group	Number of Accessions in the group	Highest similarity coefficient (%)	Lowest similarity coefficient (%)
1	Long, wide and relatively pale leaves with long intermediate. Plants are of intermediate height.	15	82.9	40
2	Dark green leaves of moderate length and width. The majority of lines are early introductions from the USA into Zimbabwe.	47	88.9	45.7
3	Long, broad relatively pale leaves. Tall plants with long internodes.	10	74.3	42.9
4	The group is composed of tall and broad leafed lines with many leaves and short internodes. Leaves are small and moderately recurved.	6	68.6	40.0
5	Early introductions except A56N(1-1-6) with long broad and relatively pale leaves.	5	80	51.4
6	The majority of lines in this group are early introductions which came through the USA. They are relatively short plants. Locally developed lines in this group have high root-knot nematode (*M. javanica*) resistance.	45	94.3	40.0
7	Early introductions from the USA with a few locally developed lines. Local lines have high root-knot (*M. javanica*) nematode resistance.	12	85.7	48.6

Group	General characteristics of the group	Number of Accessions in the group	Highest similarity coefficient (%)	Lowest similarity coefficient (%)
8	Early breeding lines developed at Kutsaga Research Station. Leaves are relatively small possessing high chlorophyll content. The lines in this group are classified as having high root-knot nematode (*M. javanica*) resistance.	24	88.6	51.4
9	Relatively short statured plants with long internodes and few leaves. Leaves and pale long and broad and moderately recurved.	26	88.6	37.1
10	This is a group of the earliest varieties to be introduced in the USA and later in Zimbabwe commonly known as TI varieties. They are inclined towards wild tobacco. They have small, dark green leaves with serrated margins. The plants are relatively tall and the flowers are dark pink.	4	97.1	40.0

4.4.2 The level of flue-cured tobacco accession similarity between groups

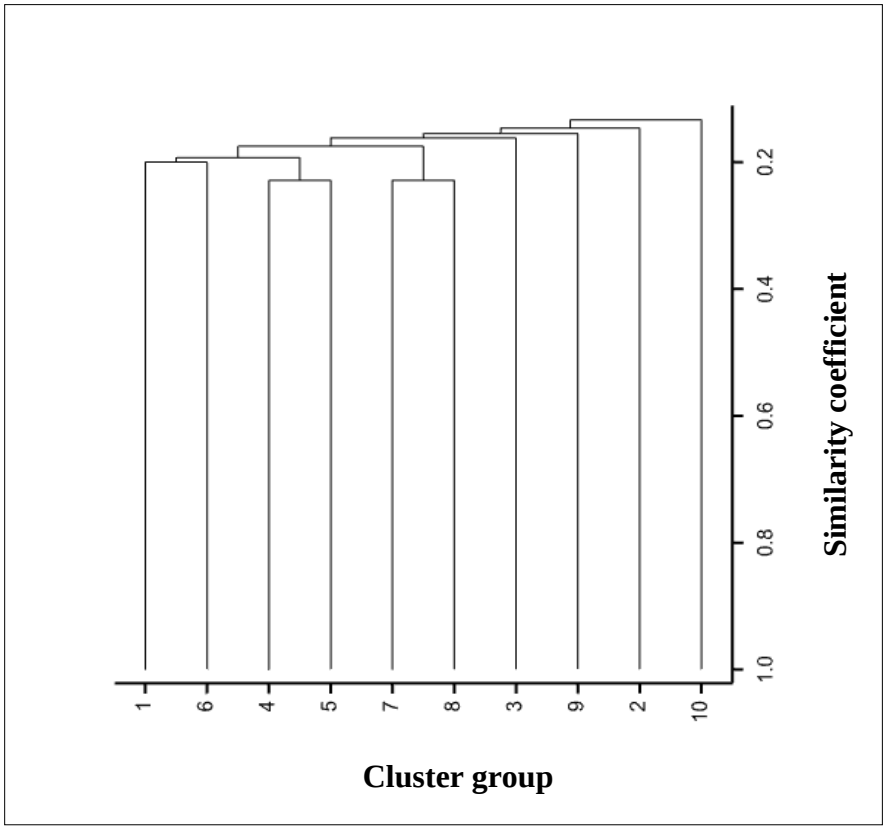


Figure 4.11: All groups average cluster analysis

Figure 4.11 shows relative similarities among the ten groups or clusters of accessions earlier formed. The range of similarity among all the groups was 11.4% to 22.9%. This showed that the groups were not similar.

4.5 Discussion

The available Zimbabwean flue-cured tobacco germplasm was effectively grouped into ten clusters based on their agro-morphological characteristics. Within each of these groups, the individual accessions had similarity coefficients of above 60% making them closely related agro-morphologically. Globally, it is generally accepted that flue-cured tobacco has limited diversity since most of it originates from the Americas (Moon *et al.*, 2009a; Raju *et al.*, 2009).

The clustering also suggested some differences in the germplasm collections available in Zimbabwe. Flue-cured tobacco germplasm in Zimbabwe was sourced from Canada, United States of America, Japan, Brazil, Germany, Belgium, South Africa and Russia. Through intercrossing among the germplasm of diverse sources, there are recombinations that truly resemble the original tobacco accessions introduced into the country during the early years of

flue-cured tobacco breeding in Zimbabwe. Because the germplasm was introduced during varying times from the late 1800s, the accessions have adapted to the local growing environment to different extents.

Group 1 is one such an important cluster in this study (Figure 4.1) as it includes some of the major progenitors of popular varieties in terms of quality and disease resistance in the Zimbabwean tobacco industry. The group had ten clusters which are formed on the basis of the origin of the accessions grouped together. The majority of the accessions in this group are early introductions, mainly for use in breeding for disease resistance, which were first utilized in tobacco improvements in the 1950s and 1960s in Zimbabwe. The early variety Florida 22 is known as one of the early donors of Alternaria resistance present in the currently popular varieties in Zimbabwe (Raeber and Smeeton, 1970). It is this American variety that was used in early breeding work for Alternaria leaf spot resistance and cured leaf quality improvement. The lines mainly originated from the United States with only one line: F111, that came from Australia. W-1-0-1 and Kutsaga 110 were developed at Kutsaga Research Station in Zimbabwe. Kutsaga 110 is the only line in the cluster that was at some point used as a commercial variety. Kutsaga 110 was also used as a donor for tobacco mosaic tabamo-virus (TMV) resistance among many of the lines used as parents in the currently grown lemon style flue-cured varieties. When cured, these lines produce lemon cured styles and are, therefore, the major donors of the genes for lemon cures in the elite flue-cured tobacco germplasm in Zimbabwe.

The majority of the accessions in Group 2 are early introductions to Zimbabwe. Apart from a few lines, the bulk of the lines were introduced from the United States of America. Bel 921 was introduced from Germany, Dixie Bright came from Canada while Nod 7, WK and MK 94 originated from South Africa. V512-MU is a mutation from the Canadian line Delcrest. Early varieties and lines developed mainly from the early introduced flue-cured lines are also in this group like Kutsaga Mammoth, Kutsaga E2, and Kutsaga 51. Some ST, TL and STNCA (ST, RKT, TL and STNCA refer to breeding lines bred for resistance to root-knot nematodes in the ealy 1980s in Zimbabwe) lines are also in this group and they formed a branch. These lines have resistance to root-knot nematodes (*M. javanica*) and share a pedigree in that their root-knot nematode resistance was derived from *N. repanda* and *N. longiflora*. . The ST, STNCA, and RKT lines are some of the lines to be bred in Zimbabwe targeting resistance to root-knot nematodes.

In Group 3, the line AW (22-10-1-1-65) which has high Alternaria leaf spot resistance, pale leaves and gives lemon cures is in this group together with Jing Chang a Chinese variety. Ruffle 1 is a mutation from the Canadian line Delcrest while PD121 is also another introduction that was extensively used in breeding for early resistance to Alternaria leaf spot. SR1 and KB 14/a are some of the early varieties developed in the 1960s.

What is common about the accessions in Group 4 is that they are all mammoth characterized by tall, giant plants with many leaves on short internodes. Mammoth varieties have relatively small and relatively pale leaves with highly serrated margins which when cured give lemon style tobacco. The Mammoth accessions including KM10 are from Zimbabwe while the others are from the USA.

Apart from A56N (1-1-6), all the other four accessions in Group 5 are early introductions (Figure 4.5). A56N (1-1-6) is one of the early lines developed for resistance to Alternaria leaf spot and lemon cured leaf styles. TI1372 is a line that was introduced from the USA and has characteristics close to wild relatives tobacco. The cluster can be summarized as composed of one line from Zimbabwe and the other four lines from USA.

Group 6 (Figure 4.6) is a cluster of accessions of diverse origins including lines developed in Zimbabwe. Accessions like Hicks, Oxford 26 and K326 originated from the USA while the lines Candel and Proctor Special came from Canada and Mcnair came from South Africa. Except for RKL (115-1-1), the rest of the accessions in this cluster are introductions from different parts of the world with the majority of the lines in the cluster having been introduced through the United States of America. Hickory Prior was introduced from Belgium while Mcnair came from South Africa and Candel came from Canada. Lines ZZ 100 and Queensland are selections made on stations in Zimbabwe. This cluster also indicated that there are morphological similarities between accessions that originated from different parts of the world which are stored in Zimbabwe and used in breeding programmes. Many early tobacco scientists reported lack of meaningful diversity in flue-cured tobacco grown in different parts of the world (Moon *et al*., 2009a).

Group 7 has four clusters. The clusters are composed of local lines and some introductions. The Accessions AE(48-18-5) and RL(1-2-1) appear to be in this group owing to their sharing a parent in their ancestry, K51. All the other accessions have American origin from which K51's progenitors were imported into Zimbabwe.

The most informative group/cluster is 8 (Figure 4.8) which includes some of the lines possessing the highest concentration of root-knot nematode resistance in Zimbabwe. A mix of early flue-cured tobacco variety introductions is found in Group 8 (Figure 4.8). Of note are the early Zimbabwe varieties, Kutsaga 51E, and Kutsaga E1, as well as the lines STNCB (2-45-15-14-22-), RL (1-2-1), STB and AKE which were bred for root-knot nematode resistance. The lines have relatively long internodes, few leaves which are long and broad as well as a short stature. The lines are the STNC and the RKE lines. The most popular early bred flue-cured tobacco variety in Zimbabwe known for giving orange cured styles, Kutsaga E1, is found in this group. Kutsaga E1 does very well in the fast growing areas like Karoi, Centenary and Chinhoyi. The most common feature between the STNC lines, Kutsaga E1 and RKE lines is the characteristic green pigmentation whose genes appear to be closely linked with the genes that confer root-knot nematode resistance in the Zimbabwean flue-cured tobacco germplasm. These lines have a common pedigree that dates back to the 1960s cross by Raeber and Schweppenhauser (1966), in an effort to breed for root-knot nematode resistance in flue-cured tobacco.

Group 9 (Figure 4.9) has accessions which are notable for long and broad leaves which are spear shaped. The majority of the accessions are early introductions which were never used in variety development programmes. When ripe the leaves in this group are pale and cure easily to give lemon style tobacco. Also, plants in this group have few leaves, hence, they have a short stature and a dark green stem that is zig-zag shaped. AW(22-10-1-1-3R) a line developed in the 1970s and now known as AW3R is one popular line in this group that gives lemon cures easily. Kutsaga MB2 and Kutsaga E1B were developed in the early 1960s and share similar morphological characteristics with AW3R. Other lines like ZZ 100 are selections made from populations developed during the early stages of flue-cured tobacco breeding in Zimbabwe. The rest of the remaining accessions in this group originated from the United states of America.

Most of the TI lines clustered together in Group 10 (Figure 4.10). The TI lines are clustering together because they are some of the earliest tobacco lines to be studied in the United States by early researchers like E. E. Clayton during the 1930s (Slana *et al*., 1976). These lines later found their way into Zimbabwe during the early 1940s when foundation experiments to breed for tobacco mosaic virus and root-knot nematode resistance in flue-cured tobacco were set up at Trelawney Research Station. The lines are still segregating for traits from wild tobacco, hence, they possess many features unacceptable to modern flue-cured tobacco farming although they are potentially rich sources of genes for resistance to many diseases.

When a cluster analysis was conducted among the ten groups of germplasm generated using hierarchical analysis, the output showed wide divergence among the groups. Similarity coefficients for eight out of the ten groups were below 20% with only two groups having similarity coefficients of slightly above 20% (Figure 4.11). The 22.9% similarity witnessed could be due to the common traits in tobacco plants which are naturally indistinguishable like leaf pigmentation and plant height. The output shows that the material clustered within each of the ten groups are closely related while there were wide intergroup differences.

4.6 Conclusions and Recommendations

4.6.1 Conclusions

1. Using TwinSpan the 196 flue-cured tobacco germplasm can be sufficiently grouped into ten groups based on agro-morphological similarity.
2. There is limited agro-morphological diversity in the flue-cured tobacco germplasm in Zimbabwe.

4.6.2 Recommendations

1. The ten clusters which are forming groups indicate degree of relatedness of the accessions. These clusters can be used as a basis for making heterotic groups in future breeding work.
2. To get highly variable offspring, it is recommended that crosses be done between lines from across the different clusters.

CHAPTER FIVE

Genetic analysis for yield, quality and agronomic traits of Zimbabwean flue-cured tobacco (*Nicotiana tabacum* L.) under normal growing conditions

5.1 Introduction

Tobacco breeding in Zimbabwe has been mainly done as an art rather than a science with most of the selection process for any traits of economic importance done traditionally using the naked eye. Evidence on the ground does not show any use of modern scientific studies such as genetic analysis of traits and their characteristics. Genetic analysis seeks to determine the gene action operating in genotypes for specific traits of interest to the plant breeder (Patel *et al*., 2012). Information like the worth of crosses, heterosis, combining ability, heritability of traits is extracted from genetic analysis data (Saroj *et al*., 2014; Yao *et al*., 2011). Knowledge of such quantitative information about traits in crops helps plant breeders to quickly make decisions about improvement of cultivars thereby increasing the rate at which varieties are made available to farmers. According to Vencovsky *et al*. (2012) genetic analysis as part of quantitative genetics allows plant breeders to make the decision of whether to use hybrids or not and to be able to select segregating populations for use in variety improvement programmes. All the information generated through genetic analysis of traits in crop improvement helps plant breeders to increase the efficiency of plant breeding programmes making it cheaper to develop new or improve existing varieties for yield and other agronomic characters.

The time it takes for a flue-cured tobacco variety to grow to eighteen leaves, which is the standard topping height in Zimbabwe, is very important in tobacco farming. Often it has been reported that varieties planted on the same day on the same farm take different times to get to topping height (TRB Flue-cured Tobacco Recommendations, 2012). This has made growers and flue-cured tobacco breeders concur that there is variation among flue cured tobacco germplasm for growth rate. Under the same field management regime, there are varieties which quickly get to topping height and those that take longer to get to the same stage (TRB Flue-cured Tobacco Recommendations, 2012). In crop husbandry, slow crop growth has always been theoretically associated with relatively high dry matter accumulation and hence high yields.

In flue-cured tobacco farming, tobacco variety growth is discussed in conjunction with the speed of ripening of the same variety. The standard acceptable rate of ripening for a variety is

two leaves per week (TRB Flue-cured Tobacco Recommendations, 2012). However, under the same management regime, it has been reported that there are some varieties which give more or less than two ripe leaves per week. Varieties that have more than two leaves ripe per week are considered fast ripening while those that give less than two leaves per week are slow (TRB Flue-cured Tobacco Recommendations, 2012). Farmers report that fast ripening varieties give low flue cured tobacco yield because they take a relatively short time in the lands to accumulate high dry matter content. They also result in losses caused by barn space limitations since hectarages planted to flue-cured tobacco are linked to the available curing barn space by assuming that varieties ripen at a rate of two leaves per week.

Too slow ripening varieties have been reported to result in inefficient utilization of flue cured tobacco curing barns (Miller, 2016 pers comm.). An ideal variety must have moderate growth and ripening rate such that under optimum management for the area, high yields and quality are obtained with minimum handling loses. To obtain a good combination of growth and ripening rate in a flue-cured tobacco variety, plant breeders must scan through the germplasm bank for material which when combined, results in the best characteristics needed by the farmer to get maximum returns.

Correct choice of material to use in a breeding programme depends on an in-depth understanding and knowledge of the characters of interest. This takes into consideration a deep understanding of the genetic properties of the traits of interest so that combinations of lines that give the desired characters in the resultant varieties are obtained (Ali and Ali, 2010). Traits such as maternal effects, dominance or epistatic effects have been found to influence the inheritance of traits (Carvalho *et al*., 2014). It is also very important to determine whether a trait is controlled by polygenes or major genes in order to be able to choose wisely the breeding method to follow in order to get the desired results in the shortest possible time and at the lowest possible cost (Crossa *et al.*, 2010).

For flue-cured tobacco breeding in Zimbabwe, although it is observed that varieties have different growth and ripening rates, no work has been done to investigate the genetic effects controlling these traits. It is, therefore, important to study the genetic effects controlling growth and ripening rate in flue-cured tobacco in order for plant breeders to successfully plan tobacco variety improvement projects.

5.2 Materials and Methods

5.2.1 Germplasm and mating design

Forty-five hybrids and their reciprocals as well as 10 parentals were generated using a 10 × 10 full diallel mating design: Method I of Griffing's (1956) using Model 1 because this method provides information about meternal and reciprocal effects within the genotypes under study. The parents were selected from lines developed from a wide range of germplasm collection spanning over 110 years improved for quality and yield over the years The diallel was set out at Kutsaga Research Station during the 2014 -15 summer season. The lines used in the diallel cross are given in Table 5.1.

Table 5.1: Lines used to constitute hybrids in the Griffing's Method 1 diallel crossing scheme

Entry code	Entry Name	Major characteristics
1	K326R1-1-28	Selection from the American line K326 known for its excellent quality. It is fast growing and fast ripening.
2	RWR3-2-12	Tall, multi-leafed line with moderate Alternaria resistance. Fast growing and ripening. This line is fast growing and fast ripening.
3	ONCR3-4-6	A high quality leaf producing line with moderate root-knot nematode resistance. It is rated as slow growing and slow ripening.
4	AW3R	Fast growing Alternaria leaf spot resistant line with big broad leaves. AW3R is fast growing and medium ripening.
5	KM10	Giant line with many leaves. Fast growing and ripening. Highly susceptible to root-knot nematodes.
6	XM26R1-1-26	A slow growing and ripening line giving predominantly lemon cures.
7	XSR4-7-10	Slow growing and ripening. Gives predominantly lemon cures.
8	BAZR3-6-18	Dark green line that gives predominantly orange to mahogany cures. It is slow growing.
9	XZR2-2-5	Slow growing and ripening. Gives predominantly lemon cures
10	KE1	TRB line known for its thick few leaves and recommended for hot environments. It gives predominantly deep orange cures. It is slow growing and ripening.

5.2.2 Land Preparation and crop management

For land preparation, raising of seedlings, planting, fertilizer application, weed and pest management refer to **Section 4.2.2.1** to **Section 4.2.2.3**.

5.2.3 Field experimental design and data collected

The constituted hybrids, reciprocals and parental lines were evaluated in field trials using a 10 × 10 α-lattice design with two replicates and one check variety during the 2015/16 and 2016/17 summer seasons at Kutsaga Research Station. The check variety was not included in the ANOVA. Data recorded included ripening rate (time taken to completely strip 50% leaves of plants of a variety in a plot), growth rate (number of days to 50% topping height), and flue-cured tobacco leaf colour.

5.2.4 Leaf priming, topping and sucker removal

The lower three seedbed leaves of the tobacco plants in the field were removed four weeks after planting in a process called “leaf priming” in order to find a clear base from which to count 18 leaves for topping. When the plants attained 18 leaves, the apical bud of each plant was removed in a process called “topping” in tobacco farming. Each time, a topping process was done in the experiment; a record of the number of topped plants was made until all the plants were fully decapitated. A mixture of systemic suckercide Acotab® and a contact suckercide N- Decanol® was applied on the topped plants to suppress sucker development. Manual sucker removal was also done whenever suckers appeared in the experiment later in the season.

5.2.5 Reaping, curing and grading

Reaping was done manually whenever the leaves were ripe on each plot. Reaped leaves were put on Tilita clips and a label with plot information attached to the clip. Loaded clips were quickly transported to the curing barns for the flue-curing process. The curing process involved a controlled moisture reduction process from the ripe leaf until the leaf was completely dry. Heat in the curing barns was supplied from burning coal and moved through the barns in flue pipes hence “flue-cured tobacco”. The curing process took an average of seven days to complete. After the curing process, the tobacco is conditioned by applying steam to facilitate easy handling without breakages during the grading process. The grading process was done following the TIMB Guidelines (TIMB, 1999). The data on grade distribution and quality distribution was recorded.

Table 5.2: List of entries evaluated in a field trial

Cross Code	Cross Identity	Cross Code	Cross Identity	Cross Code	Cross Identity
1	K326R1-1-28	34	AW3R	68	BAZR3-6-18 × XSR4-7-10
2	RWR3-2-12 × K326R1-1-28	35	K M10 × AW3R	69	XZR2-2-5 × XSR4-7-10
3	ONCR3-4-6 × K326R1-1-28	36	XM26R1-1-26 × AW3R	70	K E1 × XSR4-7-10
4	AW3R × K326R1-1-28	37	XSR4-7-10 × AW3R	71	K326R1-1-28 × BAZR3-6-18
5	K M10 × K326R1-1-28	38	BAZR3-6-18 × AW3R	72	RWR3-2-12 × BAZR3-6-18
6	XM26R1-1-26× K326R1-1-28	39	XZR2-2-5 × AW3R	73	ONCR3-4-6 × BAZR3-6-18
7	XSR4-7-10× K326R1-1-28	40	K E1 × AW3R	74	AW3R × BAZR3-6-18
8	BAZR3-6-18 × K326R1-1-28	41	K326R1-1-28 × K M10	75	K M10 × BAZR3-6-18
9	XZR2-2-5 × K326R1-1-28	42	RWR3-2-12 × K M10	76	XM26R1-1-26 × BAZR3-6-18
10	KE1 × K326R1-1-28	43	ONCR3-4-6 × K M10	77	XSR4-7-10 × BAZR3-6-18
11	K326R1-1-28 × RWR3-2-12	44	AW3R × K M10	78	BAZR3-6-18
12	RWR3-2-12	45	K M10	79	XZR2-2-5 × BAZR3-6-18
13	ONCR3-4-6 × RWR3-2-12	46	XM26R1-1-26 × K M10	80	K E1 × BAZR3-6-18
14	AW3R × RWR3-2-12	47	XSR4-7-10 × K M10	81	K326R1-1-28 × XZR2-2-5
15	K M10 × RWR3-2-12	48	BAZR3-6-18 × K M10	82	RWR3-2-12 × XZR2-2-5
16	XM26R1-1-26 × RWR3-2-12	49	XZR2-2-5 × K M10	83	ONCR3-4-6 × XZR2-2-5
17	XSR4-7-10 × RWR3-2-12	50	K E1 × K M10	84	AW3R × XZR2-2-5
18	BAZR3-6-18 × RWR3-2-12	51	K326R1-1-28 × XM26R1-1-26	85	K M10 × XZR2-2-5

Cross Code	Cross Identity	Cross Code	Cross Identity	Cross Code	Cross Identity
19	XZR2-2-5 × RWR3-2-12	52	RWR3-2-12 × XM26R1-1-26	86	XM26R1-1-26 × XZR2-2-5
20	K E1 × RWR3-2-12	53	ONCR3-4-6 × XM26R1-1-26	87	XSR4-7-10 × XZR2-2-5
21	K326R1-1-28 × ONCR3-4-6	54	AW3R × XM26R1-1-26	88	BAZR3-6-18 × XZR2-2-5
22	RWR3-2-12 × ONCR3-4-6	55	K M10 × XM26R1-1-26	89	XZR2-2-5
23	ONCR3-4-6	56	XM26R1-1-26	90	K E1 × XZR2-2-5
24	AW3R × ONCR3-4-6	57	XSR4-7-10 × XM26R1-1-26	91	K326R1-1-28 × K E1
25	K M10 × ONCR3-4-6	58	BAZR3-6-18 × XM26R1-1-26	92	RWR3-2-12 × K E1
26	XM26R1-1-26 × ONCR3-4-6	59	XZR2-2-5 × XM26R1-1-26	93	ONCR3-4-6 × K E1
27	XSR4-7 -10 × ONCR3-4-6	60	K E1 × XM26R1-1-26	94	AW3R × K E1
28	BAZR3-6-18 × ONCR3-4-6	61	K326R1-1-28 × XSR4-7-10	95	K M10 × K E1
29	XZR2-2-5 × ONCR3-4-6	62	RWR3-2-12 × XSR4-7-10	96	XM26R1-1-26 × K E1
30	K E1 × ONCR3-4-6	63	ONCR3-4-6 × XSR4-7-10	97	XSR4-7-10 × K E1
31	K326R1-1-28 × AW3R	64	AW3R × XSR4-7-10	98	BAZR3-6-18 × K E1
32	RWR3-2-4 × AW3R	65	K M10 × XSR4-7-10	99	XZR2-2-5 × K E1
33	ONCR3-4-6 × AW3R	66	XM25R1-1-26 × XSR4-7-10	100	K E1
		67	XSR4-7-10	101	K RK26R

5.2.6 Parameters measured

The parameters measured were growth rate (days to 50% topping), ripening rate (days to 50% reaping), and cured leaf colour (proportion of the cured leaf classified as orange and mahogany per plot). The parameters are as described below:

5.2.6.1 Growth rate

In tobacco farming, "growth rate" refers to the time taken by a variety to develop the standard topping height of 18 leaves from transplanting. In this study "growth rate" refers to the number of days taken by 50% of plants of a specific variety in a plot to develop to the standard topping height of 18 leaves.

5.2.6.2 Ripening rate

In tobacco framing, "ripening rate" is the time, in days; it takes for tobacco leaves to ripen to a harvestable state. The standard ripening rate of a flue-cured tobacco variety in Zimbabwe is two leaves per week. In this study, "ripening rate" was measured as the time (days) it took for 50% of plants in a plot to be stripped of leaves from the day of initial harvesting.

5.2.6.3 Dark cured leaf proportion

A flue-cured tobacco variety has the potential to produce either lemon, orange or mahogany cured leaves in different proportions. Orange and mahogany cured leaf styles are collectively referred as "Dark style tobacco" while the lemon is referred to as "light style tobacco". In this study, the harvested and flue cured leaves were graded and classified into lemon, orange and mahogany colours and weighed. Of the total tobacco leaf weight from each plot, the proportion of the "dark style" tobacco (orange and mahogany) was determined and referred to as dark style tobacco.

5.2.6.4 Yield (mass at untying and saleable yield)

In tobacco, yield refers to the mass of the cured leaf per unit area (usually reported kg/ha). When the cured leaf fresh from curing is weighed, it gives the first component of yield called mass at untying. After determining mass at untying, the cured leaf is graded to give another component of yield called saleable yield. Saleable yield is the yield component that can the taken to the market. In this study both saleable yield and mass at untying are reported.

5.3 Data analysis

The data were analysed using the software Diallel SAS 05 (Zhang *et al*., 2005). All the other analyses were done using Genstat statistical package Version 13 (Genstat, 2011). An estimation

of the effects of each pair of parents for specific combining ability (SCA) and general combining ability (GCA) was done on all traits and the relative importance of general combining ability (GCA) to specific combining ability (SCA) was estimated for all the traits measured. The GCA and SCA effects were also used to estimate the broad and narrow sense coefficient of genetic determination (R^2) (fixed parent equivalent of heritability) and the magnitude of the maternal effects.

Heritability in the narrow (h^2) and broad sense (H^2) and genetic advance (GA) for pooled data across seasons was estimated for all traits. Heritability was estimated depending on the variance of GCA and SCA , and on the variance of experimental error according to Singh and Chaudhary (1985) as follows:

$$H_{b.s} = H^2 = \frac{\sigma^2_G}{\sigma^2_P} = \frac{\sigma^2_A + \sigma^2_D}{\sigma^2_A + \sigma^2_D + \sigma^2_e} = \frac{2\sigma^2_{GCA} + \sigma^2_{SCA}}{2\sigma^2_{GCA} + \sigma^2_{SCA} + \sigma^2_e}$$

$$H_{n.s} = h^2 = \frac{\sigma^2_A}{\sigma^2_P} = \frac{\sigma^2_A}{\sigma^2_A + \sigma^2_D + \sigma^2_e} = \frac{2\sigma^2_{GCA}}{2\sigma^2_{GCA} + \sigma^2_{SCA} + \sigma^2_e}$$

Where: $H_{b.s}$: Heritability in broad sense (H^2),

$H_{n.s}$: Heritability in narrow sense (h^2),

σ^2_{GCA} : The variance of general combining ability,

σ^2_{SCA} : The variance of specific combining ability,

σ^2_e : The variance of experimental error i.e. environmental variance,

σ^2_A : Additive genetic variance,

σ^2_D : Non-additive (dominance and epistasis) genetic variance,

σ^2_G : Total genetic variance, and

σ^2_P : Phenotypic variance (genetic and environmental variance).

Heritability estimate was categorized as low when less than 40%, moderate when 40-59%, moderately high when 60-79% and very high when above 80% as suggested by Singh (2001).

GA values were determined according to formular illustrated by Jonson *et al*. (1955) where:

$GA = K\delta_p \times H^2$

Where GA is genetic advance, K = 2.063 (selection differential at 5%), δ_p = phenotypic standard deviation of mean performance of hybrid, H^2 = broad sense heritability, × = multiplied by.

The diallel fixed effects Model I for analysis of combining ability was used as given below:

$Y_{ij} = u + g_i + g_j + s_{ij} + r_{ij} + 1/bc\sum\sum e_{ijkl}$

Where,

u = population mean,

g_i and g_j = GCA effects of i^{th} and j^{th} parents,

s_{ij} = SCA effect of the cross between i^{th} and j^{th} parents ($s_{ij} = s_{ji}$),

r_{ij} = reciprocal effect involving reciprocal crosses between i^{th} and j^{th} parents ($r_{ij} = -r_{ji}$),

b = number of blocks,

c = number of crosses, and

e_{ijkl} = environmental effect associated with the $ijkl^{th}$ individual observation.

Hayman diallel analysis was used to supplement the Griffings analysis to determine genetic parameters for the traits studied. The Hayman Model (Hayman, 1954b) used is shown below:

Y = U + rep + a + b + c + d + a*rep + b*rep + c*rep + d*rep

Where,

U = grand mean

Rep = replication effects

a = additive effects

b = dominance effects, partitioned into:

b_1, which indicates direction of dominance (unidirectional if significant)

b_2, tests asymmetry of alleles and

b_3, shows that some dominance is peculiar to some crosses

b*rep partitioned into:

rep*b_1, rep*b_2 and rep*b_3

c = additive maternal effects

d = maternal interaction effects

a*rep + b*rep + c*rep + d*rep = interaction of the reps with the model components.

5.4 Results

In this chapter, data analysis was done using Griffing's diallel method 1 and the Hayman's analysis approach. Results are therefore divided into Griffing's diallel method 1 results and the Hayman's analysis results.

5.4.1 Griffing's diallel method 1

This section presents results of the across seasons Griffing's diallel method 1 analysis. In this section, combined seasons refer to the 2015/16 and the 2016/17 cropping seasons.

5.4.1.2 Griffing's diallel method 1 combined season analysis

However, the genotypes were not significant for days to 50% reaping and days to 50% topping. General combining ability (GCA) effects for flue-cured tobacco colour were also significant ($P<0.001$) whilst the GCA effects for days to 50% reaping and days to 50% topping were not significant (Table 5.3). Specific combining ability (SCA) effects, reciprocal, maternal, and non-maternal effects were not significant for all the traits studied. Within seasons, only the maternal effects were significant for cured leaf colour, the other sources of variance were not significant.

Table 5.3: Griffing's diallel combined ANOVA for days to 50% reaping, days to 50% topping and flue-cured tobacco leaf colour

Source	Degrees of freedom	Days to 50% reaping	Days to 50% topping	Proportion of dark flue-cured tobacco leaf colour
Season	1	32222.49***	4248.16***	291.62
Rep(Season)	2	50.91***	1060.31***	26.03
Genotypes	99	40.37	10.82	180.45**
GCA	9	278.86	41.26	1158.43***
SCA	45	10.24	11.50	144.66
Reciprocal	45	48.02	7.84	101.81
Maternal	9	178.38	11.42	214.42
Non Maternal	36	15.43	6.95	73.66
Season: Genotypes	99	37.18	10.06	101.91
Season: GCA	9	166.84	35.65	111.95
Season: SCA	45	11.53	7.15	117.88
Season: Reciprocal	45	47.53	11.45	123.91
Season: Maternal	9	194.29	24.90	245.95*
Season: Non Maternal	36	10.84	8.09	93.39
Residual	162	7.93	8.18	124.10

NB: GCA – General Combining Ability; SCA – Specific Combining Ability. Genotypes in the ANOVA include hybrids, reciprocals, and parentals. In tobacco there is no significant heterosis warranting separation of hybrids and parentals in experiments.

Table 5.4 shows genetic variance for the three traits under study. There were relatively high GCA variance for days to 50% reaping and flue-cured leaf colour. However, days to 50% topping had low GCA variance. SCA variance for days to 50% reaping and days to 50% topping were also low while that for flue-cured leaf colour was relatively high at 5.14. For all the traits, maternal variance was relatively low. GCA-SCA ratio showed that the GCA variance was more important than the SCA variance for days to 50% reaping and flue-cured leaf colour. For days to 50% topping, the SCA variance component were more important. Phenotypic variance was more pronounced in days to 50% reaping and flue-cured leaf colour, but low for

days to 50% topping. The trait days to 50% topping had a low broad sense heritability of only 23%. The broad sense heritability for all the traits was low with that for days to 50% reaping (ripening rate) being 44% and that for days to 50% topping (growth rate) being 23% (Table 5.4). Flue-cured tobacco colour heritability was 30%. Genetic advance values ranged from 1.57% for the trait days to 50% topping to 8.46% for the trait dark cured leaf colour (Table 5.4).

Table 5.4: Genetic effects for parameters measured across seasons

Effects	GCA effects	SCA effects	Maternal effects	Non Maternal effects	GCA-SCA ratio	Genotypic variance	Phenotypic Variance	Narrow Heritability	Broad Heritability	Genetic Advance
Days to 50 % reaping	6.77	0.58	4.07	1.87	11.73	7.35	32.08	0.42	0.44	5.14
Days to 50 % ropping	0.83	0.83	0.11	0.00	0.99	1.66	10.89	0.15	0.23	1.57
Dark flue-cured leaf Colour	25.86	5.14	3.52	0.00	5.03	31.00	187.99	0.28	0.30	8.49

NB: GCA – General Combining Ability; SCA – Specific Combining Ability; GCA-SCA ratio

GCA effects showed the contribution made by each parent to the final phenotype of the trait under study. The parents K326R1-1-28, RWR3-2-12, AW3R, and XSR4-7-10 made significant negative contributions in reducing the days to 50% reaping of tobacco in this set of germplasm (Table 5.5). Parents ONCR3-4-6, BAZR3-6-18, K M10, XZR2-2-5, and K E1 made significant positive contributions by increasing the days to 50% reaping required in the hybrids possessing them as parents. K M10 and XZR2-2-5 were the only parents that showed significant contributions to days to 50% topping with K M10 contributing by reducing days to 50% topping by 1.6 days and XZR2-2-5 increasing days to 50% topping by 1 day. Parents ONCR3-4-6, AW3R, and XM26R1-1-26 while KE1 tended to give a decrease in dark flue-cured leaf colour.

Table 5.5: Parental GCA effects for days to 50% reaping, days to 50% topping and flue-cured leaf colour measured across seasons

Parent	Days to 50% reaping	Days to 50% topping	Dark flue-cured leaf colour
K326R1-1-28	-2.45***	-0.59	0.04
RWR3-2-12	-1.39**	0.24	-1.28
ONCR3-4-6	2.60***	0.32	4.58**
AW3R	-1.54***	-0.20	5.16**
KM10	1.15**	-1.55***	-0.06
XM26R1-1-26	-0.57	0.32	3.54*
XSR4-7-10	-2.25***	-0.28	-0.27
BAZR3-6-18	0.80*	0.27	-1.72
XZR2-2-5	2.14***	1.17**	-2.19
KE1	1.52***	0.30	-7.79***
Standard error	0.30	0.30	1.18

5.4.1.3 Griffing's diallel method 1 season 1 analysis results

In this section, results of the Griffing's diallel method 1 analysis are presented. Throughout this section, season 1 refers to the 2015/16 cropping season.

Table 5.6: Griffing's ANOVA for season 1 (2015/16) for the traits: days to 50% reaping, days to 50% topping, and dark flue-cured leaf colour measured at Kutsaga research station

Source	Degrees of freedom	Days to 50% reaping	Days to 50% topping	Dark flue-cured leaf colour
Rep	1	95.01***	21.34***	33.74***
Genotypes	99	17.36**	9.06	137.16*
GCA	9	64.42***	27.01**	612.57***
SCA	45	10.94	7.378724	110.8
Reciprocal	45	17.31*	7.393603	72.22
Maternal	9	14.89	15.92**	66.89
Non Maternal	36	17.92*	5.26	73.55
Residual	81	9.85	9.93	91.67

NB: GCA – General Combining Ability; SCA – Specific Combining Ability

At Kutsaga Research Station during season 1, the genotypes were significant for days to 50% reaping ($P<0.01$) and dark flue-cured leaf colour ($P<0.05$) (Table 5.6). The GCA effects were significant ($P<0.01$) for all the traits measured while reciprocal effects were significant($P<0.05$) only for the trait days to 50% reaping. Maternal effects were significant for days to 50% topping. Non maternal effects were significant ($P<0.05$) only for the trait days to 50% reaping.

Table 5.7: Genetic effects for season 1 (Kutsaga Research Station) for the traits days to 50% reaping, days to 50% topping, and dark flue-cured leaf colour

Effects	GCA effects	SCA effects	Maternal effects	Non-Maternal effects	GCA-SCA ratio	Genotypic variance	Phenotypic Variance	Narrow Heritability	Broad Heritability	Genetic Advance
Days to 50 % reaping	1.36	0.27	0.00	2.02	5.04	1.63	14.87	0.18	0.20	1.59
Days to 50 % topping	0.43	0.00	0.27	0.00	0.00	0.43	11.32	0.08	0.08	0.55
Dark flue-cured leaf colour	13.02	4.79	0.00	0.00	2.72	17.81	122.5	0.21	0.25	5.71

NB: GCA - General Combining Ability; SCA - Specific Combining Ability; GCA-SCA ratio

There were very small SCA variance for days to 50% reaping and zero SCA variance for days to 50% topping. Dark flue-cured tobacco colour had a relatively high positive SCA variance of 4.79 (Table 5.7). Maternal variance was only present for days to 50% topping although effects were small. Non-maternal variance was only present for days to 50% reaping. GCA-SCA ratio shows the importance of GCA over SCA variance for days to 50% reaping and dark flue-cured leaf colour. Relatively low narrow sense and broad sense heritability estimates were reported for all the traits. Phenotypic variance component was high for dark flue-cured leaf colour. Calculation of genetic advance showed low values of not more than 5.71% for all the traits studied (Table 5.7).

Table 5.8: Parental GCA effects for the traits days to 50% reaping, days to 50% topping and dark flue-cured leaf colour measured at Kutsaga research station during 2015/2016 season

Parent	Days to 50 % reaping	Days to 50 % topping	Dark flue-cured leaf colour
K326R1-1-28	-0.39	0.62	0.17
RWR3-2-12	-0.88	0.37	-1.54
ONCR3-4-6	2.78***	0.96	5.20**
AW3R	-1.08*	-0.40	4.82**
KM10	-0.90	-1.46*	-1.68
XM26R1-1-26	1.02	0.24	3.52*
XSR4-7-10	-1.01	-0.20	2.23
BAZR3-6-18	-0.47	-0.78	-3.72*
XZR2-2-5	1.23*	1.17*	-2.17
KE1	-0.29	-0.53	-6.82**
Standard error	0.47	0.47	1.44

The parental lines ONCR3-4-6 and XZR2-2-5 had significant ($P<0.05$) positive contributions to prolonging the days to 50% reaping during season 1 (Table 5.8). The line AW3R had the effect of quickening ripening during the same season. The line KM10 significantly ($P<0.05$) quickened growth by 1.5 days while the line XZR2-2-5 slowed growth by one day (Table 5.8). The lines AW3R, ONCR3-4-6 and XM26R1-1-26 made the cured leaf darker while the lines BAZR3-6-18 and K E1 made the cured leaf lighter.

5.4.1.4 Griffing's diallel method 1 Season 2 analysis results

In this section, results of the Griffing's diallel method 1 analysis are presented. Throughout this section, season 2 refers to the 2016/17 cropping season.

Table 5.9: Griffing's ANOVA for season 2 for the traits days to 50% reaping, days to 50% topping and flue-cured leaf colour measured at Kutsaga

Source	Degrees of freedom	Days to 50% reaping	Days to 50% topping	Dark flue Cured leaf colour
REP	1	1.31	1062.24***	33.05
Genotypes	99	61.54***	11.31**	157.36
GCA	9	379.82***	48.62***	748.16
SCA	45	10.85**	9.22	141.76
Reciprocal	45	78.53***	11.96**	155.94
Maternal	9	359.16***	21.13*	388.96**
No Maternal	36	8.37	9.67*	97.68
Residual	81	5.99	6.51	159.15

NB: GCA – General Combining Ability; SCA – Specific Combining Ability

The genotypes, GCA, reciprocal effects, and maternal effects were all significant ($P<0.01$) for the traits days to 50% reaping and days to 50% topping (Table 5.9). SCA effects were significant ($P<0.01$) for days to reaping while non maternal effects were significant ($P<0.05$) for days to 50% topping. However, for dark flue-cured leaf colour, only maternal effects were significant ($P<0.01$).

Table 5.10: Genetic effects for season 2 for the traits days to 50% reaping, days to 50% topping, and dark flue-cured leaf colour

Effects	GCA effects	SCA effects	Maternal effects	Non Maternal effects	GCA-SCA ratio	Genotypic variance	Phenotypic Variance	Narrow Heritability	Broad Heritability	Genetic Advance
Days to 50 % reaping	14.73	0.00	7.28	0.00	0.00	14.73	203.17	0.14	0.14	4.12
Days to 50 % topping	9.35	1.21	8.77	0.59	7.70	10.56	44.03	0.42	0.45	6.16
Dark flue-cured leaf colour	1.05	0.68	0.29	0.79	1.55	1.73	10.65	0.20	0.26	1.75

NB: GCA – General Combining Ability; SCA – Specific Combining Ability; GCA -SCA ratio = Baker's ratio

The GCA variance for days to 50% reaping were high at 14.73 during the 2016-17 season (Table 5.10). Dark flue-cured leaf colour was the least with a GCA variance of 1.05. There were no SCA variance for days 50% reaping while SCA effects were very small for dark flue-cured leaf colour. Maternal components were relatively high for the traits days to 50% reaping and topping while there were very small maternal components for dark flue-cured leaf color. GCA-SCA ratio was 0 for the trait days 50 % reaping while it was above 1 for the traits days to 50% topping and dark flue cured leaf colour. Narrow sense heritability was moderate for days to 50% topping at 42% while it was very low for days to 50% reaping at 14 %. The narrow sense heritability for dark flue-cured leaf colour was also very low at 20%. Values for genetic advance were relatively low for all the traits studied ranging from 1.75% for the trait dark flue cured leaf colour to 6.16% for the trait days to 50% topping.

Table 5.11: Parental GCA effects for season 2 for the traits days to 50% reaping, days to 50% topping and dark flue-cured leaf colour

Parent	Days to 50% reaping	Days to 50 % topping	Dark flue-cured leaf colour
K326R1-1-28	-4.51***	-1.80**	-0.35
RWR3-2-12	-1.90***	0.04	-1.68
ONCR3-4-6	2.41***	-0.63	3.92
AW3R	-1.99***	0.03	5.93*
KM10	3.18***	-1.61**	1.78
XM26R1-1-26	-2.16***	0.25	4.14
XSR4-7-10	-3.49***	-0.02	-2.88
BAZR3-6-18	2.09***	1.27**	-0.13
XZR2-2-5	3.07***	1.33**	-1.56
KE1	3.31***	1.14*	-9.16***
Standard error	0.36	0.38	1.89

The parents K326R1-1-28, RWR3-2-12, XM26R1-1-26, XSR4-7-10 and AW3R showed negative and significant GCA effects on number of days to 50% reaping during the 2016-17 season (Table 5.11). The parents ONCR3-4-6, KM10, BAZR3-6-18, XZR2-2-5 and K E1 showed positive and significant GCA effects on days to 50% reaping. The lines K326R1-1-28, and K M10, also significantly increased the growth rate as shown by reduced time from planting to topping. BAZR3-6-18, XZR2-2-5 and K E1 significantly slowed down the growth of the varieties as shown by an increase in the days to 50% topping recorded (Table 5.11). AW3R is the only parent that significantly contributed by increasing the dark colour component of the cured leaves while K E1 did the opposite i.e. increasing the light colour of the cured leaves.

5.4.2 Hayman diallel analysis results

Table 5.12: General analysis of variance

Change	d.f.	Days to 50% reaping	Days to 50% topping	Dark cured leaf colour
Year	1	58418.89***	15500.28***	1887.90***
Year.Rep	2	92.31***	3882.86***	168.50
Year.Rep.Block	36	107.56***	32.05***	665.60***
Genotype	99	28.55***	10.78*	183.10*
Year.Genotype	99	30.43***	9.96	107.00
Residual	162	7.86	8.11	124.10
Total	399	174.40	70.11	188.00

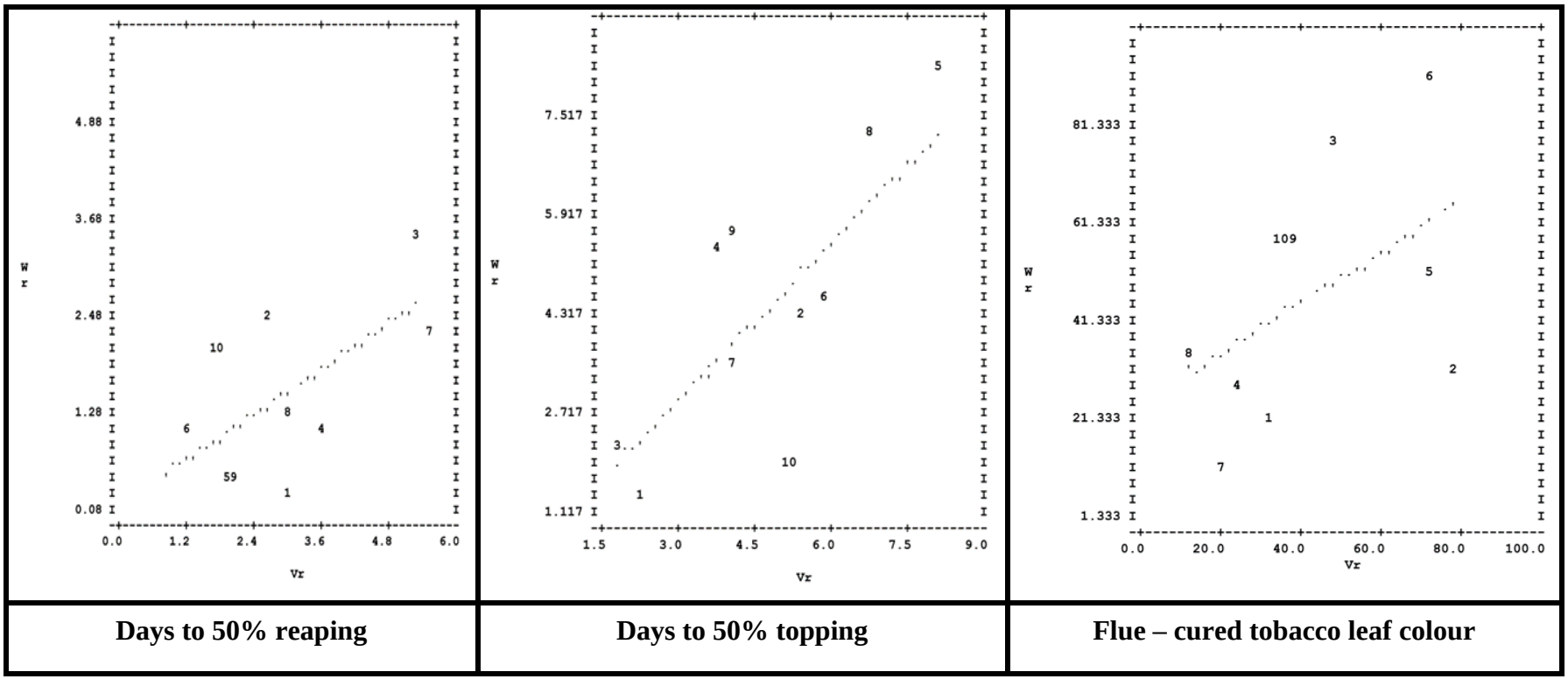
Error df for days to 50% topping was 159 while the df for year.rep, year.rep.block, year x genotype and error were 1, 27, 90 and 81 respectively.

There were significant differences among genotypes for all the three traits measured: days to 50% topping (growth rate) ($P<0.05$), days to 50% reaping (ripening rate) ($P<0.001$) and dark cured leaf colour ($P<0.05$). Hence, the genetic parameters for all the three traits can be estimated using the Hayman method of analysis since they were significant.

Table 5.13: Hayman diallel mean squares for dark colour, reaping and topping for individual and across years

Item	Degrees of freedom	Season one			Season two			Across years		
		Days to 50% reaping	Days to 50% topping	Dark flue-cured leaf colour	Days to 50% reaping	Days to 50 % topping	Dark flue-cured leaf colour	Days to 50% reaping	Days to 50% topping	Dark flue-cured leaf colour
Additive effect (a)	9	64.07*	26.97	616.90**	369.27****	67.00	784.46	136.30	30.82	583.24*
Dominance direction (b1)	1	30.68	9.26	147.85	3.47	0.14	41.80	13.69	0.42	8.12
Allele asymmetry (b2)	9	7.91	5.89	202.82	6.91	6.74	35.85	4.86	4.11	37.69
Dominance peculiarity (b3)	35	11.25	7.71	86.30	12.07*	9.85*	171.85	5.03	5.01	71.38
Dominance (b)	45	11.02	7.38	110.97	10.85	9.02*	141.76	5.19	4.73	63.29
Additive maternal (c)	9	15.1	15.89	66.41	369.87***	33.03	614.41	91.90	7.01	163.87
Maternal interaction (d)	36	18.00***	5.26	74.89	8.37	9.67	97.68	7.69	3.48	43.82
Genotypes (t)	99	18.75	9.16	139.79	75.17	16.71	227.13	25.90	6.86	112.60
Block	1	182.41	24.68	38.06	2.21	7707.85	298.97	29209.45	7849.54	943.78
Block. Additive effects	9	12.73	13.55	79.41	9.73	34.37	580.45	80.37	19.12	117.47
Block. Dominance direction	1	63.85	0.22	185.96	17.80	1.72	129.32	3.38	5.74	86.70
Block. Allele asymmetry	9	10.55	14.08	174.05	7.13	4.16	148.06	2.55	2.38	81.67
Block. Dominance Peculiarity.	35	13.19	9.30	88.57	6.39	5.53	173.83	6.63	3.75	57.70
Block. Dominance	45	13.79	10.05	107.83	6.79	5.18	167.69	5.74	3.52	63.14
Block. Additive maternal effects	9	10.82	8.69	114.36	6.35	20.77	1018.62	100.58	17.29	176.52
Block. Maternal interaction	36	6.18	8.63	71.76	5.00	8.04	148.49	5.50	3.98	42.45
Block. Genotypes	99	10.66	9.73	92.72	6.37	10.29	275.59	21.06	6.36	70.86

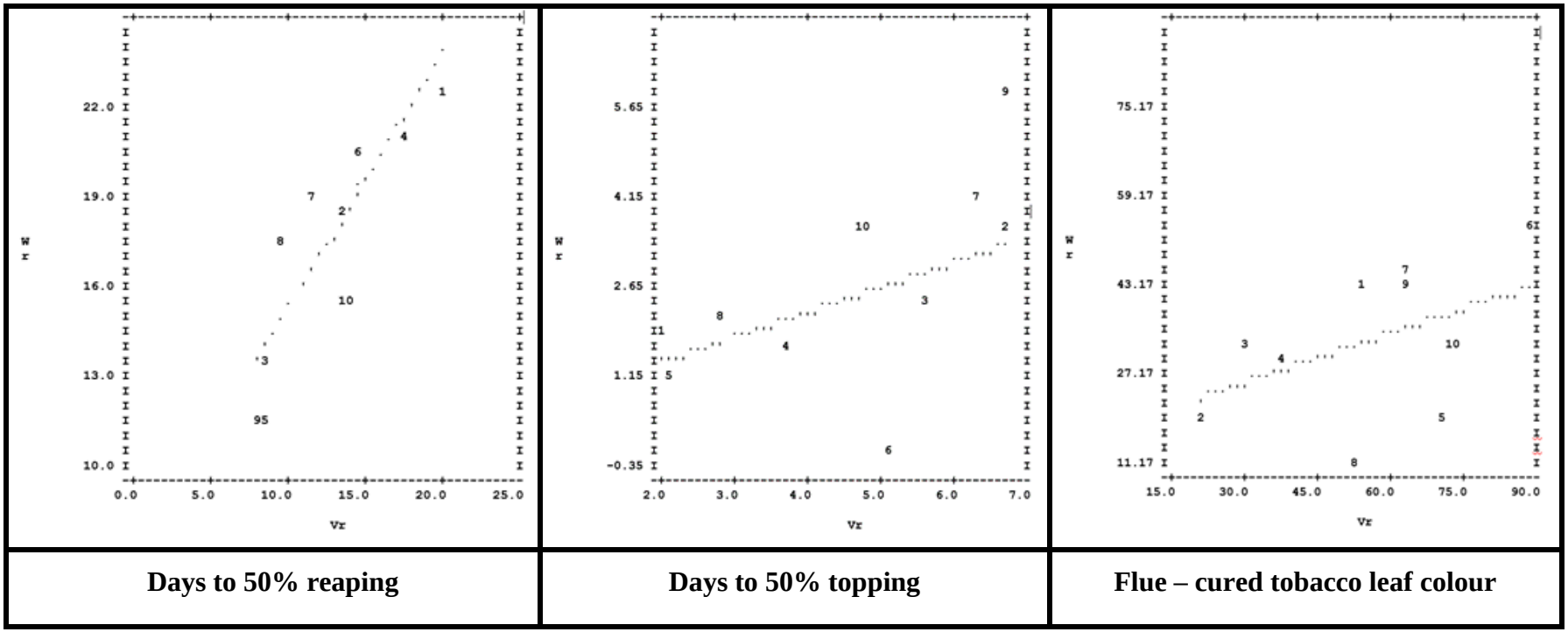
Hayman's diallel analysis shows that additive gene action was significant for days to 50% reaping and was significant ($P<0.05$) for both season 1 and season 2 but not significant for across year analysis (Table 5.13). Additive gene action for dark flue-cured leaf colour was also significant ($P<0.05$) for year 1 and across years. There was some dominance gene action, though, peculiar to some crosses as shown by significant b_3 during season two for days to 50% reaping and days to 50% topping. There were some additive maternal effects for days to 50% reaping during season two only and maternal interaction effects were significant during season one.



Days to 50% reaping	Days to 50% topping	Flue – cured tobacco leaf colour

Figure 5.1: Hayman's graphical display of days to 50% reaping, days to 50% topping and dark flue-cured leaf colour production for the 2015-16 season.

The graphs in Figure 5.1 show that partial dominance gene action on all the traits investigated since the regression line cuts the Wr line above the origin. Parents 6 (XM26R1-1-26, 5 (KM10), 9 (XZR2-2-5) and 1(K326R1-1-28) were closest to the origin the trait days to 50% reaping while parents 2 (RWR3-2-12) are furthest to the origin for the same trait. For days to 50% topping, parent 1 (K326R1-1-28), 3 (ONCR3-4-6) and 10 (KE1) are closest to the origin. Parents 4 (AW3R), 1 (K326R1-1-28), and 7 (XSR4-7-10) are also closest to the origin for the trait dark flue-cured leaf colour.

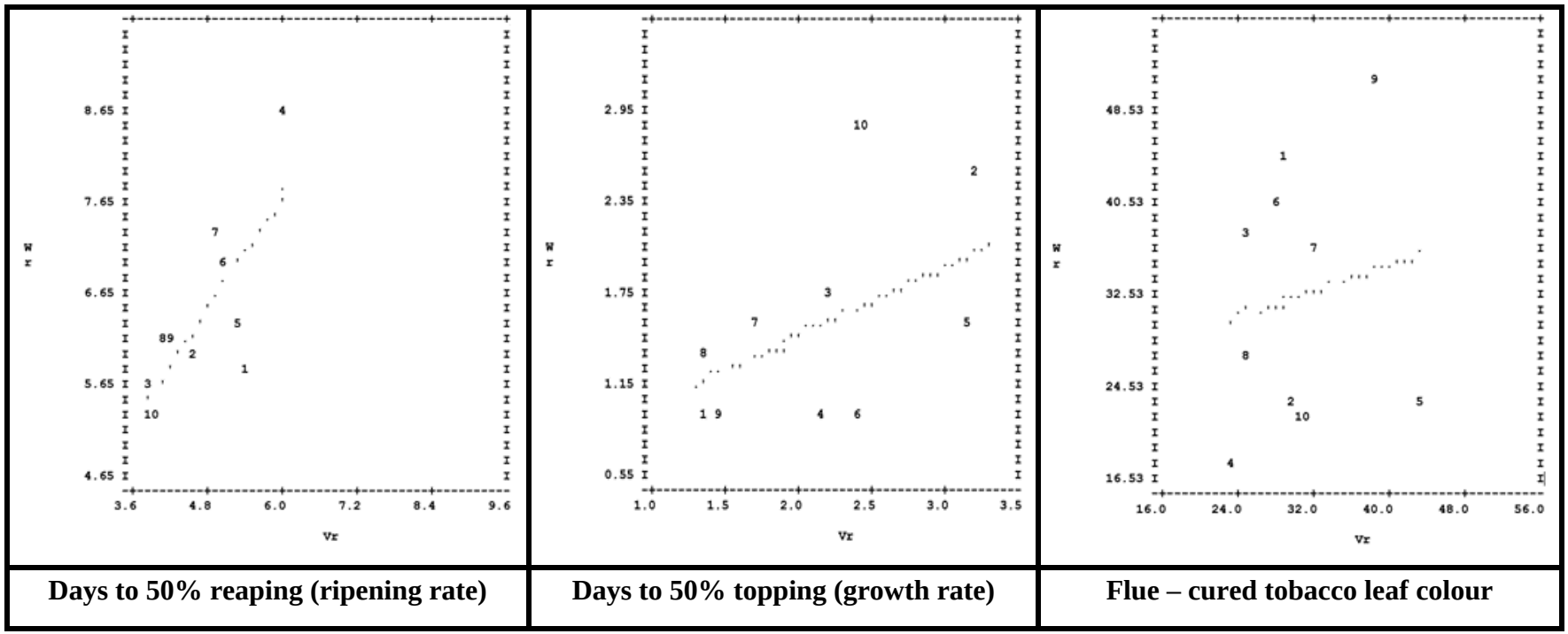


Days to 50% reaping	Days to 50% topping	Flue – cured tobacco leaf colour

Figure 5.2: Hayman's Wr-Vr graphical display of days to 50% reaping, days to 50% topping and dark flue-cured leaf colour production for the 2016-17 season.

The regression line cuts Wr line above the origin, the output showing that partial dominance for all the traits under investigation during the 2016-17 season (Figure 5.3). The graphical display show that parent 2 (RWR3-2-12) is closest to the origin for the trait dark flue-cured leaf colour while parents 6 (XM26R1-1-26) and 5 (KM10) are closest to the origin for days to 50% topping. Also, parents 5 (KM10), 9 (XZR2-2-5) and 3 (ONCR3-4-6) are closest to the origin for the trait days to 50% reaping.

The combined seasons Hayman's graphical gene action display output suggests partial dominance for the traits days to 50% reaping, days to 50% topping and dark flue-cured leaf colour production (Figure 5.3).



Days to 50% reaping (ripening rate)	Days to 50% topping (growth rate)	Flue – cured tobacco leaf colour

Figure 5.3: Combined Hayman's graphical display for days to 50% reaping, days to 50% topping and dark flue-cured leaf colour production across two seasons

Parents 1 (K326R1-1-28) and 9 (XZR2-2-5) have the highest proportion of dominant genes for days to 50% topping while parents 3 (ONCR3-4-6) and 10 (KE1) have the highest number of dominant genes for days to 50% reaping. Parent 4 (AW3R) had the highest number of genes favouring production of dark flue cured tobacco leaves.

5.5 Discussion

5.5.1 Griffing's diallel analysis

Many studies have been conducted on the inheritance of quantitative traits in self-pollinating crops like flue-cured tobacco in the past. In the majority of the cases, it was reported that additive genetic effects were of major importance in quantitative character expression although non-additive effects showed to be of value in other instances (Table 5.3). This means that seasons were different and they influenced the response of the genotypes differently.

GCA effects were significant for flue-cured leaf colour suggesting additive gene action for the trait (Acquaah, 2012). Additive gene action is associated with average effects of genes in crop improvement. The non-significance of the SCA for all the traits means that there is no dominance gene action at play for the traits (Nduwumuremyi *et al*., 2013).

Not much literature on studies of the inheritance of traits such as days to 50% topping (growth rate) and days to 50% reaping (ripening rate) and cured leaf colour in tobacco is available but earlier studies on inheritance of other quantitative traits in flue-cured tobacco were reviewed. For instance, Robinson *et al.* (1954) reported significant estimates of additive genetic variances for growth rate (measured as days to flowering) and for yield in F_2 tobacco hybrids. Marani and Sachs (1966) also found the predominance of additive genetic effects for growth rate in oriental tobacco a type of tobacco also grown for smoking and of diverse origins they studied just like the findings in this study. In addition, Manzinger *et al.* (1963) reported significant GCA effects for ten traits in a diallel study of eight tobacco varieties and also showed that SCA effects were not significant for all the characters studied just like what has been found in this study with traits dark cured leaf colour and ripening rate. All these early findings corroborate with the findings in this study for the trait cured-leaf colour. However, Butorac *et al.* (2004) reported over-dominance inheritance of growth rate (days to flowering) in burley tobacco which is another type of tobacco closely related to flue-cured tobacco which contradicted the findings of Forsberg and Carlborg (2017) who found out that additive effects were significant for growth rate which they measured as number of days to flowering in flue-cured tobacco.

The reciprocal effects were not significant for all the three traits studied meaning that either of the parents can be used as female or male without significant changes in the performance of the hybrids constituted. In constituting tobacco hybrids, it has been difficult to predict the effect of interchanging parents on the overall performance of the constituted hybrid. This result,

therefore, confirms that interchanging the parents of a flue-cured tobacco hybrid does not negatively affect its performance in terms of days to 50% topping (growth rate) and days to 50% reaping (ripening rate) and cured leaf colour obtained**.** Reciprocal effects were not significant for days to 50% reaping suggesting that there could be no changes in the ripening rate of flue-cured tobacco hybrids depending on the parent made the female or male in a cross. Ripening rate is a trait that is of interest to flue-cured tobacco as it influences the quality of the cured crop. Farmers prefer relatively slow ripening varieties that do not give them reaping and curing pressure. This finding concurs with Matzinger *et al.* (1971) who studied the inheritance of alkaloid content, plant height, leaf number and leaf size and reported non-significant reciprocal effects for the traits

Within season, maternal effects were significant (Table 5.3) for cured leaf colour in tobacco. Maternal effects are heritable effects that may influence the behaviour of offspring. They are a result of cytoplasmic differences coming from differences in the gene dosages contributed by the male and female parents from their cytoplasms. Female parents contribute twice the dosage due to the presence of polar nuclei. The female parents always contribute the mitochondria and chloroplasts in a cross. These differences in contributions have been found to influence the performance of certain offspring for specific traits in seed propagated crops (Wolf and Wade, 2009).

The difference in days to 50% reaping among the hybrids shows that crosses differ in the time they take to reach maturity and reaping stage. Traits which fluctuate across seasons point to the possibility of the presence of environmental influences which have no genetic basis. On the combined season analysis, genotypes were significant for dark flue-cured tobacco leaf colour only indicating that the genotypes differed in the amount of dark cured leaf they produce. Inheritance of quantitative traits is a complex process that is greatly influenced by many factors which are genetic or environmental. Quantitative traits are controlled by many genes each of which make a contribution towards the overall phenotype of an organism (Forsberg and Carlborg, 2017). Where the genetic effects other than the genes of interests are at play, pleiotropic effects, linkage and epistasis can mask the expression of genes of interests to different extents. So, the differences in the composition of dark cured leaf colour can be explained as caused by the composition of the genes controlling cured leaf colour at different loci in the genotypes investigated.

Maternal influences for days to 50% topping were only significant during the 2015/16 season but none significant during the 2016/17 and across the seasons while the flue-cured leaf colour is consistently non-significant across seasons suggesting the absence of maternal effects for the trait (Table 5.6 and 5.9). Days to 50% topping is an indicator of the rate of growth in flue-cured tobacco. Tobacco varieties are chosen for growing in specific environments depending on their growth and ripening rate. Varieties that grow fast have been reported to give relatively low yields (Miller, 2016). Significant GCA effects across the seasons for the traits days to 50% reaping and days to 50% topping suggests the presence of additive gene action for the traits. Additive genetic effects are associated with average effects of genes in crop improvement and traits that exhibit additive gene action have been found to be heritable (Wolf and Wade, 2009). The results here therefore confirm that the traits days to 50% reaping and days to 50% topping can be selected for with relatively high chances of success.

5.5.2 Variance components

The high phenotypic variances for traits within and across seasons suggest that the traits could be highly sensitive to environmental factors although the presence of genetic differences among genotypes for the same trait could be another reason. High phenotypic variance has led to reduced heritability for the traits under study with the highest narrow sense heritability across the seasons of 42% recorded for days to 50% reaping (Table 5.4). This is because environmental contribution to the phenotype was relatively high and this could be the reason why heritability and genetic advance were low. The estimate of genetic advance is useful as a selection tool when considered jointly with heritability estimates (Johnson *et al*., 1955).

On the other hand, this could mean that we can make considerable progress in improving tobacco for days to 50% reaping versus other traits under study since it had moderate narrow sense heritability. Narrow-sense heritability is determined by additive gene action which in turn is governed by average effects of genes. Flue-cured tobacco colour has a relatively low narrow sense heritability meaning it can be bred for but with relatively high difficulty. The narrow sense heritability for days to 50% topping is very low at 8% during the 2016 season maybe because of drought and heat stress experienced during that season. It has been reported in earlier studies that genetic variance and broad sense heritability declines with increase in moisture stress because of the complexity of genotype x environment interactions (Rosielle and Hamblin, 1981; Edmeades *et al*., 1997). Estimation of heritability for yield and its components is important in determining the influence of the environment on the expression of

the characters and the extent to which improvement would be successful after a breeding process (Parajuli *et al*., 2015).

Forsberg and Carlborg (2017) found out that growth rate in tobacco was influenced by additive genetic effects and this was in agreement with findings of this study. This could suggest the presence of dominance effects for the trait. The findings were contrary to findings by earlier researchers in quantitative trait inheritance in flue-cured tobacco who found no dominance gene action for traits like growth rate, which they measured as days to flowering. To determine the relative importance of GCA and SCA effects on the traits under investigation, GCA-SCA ratios was calculated. GCA-SCA ratios in this study also indicated that GCA components were relatively more important for the all the traist studied like in the findings of Sayyed-Nazari *et al.* (2016) who also showed that additive gene action was more important for yield in oriental tobacco. The GCA effects were important across the two seasons for dark flue-cured tobacco colour suggesting the presence of additive gene action in the control of the trait in flue- cured tobacco. On the combined analysis, the ratio was highest indicating that the trait days to 50% reaping has relatively high additive gene activity than the other traits.

5.5.3 General combining ability of parental lines

For each trait within the season and across the seasons, parents contributed differently to the overall phenotype of the F_1 hybrids they constituted. Parents with negative general combining ability (GCA) effects reduced the trait mean by the value given while those with a positive GCA effect increased the mean by the given GCA effect. Identifying parents that have the potential to increase or decrease growth and ripening rate is very important in tobacco breeding as it allows breeders to constitute hybrids that suit specific growing conditions within the tobacco growing belt of Zimbabwe. Slow growing and ripening varieties would do well in the hot areas like Hurungwe, Centernary, Chinhoyi and Mt Darwin while the fast growing and ripening ones would be developed targeting cool growing environments in the Highveld areas e.g around Harare, Marondera, Norton and Chegutu. Toning down ripening rate by even a day is quite significant as it reduces harvesting and curing pressure on the farmer and the losses associated with such pressures. Fast growing varieties are generally associated with relatively low cured tobacco yields and poor quality tobacco. Varieties with moderate to slow growth rates are preferred as they tend to give a balance between the yield and the proportion of dark cured styles which are preferred by most buyers on the market.

The study showed that parents like K326R1-1-28, RWR3-2-12, AW3R, XM26R1-1-26, BAZR3-6-18 and XSR4-7-10 reduce the number of days needed for a variety to ripen. They, therefore, quickened the rate of ripening of hybrids possessing them as parents thus reducing the days to 50% reaping (Table 5.5). Conversely, lines ONCR3-4-6, KM10, BAZR3-6-18, XZR2-2-5 and KE1 increased the number of days it takes for a hybrid to get stripped off the field. The lines therefore reduced the ripening rate of the hybrids possessing them. Similarly, KM10 and XZR2-2-5 increased and reduced the growth rate of hybrids respectively. Lines ONCR3-4-6, AW3R, and XM26R2-2-26 increased the proportion of dark cures in the hybrids where they were used as parents while K E1 had the effect of reducing the proportion of dark cures. These observations concur with the general observations by growers of flue-cured tobacco in Zimbabwe.

5.5.4 Hayman diallel analysis

In many cases, Griffings diallel analysis is combined with Hayman analysis to get a more comprehensive understanding of the gene action at play in controlling the inheritance of particular traits in crop breeding. In this study the Wr-Vr Hayman graphical analysis was used to explore the gene action controlling days to 50% topping, days to 50% reaping and dark flue-cured colour in tobacco. Singh and Chaudhary (1979) indicated that estimation of genetic parameters by the cross diallel analysis is possible if there are significant differences among the genotypes in the study based on the F test of the observed variables. Like what was reported by Pandeya *et al.* (1983), the Wr-Vr graph cuts the Wr axis above the zero point indicating partial dominance and lack of non-allelic interaction for all the three traits. Hancock and Lewis (2017) also reported similar findings for growth rate of crosses between *N. tabacum* and synthetic tobacco. Dominance of parents in hybrid combinations for specific traits can also be determined from the relative positions of the lines on the Wr-Vr graph. Parents close to the origin have relatively high proportions of dominant genes for the trait under investigation while those furthest from the origin have relatively high proportions of recessive genes. In this study, parent 10 (KE1) had highest proportion of dominant genes while parent 4 (AW3R) had the highest proportion of recessive genes for days to 50% reaping. These parents can be used in breeding programmes to develop varieties that ripen relatively fast. More so, parent 1 (K326R1-1-28) had the greatest proportion of dominant genes while parent 10 (KE1) had the highest proportion of recessive genes for days to 50% topping. Where the intention is to develop fast growing varieties, the parent 1 (K326R1-1-28) can be of use since it has a relatively high proportion of dominant genes for the trait. For the trait dark flue-cured colour,

AW3R had the highest proportion of dominant genes while XZR2-2-5 had the highest proportion of recessive genes making AW3R a good parent to use when breeding for dark cured leaf flue-cured tobacco varieties.

5.6 Conclusions and recommendations

5.6.1 Conclusions

1. The GCA effects for dark cured leaf colour were significant suggesting that the trait is influenced by additive gene action.
2. SCA effects were not significant suggesting the absence of dominance gene action for all the traits under study.
3. Reciprocal effects for all the traits were not significant meaning that the male or female parents can be interchanged without significant changes in the performance of resultant hybrids
4. There were no maternal effects for dark cured leaf colour, days to 50% topping and days to 50% reaping.
5. Hayman graphical analysis exhibited partial dominance mode of inheritance for dark cured leaf colour, days to 50% topping and days to 50% reaping.
6. The traits dark cured leaf colour, days to 50% topping and days to 50% reaping had generally low to moderate heritability.
7. Genetic advance values were low for all the traits studied across all the seasons.

5.8.2 Recommendations

1. Breeders can include selection for traits like days to 50% topping and days to 50% reaping in tobacco breeding programmes and make selection progress but less benefits can be realized through hybridization.

CHAPTER SIX

Genotype × environment interaction analysis for yield and quality in elite flue-cured tobacco (*Nicotiana tabacum* L.) varieties in Zimbabwe

6.1 Introduction

Even if farms occur in the same province or district, it has been observed that varieties of a crop perform differently across seasons on each farm. The same varieties have been reported to yield differently on different farms during the same season leading farmers in the same growing area preferring to grow different varieties of the same crop (Ding *et al*., 2008; Dolatabad, 2010). Varieties have also been found to change rankings in different growing environments with otherwise popular varieties in one location performing poorly in other locations (Farshadfar *et al*., 2012; Hassanpanah 2011). There are, however, other varieties that perform well across different growing conditions and are considered stable (Karimizadeh *et al*., 2013). Situations where varieties perform differently in different management regimes or edaphic and climatic conditions are a result of genotype × environment interaction. A genotype is defined as a set of genes that make up the characteristics of a specific cultivar while the environment includes all the factors that may influence the performance of a cultivar (Mendes *et al*., 2012; Mohammadi *et al*., 2010). Environment is not limited to climatic factors but also includes field management practices by the farmer in response to prevailing climatic conditions since different soil and climatic conditions determine the farmers' crop management behaviour (Munafò *et al*., 2009; Zobel, 2016).

In general, flue-cured tobacco has always been regarded as non-responsive to changes in environments (Mwololo *et al*., 2009). Although flue-cured tobacco varieties have always been referred to as either good performers across growing environments, modern farmers in different parts of Zimbabwe now prefer some varieties against others either for their quality or yield performance under their specific management regime in the location they farm. The current practice has, therefore, been to develop varieties targeting different farmer growing regions and management styles (TRB Annual Report, 2015). The choice of varieties has always been directed by farmers based on trial and error rather than objective scientific testing. It is important in crop breeding to evaluate varieties across many different environments to determine whether genotype × environment interaction influences their performance since, in Zimbabwe, for flue-cured tobacco, no such work has been done before. Understanding the extent of the genotype × environment interaction in crop production allows farmers to choose

varieties that best suit their growing conditions and management regimes since growing conditions vary from one farm to another (Sabaghnia *et al*., 2012). In addition, such an understanding allows plant breeders to breed varieties targeting specific areas so that the best performance of crops in specific areas is maximised thus resulting in maximum possible yields of a crop in a given environment (Misra *et al*., 2009). The aim of this study was to determine the absence or presence of genotype × environment interaction effects on tobacco yield and quality. If genotype × environment interaction effects were present, the study would determine the best varieties for specific areas as well as those that have broad adaptation and can be grown across the whole country without significant yield and quality losses. Variety characteristics will be linked to performance in different growing areas using climatic conditions and altitude.

6.2 Materials and Methods

6.2.1 Experiment sites

Experiments were set out in Centenary, Rusape, Harare, Makoholi, Karoi North, Matepatepa, Tengwe and Trelawney during the 2015/2016 and the 2016/17 growing seasons. These sites are dotted around the traditional tobacco growing belt of Zimbabwe which is split into the fast, medium and slow growing areas. In general, the areas get average rainfall of not less than 800 mm annually and are characterised by granite sandy soils devoid of fertility. The specific characteristics of each of the locations are given below:

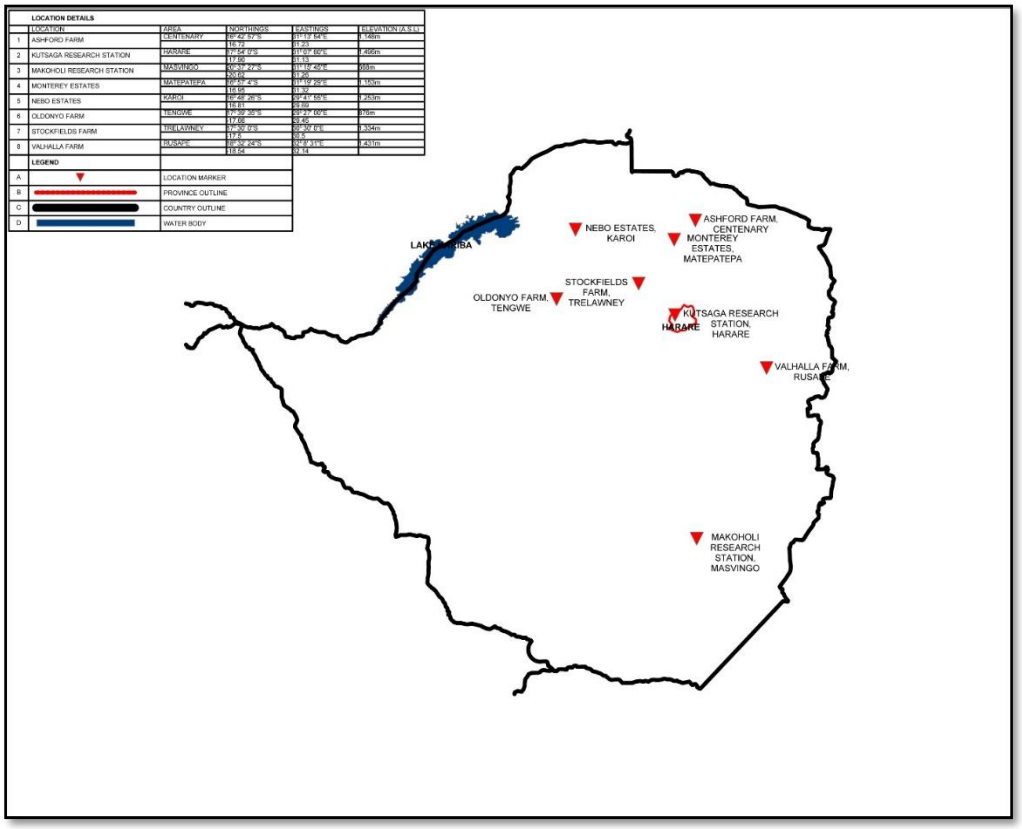


Figure 6.1: Map showing the eight experimental locations for the study

6.2.1.1 Centenary

Centenary is located in the north corner of Mashonaland Central Province of Zimbabwe towards the Zambezi Valley at an altitude of 1156 m above sea level. In Centenary the study was done at Herendon Farm in the 2015/16 season and at Dollar Farm during the 2016/17 farming season. Centenary is regarded as a special tobacco growing area as it produces spotted flue-cured leaf of flavour character. It averages 840.7 mm of rain per year. Temperatures in this area average a high of 28.4°C and a low of 14.8°C. The area is rocky, hilly and sloppy but the soils have medium fertility that favours the production of high value flavour style flue cured tobacco.

6.2.1.2 Harare

In Harare, the experiment was carried out at Kutsaga Research Station. Kutsaga Research Station receives mean annual rainfall ranging 800 mm to 1000 mm per annum. The rainfall usually occurs during a single rainy season from November to April. The mean annual temperature at Kutsaga is 21°C with insignificant frost occurrences in the months of June and July. The research station is located at an elevation of 1480m above sea level. Soils at Kutsaga

are deep well drained granite sands requiring additional fertilization in order to get decent tobacco crops.

6.2.1.3 Karoi North

The Karoi North experiment was set at Nebo Estates near Hurungwe Mountain in the Karoi area of Mashonaland West Province. The area has sandy-loam soils with relatively high proportions of clay making them fertile. The annual average rainfall received in this area is 804.1mm with the month of February being the wettest month receiving 191.6 mm on average. The annual average temperatures are 19°C. The area is known for being very hot in summer and production of deep orange flue cured tobacco of flavour style. Karoi North represents the lower extreme of the tobacco growing belt of Zimbabwe with a low altitude of 1 124 m above sea level with a combination of extremely hot conditions and high rainfall.

6.2.1.4 Makoholi

The Masvingo experiment was located at Makoholi Research Station in Masvingo at an altitude of 1204 m above sea level. Masvingo is a relatively low potential area for agriculture production. It receives low erratic rainfall and the soils are infertile sands. This semi-arid hot climate of Masvingo receives an annual average rainfall of 614.4 mm and experiences annual average temperatures of 19.4°C but the summers can be scorching hot. Makoholi represents the very least potential area for tobacco production since it experiences frequent droughts.

6.2.1.5 Matepatepa

Matepatepa is situated in the middle part of Mashonaland Central Province near the town of Bindura. It is a mountainous area with soils being infertile granite sands needing additional inorganic fertilizers to guarantee better yields. The area is low lying at an altitude of 1070 m above sea level but receives high rainfall of above 800 mm per annum. Temperatures in summer are high, averaging 28 ºC but winters can be freezing. The area is also known for producing high tobacco yields of flavour styles. The experiment was at Pimento Farm in Matepatepa.

6.2.1.6 Rusape, Valhalla Farm

The study was carried out at Valhalla farm in Rusape District, Manicaland Province, in Zimbabwe. Rusape receives an annual average rainfall of 789 mm and annual average temperatures of 17.5°C. The soils at Valhalla Farm are deep well drained, sandy granitic soils with relatively high organic matter and the farm is located at an elevation of 1415 m above sea level. Rusape is at a relatively low altitude hence the environment is relatively hot in late summer. However, the area is classified as a high potential area in tobacco farming terms.

6.2.1.7 Tengwe

Tengwe is a rocky, hilly and gravelly area situated in the southern part of Hurungwe District in Mashonaland West Province of Zimbabwe. The experiment was located at Oldonyo Farm in Tengwe whose soils are rocky but deep and relatively fertile. Tengwe falls at an altitude of 1276m above sea level and is directly opposite Karoi North. Although the area is in the same district of the vast Hurungwe with Karoi North, their climatic and soil conditions are different making them produce different flue cured tobacco types. Tengwe receives high rainfall of above 800mm but it is a very hot environment with average annual temperatures of 29 ºC.

6.2.1.8 Trelawney

In Trelawney, the study was carried out at Stockfield Farm which is situated in 75 km north-west of Harare in Zimbabwe. This area receives an annual average rainfall of 824.9 mm and annual average temperatures of 20.4°C. Soils are deep well drained sands. Trelawney is a traditional tobacco growing area in Zimbabwe. At an altitude of 1274m above sea level, it is regarded as medium in terms of growing conditions in between the cool high altitude areas and the hot low lying areas where tobacco is grown in Zimbabwe.

6.2.2 Germplasm and data collection

Ten new elite varieties (Table 6) and two control varieties (AW3R × KM10 and K RK26R) were selected based on their high yield potential from the newly developed varieties and evaluated in a Randomized Complete Block Design (RCBD) with three replicates for two years at eight sites. Data on leaf yield (dry mass of cured leaf in kg/ha), and leaf quality, dark flue-cured leaf colour and Grade Index). The entries and their basic characteristics are given in Table 6.1.

Table 6.1: List of entries evaluated in the study

Entry No.	Entry Name	Basic Characteristics of the varieties
1	KE1 × K M10	A moderate variety growth and ripening variety resulting from a cross of a slow and a fast line
2	BAZR3-6-18 × K M10	Relatively fast growing and ripening variety with high leaf potential
3	AW3R × KM10 (Control)	A very fast growing and ripening variety with medium Alternaria leaf spot resistance.
4	KE1 × XZR2-2-4	Medium growing and ripening variety.
5	RWR3-2-12 × KM10	Medium growing variety with high leaf potential. It ripens fast.
6	ONCR3-4-6 × K E1	Medium leaf potential variety with thick leaves. It is slow to medium ripening with medium growth rate.
7	K326R1-1-28 × K E1	Long pointed leaves. It exhibits fast growth and ripening rates.
8	KE1 × AW3R	Big broad leaves which are thick with medium growth and ripening rate.
9	K326R1-1-28 × XZR2-2-4	Long pointed leaves giving mainly deep lemon cures. This is a medium growing and ripening variety.
10	AW3R × BAZR3-6-18	Relatively green variety with long pointed leaves. This variety is fast growing and slow ripening.
11	BAZR3-6-18 × XSR4-7-10	Medium green variety with relatively long pointed leaves. It is medium growing and ripening giving dark cures.
12	K RK26R (Control)	A medium yielding variety regarded as possessing wide adaptation. It is fast growing and fast ripening giving lemon cures.

6.2.3 Agronomic trial management

6.2.3.1 Land preparation, seedlings, planting and weed and pest management

For land preparation, raising of seedlings, planting, fertilizer application, weed and pest management refer to **Section 4.2.2.1** to **Section 4.2.2.3**.

6.2.3.2 Leaf priming, topping, sucker removal, reaping, curing and grading

For leaf priming, topping, sucker removal, reaping, curing, and grading refer to **Section 5.2.4** to **Section 5.2.5**.

6.2.3.3 Parameters measured

The parameters measured were ripening growth rate (days to 50% topping), ripening rate (days to 50% reaping), and cured leaf colour (proportion of the cured leaf classified as orange and mahogany per plot), grade index, saleable yield and mass at untying. The parameters are as described below:

6.2.3.3.1 Growth rate

In tobacco farming, "growth rate" refers to the time taken by a variety to develop the standard topping height of 18 leaves from transplanting. In this study "growth rate" refers to the the number of days taken by 50% of plants of a specific variety in a plot to develop to the standard topping height of 18 leaves.

6.2.3.3.2 Ripening rate

In tobacco framing, "ripening rate" is the time, in days, it takes for tobacco leaves to ripen to a harvestable state. The standard ripening rate of a flue-cured tobacco variety in Zimbabwe is two (2) leaves per week. In this study "ripening rate" was measured as the time (days) it took for 50% of plants in a plot to be stripped of leaves from the day of initial harvesting.

6.2.3.3.3 Dark Cured leaf colour

A flue-cured tobacco variety has the potential to produce either lemon, orange or mahogany cured leaves in different proportions. Orange and mahogany cured leaf styles are collectively referred as "Dark style tobacco" while the lemon is referred to as "light style" tobacco. In this study, the harvested and flue cured leaves were graded and classified into lemon, orange and mahogany colours and weighed. Of the total tobacco leaf weight from each plot, the proportion of the "dark style" tobacco (orange and mahogany) was determined and referred to as dark style tobacco in this study.

6.2.3.3.4 Quality

Tobacco quality refers to the overall market acceptability of the cured leaf. Flue-cured tobacco quality is measured over a scale of 1(best quality) to 5 (worst quality). Scores of 1-3 are acceptable on the market while scores of 4-5 are unacceptable and fetch relatively low prices. Quality in tobacco farming is a function handling which is lined to the genetic composition of

a variety. In this study, the cured leaf from each plot was graded and scored into the five grades of quality and the weight of tobacco leaf in each grade determined. Weights of tobacco in the scores 1-3 were lumped together and their proportion relative to the total yield per plot determined and referred to as "Quality" throughout this study.

6.2.3.3.5 Grade Index

An index that ranges from 1 to 100 that is used to evaluate the overall acceptability of cured tobacco leaf. Grade index takes into account all the cured tobacco leaf factors like quality grade, cured leaf colour and general leaf acceptability.

6.2.3.3.6 Yield

In tobacco, yield refers to the mass of the cured leaf per unit area (usually reported kg/ha). When the cured leaf fresh from curing is weighed, it gives the first component of yield called Mass at Untying. After determining mass at untying, the cured leaf is graded to give another component of yield called Saleable yield. Saleable yield is the yield component that can the taken to the market. In this study both saleable yield and mass at untying are reported.

6.3 Data analysis

Analysis of variance was conducted in Genstat Version 13 to determine the effect of environment (E), genotype (G) and genotype x environment interaction. Data was analysed using the AMMI model (Yan and Hunt, 2001). AMMI analysis first fits additive effects for genotypes (G) and environments (E) by the usual additive analysis of variance procedure, and then fits multiplicative effects for genotype x environment interaction by principal component analysis (PCA).

The model used was:

$$\mathbf{Y_{ij} = \mu + g_i + e_j + \sum_{k=1}^{r} \lambda_k \alpha_{ik} \gamma_{jk} + R_{ij}} \quad \text{where,}$$

$\mathbf{Y_{ij}}$ is the yield of the i^{th} genotype in the j^{th} environment.

$\boldsymbol{\mu}$ is the grand mean

$\mathbf{g_i}$ is the is the mean of the i^{th} genotype minus the grand mean.

$\mathbf{e_j}$ is the is the mean of the j^{th} environment minus the grand mean.

$\boldsymbol{\lambda_k}$ is the square root of the eigenvalue of the PCA axis k.

$\boldsymbol{\alpha_{ik}}$ and $\boldsymbol{\gamma_{jk}}$ are the principal component scores for PCA axis k of the i^{th} genotype and the j^{th} environment respectively.

$\mathbf{R_{ij}}$ is the residual.

The first two components which resulted from model diagnostic of the AMMI analysis were used to obtain GGE biplots in Genstat Version 13. The basic model for a GGE biplot is:

$$Y_{ij} - \mu - \beta_j = \sum_{i=1}^{k} \lambda_l \xi_{il} \eta_{lj} + \varepsilon_{ij}$$

Where,

Y_{ij} = the mean yield of genotype i (=1, 2, …, n) in environment j (=1, 2, …m),

μ = the grand mean,

β_j = the main effect of environment j, ($\mu + \beta_j$) being the mean yield of environment j,

λ_l = the singular value (SV) of l^{th} principal component (PC), the square of which is the sum of squares explained by PC_l = (l = 1, 2, …, k with k ≤ min (m, n) and k = 2 for a two- dimensional biplot),

ξ_{il} = the eigenvector of genotype i for PC_l,

η_{lj} = the eigenvector of environment j for PC_l,

ε_{ij} = the residual associated with genotype i in environment j.

Analysis of the data was carried per season and across seasons. Environmentes were made up of season – site combinations hence the the which won where analysis was not carried out.

Genotypic (GCV%) and environmental coefficient (ECV%) of variation was calculated as percentage using the formulae below:

$$GCV\ \% = \sqrt{\frac{\mathrm{V}g}{\bar{\mathrm{x}}}} \times 100$$

$$ECV\ \% = \sqrt{\frac{Ve}{\bar{x}}} \times 100$$

Where, GCV% = Genotypic Coefficient of variation; Vg = Genotypic Variance; ECV % = Environmental Coefficient of variation; Ve = Environmental Variance.

Broad Sense Heritability (H^2) was calculated using the formula: $H^2 = \frac{Vp}{Ve}$

Where, Vp = phenotypic variance and Ve = environmental variance. Heritability estimate was categorized as low when less than 40%, moderate when 40-59%, moderately high when 60-79% and very high when above 80% as suggested by Singh (2001).

GA values were determined according to formular illustrated by Jonson *et al*. (1955) where:

$$GA = K\delta_p \text{ x } H^2$$

Where GA is genetic advance, K = 2.063 (selection differential at 5%), δ_p = phenotypic standard deviation of mean performance of hybrid, H^2 = broad sense heritability, × = multiplied by.

6.4 Results

6.4.1 Seasonal analysis of variance

During the 2015/16 season, there were highly significant differences ($P < 0.001$) among locations for all traits (Table 6.2). Varieties showed significant differences ($P < 0.01$) for cured leaf colour, quality grade, saleable yield, and mass at untying but no significant differences were observed for grade index. Genotype × environment interaction was not significant for all the traits except for quality grade. Heritability for all the traits was high, with cured leaf colour recording the least estimate of 61% and mass at untying recording the highest value of 80.5%.

Table 6.2: Mean squares for tobacco traits evaluated in eight locations during the 2015/16 season

Source	Degrees of freedom	Grade index	Cured leaf colour	Quality grade	Saleable yield	Mass at untying
Locations	7	648.07***	4131.20***	1517.23***	21607782.00***	28394549.00***
Locations. Block	16	90.81***	160.80**	240.89***	966765.00***	1069115.00***
Variety	11	172.24ns	377.40**	388.49***	799651.00***	1167920.00***
Locations × Variety	77	35.26ns	145.30ns	94.84*	168826.00ns	227853.00ns
Residual	176	20.41	109.40	63.99	250376.00	167988.00
Total	287	49.44	230.20	130.01	810400.00	961064.00
Genotypic variance		5.71	9.67	12.24	26284.38	39169.46
G x E variance		4.95	11.97	10.28**	0.00	19955.00
Error variance		20.41	109.40	63.99	250376	167988.00
Means		59.38	49.60	84.95	2399.42	2206.67
GCV %		4.02	6.27	4.12	6.76	8.97
ECV %		7.61	21.09	9.42	20.85	18.57
Broad sense heritability %		79.53	61.50	75.59	71.59	80.49

GxE - genotype by environment interaction, GCV-genotypic coefficient of variation, ECV-error coefficient of variation

Table 6.3 shows that there were significant differences ($P < 0.001$) among varieties for all the traits measured during the 2015/16 season. The cross KE1 × AW3R performed just like the check variety K RK26R gave the highest overall grade index value of 63.17 but the cross had the lowest proportion of dark cures, which are mostly preferred on the market. This cross also had the highest saleable yield (2440.4 kg/ha) and overall mass at untying (2834.33 kg/ha). The cross BAZR3-6-18 × K M10 gave cured tobacco with the highest proportion of dark cures of 55.87%.

Table 6.3: Mean performance of the tobacco varieties during the 2015/16 season

Variety	Grade index	Dark cured leaf colour (%)	Quality grade	Mass at untying (kg/ha)	Saleable yield (kg/ha)
KE1 × K M10	55.25^{a}	50.10^{bcde}	80.33^{a}	2095.68^{a}	1894.66^{ab}
BAZR3-6-18 × K M10	56.39^{a}	55.87^{e}	81.83^{abc}	2212.88^{abc}	1991.8^{abcd}
AW3R × KM10 (Control)	56.78^{ab}	47.45^{abcd}	78.84^{a}	2274.03^{abcde}	2027.8^{a}
KE1 × XZR2-2-4	59.53^{cd}	46.43^{abc}	87.13^{defg}	2560.18^{ef}	2323.69^{ef}
RWR3-2-12 × KM10	58.72^{bc}	54.90^{de}	82.25^{abcd}	2217.44^{ab}	1976.78^{abc}
ONCR3-4-6 × K E1	59.48^{cd}	50.16^{bcde}	88.88^{efg}	2247.84^{abcd}	2048.1^{bcde}
K326R1-1-28 × KE1	57.57^{ab}	51.70^{cde}	81.30^{ab}	2323.8^{bcde}	2142.9^{cdef}
KE1 × AW3R	63.17^{ef}	41.80^{a}	90.52^{fg}	2834.33^{f}	2440.4^{g}
K326R1-1-28 × XZR2-2-4	60.00^{cd}	45.73^{ab}	86.00^{cdef}	2407.65^{cdef}	2181.78^{def}
AW3R × BAZR3-6-18	60.42^{d}	49.77^{bcd}	86.37^{cdefg}	2396.86^{bcde}	2178.85^{cdef}
BAZR3-6-18 × XSR4-7-10	60.87^{de}	52.62^{cde}	85.13^{bcde}	2417.52^{cdef}	2163.66^{cdef}
K RK26R (Control)	64.32^{f}	48.62^{bcd}	90.83^{g}	2449.53^{def}	2252.97^{f}
Grand mean	59.38	49.60	84.95	2369.81	2135.28
5% LSD	2.57	5.96	4.56	206.78	195.08
P value	<.001	<.001	<.001	<.001	<.001
CV (%)	7.61	21.09	9.42	12.75	15.06

LSD-least significant difference, CV (%)-coefficient of variation. Means with the same letter are not significantly different.

Locations were highly significant ($P < 0.001$) for all the traits studied in 2016/17 (Table 6.4). Significant differences were obtained for varieties for the traits grade index, quality grade, saleable yield and mass at untying. There was no significant ($P > 0.05$) genotype × environment interaction for all the traits during this season except quality grade. Heritability values for traits measured during the 2016/17 season ranged from low to high with the least value being 32% for quality grade and the highest value being 77.5% for grade index.

Table 6.4: Mean squares for tobacco traits evaluated in seven locations during the 2016/17 season

Source	Degrees of freedom	Grade index	Cured leaf colour	Quality grade	Saleable yield	Mass at untying
Locations	6	548.35***	11327.20***	6184.20***	8783405.00***	5574942.00***
Locations. Block	14	39.56*	131.10ns	185.40ns	540588.00***	433125.00*
Variety	11	119.98***	381.90ns	366.10***	610657.00**	809793.00***
Locations×Variety	66	27.03ns	169.10ns	214.60***	260627.00ns	292393.00ns
Residual	154	20.09	212.70	108.90	244633.00	259888.00
Total	251	40.00	469.80	297.50	485501.00	429250.00
Genotypic variance		4.43	10.13	7.21	16668.10	24638.10
GxE variance		2.31	0.00	35.23	5331.33	10835.00
Error variance		20.09	212.70	214.60	244633.00	259888.00
Means		56.96	37.40	78.89	2580.92	2411.08
GCV %		3.69	8.51	3.40	5.00	6.51
ECV %		7.87	38.99	18.57	19.16	21.14
Broad sense heritability %		77.47	50.01	32.11	57.32	63.89

GxE-genotype by environment, GCV-genotypic coefficient of variation, ECV-error coefficient of variation

Table 6.5 shows that there were significant differences ($P < 0.01$) among varieties for all the traits investigated during the 2016/17 season except dark cured leaf colour. None of the test varieties gave higher values than the control variety K RK26R in terms of grade index. Crosses with either BAZR3-6-18 or KE1 tended to give cured leaf tobacco with relatively high proportions of dark cured leaf styles although the cross RWR3-2-12 × KM10 was an exception. The K326R1-1-28 × XZR2-2-4 gave the highest quality grade and the cross KE1 × KM10 gave least quality grades although the quality levels were all very high. K326R1-1-28 × XZR2-2-4 gave the highest saleable yield and mass at untying during the 2016/17 season.

Table 6.5: Mean performance of the tobacco varieties during the 2016/17 season

Variety	Grade index	Cured leaf colour	Quality grade	Mass at untying	Saleable yield
KE1 × K M10	52.92^{a}	32.45^{ab}	73.87^{a}	2210.82^{a}	2267.53^{a}
BAZR3-6-18 × K M10	54.89^{ab}	42.16^{cd}	76.69^{abc}	2499.25^{abcd}	2353.54^{b}
AW3R × KM10	58.49^{cd}	33.03^{ab}	83.16^{d}	2677.07^{de}	2445.11^{bc}
KE1 × XZR2-2-4	55.20^{ab}	39.74^{abcd}	74.00^{a}	2567.86^{bcd}	2407.01^{bc}
RWR3-2-12 × KM10	55.60^{ab}	44.63^{d}	73.37^{a}	2513.52^{abc}	2347.02^{b}
ONCR3-4-6 × K E1	55.70^{b}	40.24^{bcd}	74.87^{ab}	2504.99^{abcd}	2360.41^{b}
K326R1-1-28 × KE1	56.65^{bc}	34.81^{abc}	82.73^{cd}	2428.77^{abcd}	2351.89^{ab}
KE1 × AW3R	57.31^{bc}	31.12^{a}	80.98^{bcd}	2606.09^{bcd}	2432.9^{bc}
K326R1-1-28 × XZR2-2-4	61.10^{d}	37.42^{abcd}	83.70^{d}	2806.82^{e}	2625.77^{c}
AW3R × BAZR3-6-18	57.36^{bc}	38.1^{abcd}	78.06^{abcd}	2750.46^{cde}	2433.62^{bc}
BAZR3-6-18 × XSR4-7-10	57.37^{bc}	40.81^{bcd}	83.59^{d}	2508.58^{ab}	2311.6^{b}
K RK26R (Control)	60.92^{d}	34.3^{abc}	81.69^{cd}	2672.27^{de}	2450.45^{bc}
Grand mean	56.96	37.40	78.89	2562.21	2398.9
5% LSD	2.73	8.89	6.36	223.66	185.35
P-value	<.001	0.059	<.001	0.003	0.007
CV (%)	7.87	38.99	13.23	20.53	19.05

LSD-least significant difference, CV (%)-coefficient of variation. Means with the same letter are not significantly different.

Genotype x environment interaction was significant ($P < 0.001$) for quality grade hence an AMMI ANOVA was carried out to explore the interaction effects (Table 6.6). The AMMI analysis of variance output showed highly significant differences ($P<0.001$) among tobacco varieties and environments studied during the 2015/16 farming season. Genotype x environment interaction mean squares were significant ($P<0.05$) and only IPCA 1 significantly contributed to the genotype x environment interaction variation (contributed 68% of interaction sum of squares).

Table 6.6: AMMI ANOVA for quality grade measured during the season 2015/2016

Source of variation	Degrees of Freedom	Sum of squares	Mean of squares	Fisher's value	Fisher's Probability
Treatments	95	22197	233.70	3.65	0.00000
Genotypes	11	4273	388.50	6.07	0.00000
Environments	7	10621	1517.20	6.30	0.00000
Block	16	3854	240.90	3.76	0.00001
Interactions	77	7303	94.80	1.48	0.01767
Interactive principal component axis 1	17	4963	292.00	4.56	0.00000
Interactive principal component axis 2	15	1124	74.90	1.17	0.29831
Residuals	45	1216	27	0.42	0.99948
Error	176	11263	64		
Total	287	37314	130.00		

IPCA-interactive principal component axis

The polygon view of the GGE biplot (which-won-where) for quality grade (which had significant genotype x environment interaction during the 2015/16 season) was constructed using the first two principal components (IPCA1 and IPCA2) derived from subjecting the environment-centered data to singular-value decomposition (Figure 6.2). The biplot puts the test locations into two mega-environments. Centenary and Makoholi were grouped together into one mega-environment while the rest fall under a second mega-environment cluster. The best variety for the mega-environment incorporating the locations; Trelawney, Matepatepa, Kutsaga, Rusape, Nebo and Tengwe is K RK26R. Variety 6 (ONCR3-4-6 × K E1) was the best variety for the mega-environment incorporating the locations Makoholi and Centenary.

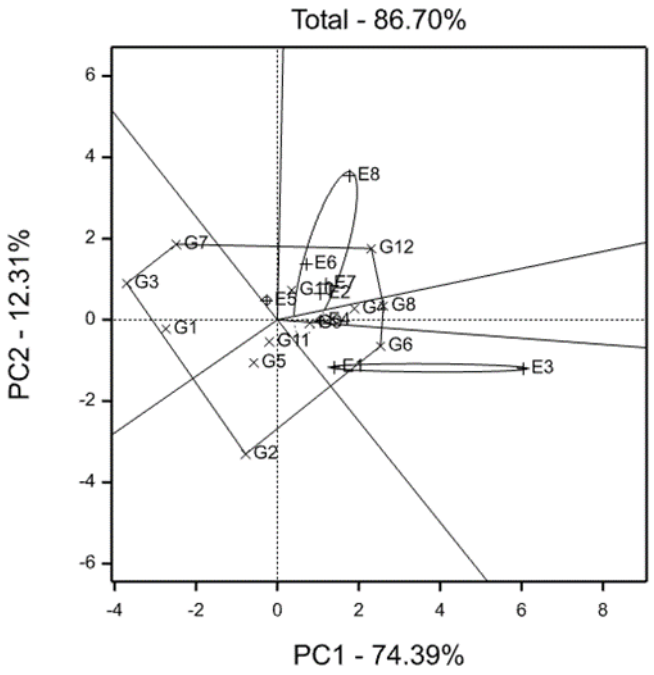


Figure 6.2: Which-won-where illustration for quality grade in the season 2015/16

The comparison biplot for the quality grade trait on Figure 6.3 show that genotypes 4 (KE1 × XZR2-2-4), 8 (K E1 × AW3R) 6 (ONCR3-4-6 × K E1) and 9 (K326R3-6-18 × XZR2-2-4) are in the innermost concentric rings of the graph together with the control variety K RK26R. Almost half of the genotypes had above average quality grade.

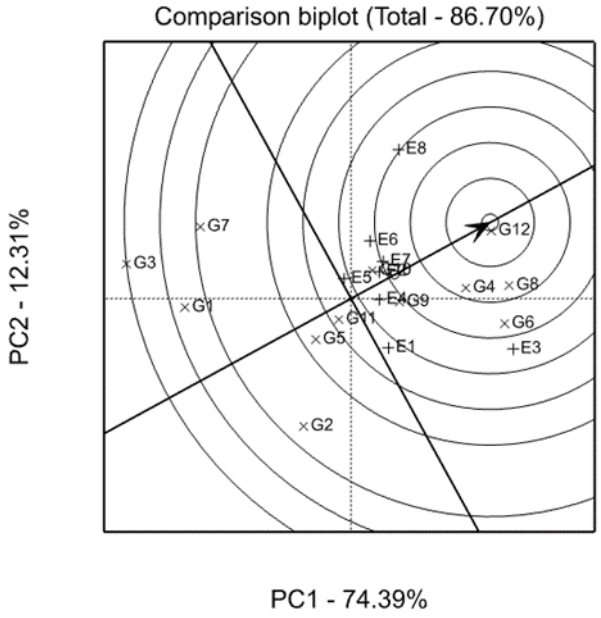


Figure 6.3: Comparison biplot for the trait quality grade during the season 2015/16

Table 6.7 below gives the best four varieties for quality grade measured across all the environments during the 2015/16 season. The general trend was that varieties 12 (K RK26R), 8 (K E1 × AW3R) and 10 (AW3R × BAZR3-8-18) were among the best four in performance at all the test locations. Variety 12 (K RK26R) was the best performer at five of the eight locations while genotype 8 (K E1 × AW3R) was the best performer at two of the locations.

Table 6.7: The best four environment selections for the 2015/16 season

Number	Environment	Mean	Score	1	2	3	4
1	Centenary	84.31	0.08	G8	G2	G6	G11
2	Kutsaga	88.72	0.75	G12	G8	G6	G10
3	Makoholi	77.03	-5.72	G6	G8	G4	G12
4	Matepatepa	91.48	0.54	G8	G12	G6	G10
5	Nebo	89.56	2.26	G12	G11	G1	G3
6	Rusape	91.01	1.24	G12	G8	G10	G6
7	Tengwe	83.23	0.61	G12	G8	G6	G10
8	Trelawney	74.27	0.24	G12	G4	G8	G10

G = genotype ie G8 means genotye 8.

Table 6.8 below shows, the AMMI ANOVA done for quality grade measured during the 2016/17 growing season. Tobacco varieties and test environments were highly significant (P<0.001) for quality grade. There were highly significant genotype × environment interactions (P < 0.001) for the trait quality grade and the output shows that only IPCA 1 is sufficient to describe the interactions present.

Table 6.8: AMMI ANOVA for quality grade measured during the 2016/2017 season

Source	Degrees of Freedom	Sum of squares	Mean of squares	Fisher's value	Fisher's Probability
Treatments	83	55299	666.30	6.12	0.00000
Genotypes	11	4027	366.10	3.36	0.00035
Environments	6	37105	6184.20	33.36	0.00000
Block	14	2596	185.40	1.70	0.06022
Interactions	66	14166	214.60	1.97	0.00033
Interactive principal component axis 1	16	10123	632.70	5.81	0.00000
Interactive principal component axis 2	14	1920	137.10	1.26	0.23924
Residuals	36	2123	59.00	0.54	0.98355
Error	154	16774	108.90		
Total	251	74669	297.50		

The which-won-where GGE biplot for quality grade delineated two mega-environments: Rusape and Trelawney. These two mega environments however, merged on the sites Kutsaga, Matepatepa, Centenary and Makoholi (Figure 6.3). Tengwe stands out alone as a unique environment. Varieties 12 (K RK26R) and 3 (AW3R × KM10) were the outstanding performers in the two interlocking mega-environments. Varieties 9 (K326R3-6-18 × XZR2-2-4) and 7 (K326R3-6-18 × K E1) were the best varieties for quality grade in Tengwe.

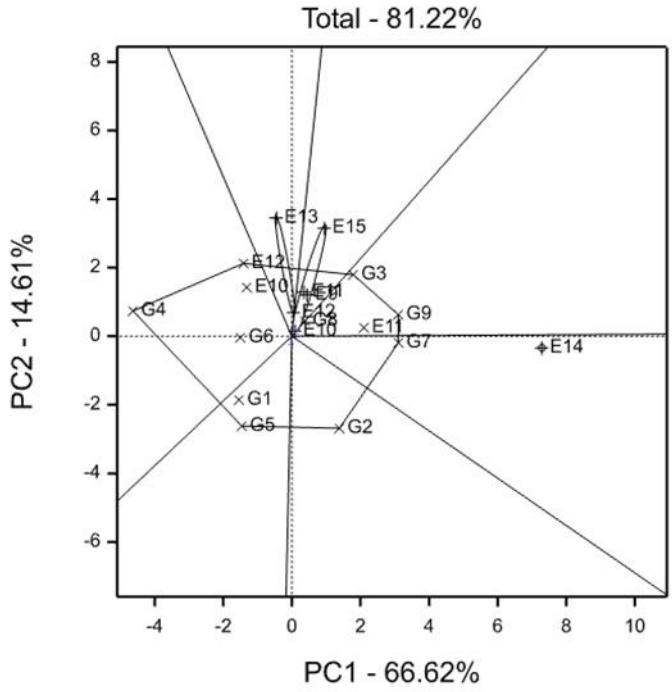


Figure 6.4: Which-won-where biplot for the trait quality grade across sites during the 2016/17 cropping season

Figure 6.5 shows that genotypes 3 (AW3R × KM10) is located in the innermost part of the concentric circles in the comparison biplot during the 2016/17 season. Genotypes 11 (BAZR3-6-18 × XSR4-7-10), 9 (K326R3-6-18 × XZR2-2-4) and 8 (K E1 × AW3R) follow as the pattern moves from the centre of the of biplot. All these genotypes which were located close to the centre of the the concentric circles had above average quality grades.

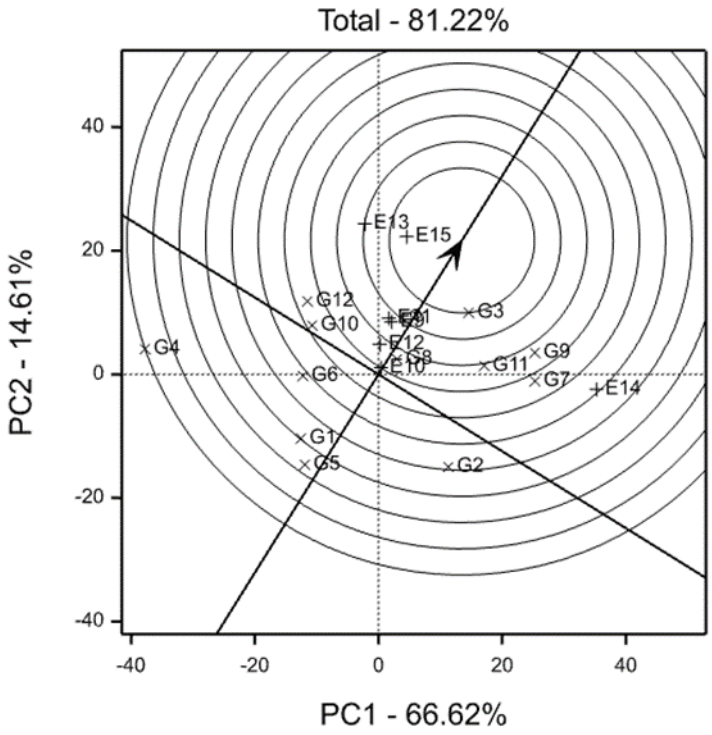


Figure 6.5: Comparison biplot for quality grade for season 2016/17

During the 2017 farming season, varieties G12 (K RK26R), G8 (K E1 × AW3R), G11 (BAZR3-6-18 × XSR4-7-10) and G9 (K326R3-6-18 × XZR2-2-4) dominated in performance across the test locations (Table 6.9). Variety 11 (BAZR3-6-18 × XSR4-7-10) showed great promise although it was not the highest performer for quality grade. However, the which-won-where pattern is not repeatable across the sesons tested.

Table 6.9: The best four selections in each environment for the 2016/2017 season

Number	Environment	Mean	Performance Score	1	2	3	4
1	Centenary	77.51	0.89	G12	G3	G9	G11
2	Kutsaga	87.04	1.20	G7	G12	G8	G11
3	Makoholi	91.90	0.97	G12	G11	G8	G7
4	Matepatepa	92.59	1.22	G12	G11	G8	G3
5	Rusape	75.37	2.20	G12	G4	G8	G11
6	Tengwe	55.41	-6.96	G7	G9	G11	G3
7	Trelawney	72.42	0.48	G3	G9	G6	G12

6.4.2 Across site ANOVA for the traits grade index, cured leaf colour, quality grade, saleable yield and mass at untying

Environments and varieties were highly significant (P<0.001) for all the traits measured (Table 6.10). Genotype × environment interaction was, however, only significant for the traits grade index, and quality grade. Heritability for all traits measured was moderate and it ranged from 58.91% for quality grade to 86.14% for grade index.

Table 6.10: Combined ANOVA across sites for grade index, dark cured leaf colour, quality grade, saleable yield and mass at untying

Source	Degrees of freedom	Grade index	Cured leaf colour	Quality grade	Saleable yield	Mass at untying
Environment	14	615.03***	8347.90***	3761.42***	14884761.00***	16987852.00***
Environment. Block	30	66.89***	146.90ns	215.00***	767883.00***	772320.00***
Variety	11	238.44***	636.40***	400.79***	1122814.00***	1366762.00***
Environment × Variety	154	33.05***	153.90ns	164.68***	216645.00ns	282877.00ns
Residual	330	20.26	157.60	84.96	247696.00	210874.00
Total	539	46.41	378.40	216.91	665820.00	722051.00
Genotypic variance		4.56	10.72	5.25	20046.36	31974.55
GxE variance		4.26**	0.00	26.57**	27637.89	19120.49
Error variance		20.26	157.60	84.96	156171.81	186034.22
Means		116.57	1109.80	551.17	2267.09	2466.01
GCV %		1.83	0.30	0.42	2.97	3.60
ECV %		3.86	1.13	1.67	12.33	11.57
Heritability (H^2)		86.14	75.38	58.91	75.00	83.00

GxE-genotype × environment interaction, GCV-genotypic coefficient of variation, ECV-error coefficient of variation

Table 6.11 shows that there were significant differences among varieties for all the traits measured across seasons and test locations. The control variety K RK26R had the highest across site grade index followed by the cross BAZR3-6-18 × XSR4-7-10. None of the experimental crosses could surpass the quality grade of the control variety K RK26R across the seasons and locations. The KE1 × K M10 had the least grade index of 54.16. RWR3-2-12 × KM10 had the highest proportion of dark cured leaf colour compared to the rest of the evaluated varieties. The cross KE1 × AW3R had the highest saleable yield and mass at untying although it was not significantly different from the control variety K RK26R.

Table 6.11: Mean performance of tobacco varieties evaluated across sites for all traits measured

Variety	Grade index	Dark cured leaf colour	Quality grade	Saleable yield	Mass at untying
KE1 × K M10	54.16^{a}	41.86^{abc}	77.32^{a}	2053.1^{a}	2122.57^{a}
BAZR3-6-18 × K M10	55.69^{ab}	49.47^{de}	79.43^{ab}	2163.58a^{bc}	2350.87^{b}
AW3R × KM10 (Control)	57.58^{cd}	40.72^{ab}	80.85^{abc}	2248.15^{bcd}	2486.11^{bcdef}
KE1 × XZR2-2-4	57.51^{bcd}	43.30^{bc}	81.00ab^{c}	2365.4^{de}	2563.98^{def}
RWR3-2-12 × KM10	57.27^{bc}	50.11^{e}	78.10^{a}	2151.46^{ab}	2361.56^{b}
ONCR3-4-6 × K E1	57.72^{cd}	45.53^{bcde}	82.34^{bcde}	2196.21^{bcd}	2371.64^{bc}
K326R1-1-28 × KE1	57.14^{bc}	43.81^{bcd}	81.97^{bcd}	2236.47^{bcd}	2364.47^{bcd}
KE1 × AW3R	60.43^{e}	36.81^{a}	86.07^{ef}	2441.56^{f}	2722.81^{f}
K326R1-1-28 × XZR2-2-4	60.51^{e}	41.85^{abc}	84.93^{def}	2455.68e^{f}	2629.03^{ef}
AW3R × BAZR3-6-18	58.99^{cde}	44.32^{bcde}	82.49^{bcde}	2313.82^{cde}	2590.44^{bcdef}
BAZR3-6-18 × XSR4-7-10	59.24^{de}	47.11^{cde}	84.41^{cdef}	2217.21^{bcd}	2458.11^{bcde}
K RK26R (Control)	62.74^{f}	41.94^{abc}	86.56^{f}	2362.47^{ef}	2570.53^{cdef}
Grand mean	58.25	43.90	82.12	2267.09	2466.01
5% LSD	1.87	5.82	3.82	155.3	172.03
P-value	<.001	<.001	<.001	<.001	<.001
CV (%)	7.73	28.59	11.22	17.43	17.49

5% LSD-least significant difference. Means with the same letter are not significantly different.

Table 6.12 shows that the genotypes tested and the test environments were significantly different ($P < 0.001$). There was significant ($P < 0.05$) genotype × environment interaction for the traits grade index, quality grade, and mass at untying. Genotype x environment interaction mean squares were highly significant ($P<0.001$). Both IPCA1 and IPCA2 were highly significant ($P<0.01$) and contributed 45.34% and 17.60% respectively of the genotype x environment interaction sum of squares.

Table 6.12: AMMI Anova for grade index measured across all environments.

Source	Degrees of Freedom	Sum of squares	Mean of squares	Fisher's value	Fisher's Probability
Treatments	179	16323.00	91.20	4.50	0.00000
Genotypes	11	2623.00	238.40	11.77	0.00000
Environments	14	8610.00	615.00	9.19	0.00000
Block	30	2007.00	66.90	3.30	0.00000
Interactions	154	5090.00	33.10	1.63	0.00013
Interactive principal component axis 1	24	2308.00	96.20	4.75	0.00000
Interactive principal component axis 2	22	896.00	40.70	2.01	0.00506
Residuals	108	1886.00	17.50	0.86	0.81794
Error	330	6685.00	20.3		
Total	539	25015.00	46.40		

Table 6.13 shows that there were significant differences ($P < 0.05$) among genotypes for the trait quality grade across the locations. Environments also showed significant differences across the seasons ($P < 0.001$). There were significant genotype × environment interactions ($P < 0.001$) for the same trait across the two seasons and the 15 locations (One location was lost during the 2016/17 season because the farm manager got crippled by an accident that he could not look after the trial). The output suggests that only two IPCA is sufficient to fully describe variation due to genotype x environment interaction.

Table 6.13: AMMI Anova for quality grade evaluated across sites

Source	Degrees of Freedom	Sum of squares	Mean of squares	Fisher's value	Fisher's Probability
Treatments	179	82430.00	460.50	5.42	0.00000
Genotypes	11	4409.00	400.80	4.72	0.00000
Environments	14	52660.00	3761.40	17.50	0.00000
Block	30	6450.00	215.00	2.53	0.00003
Interactions	154	25361.00	164.70	1.94	0.00000
Interactive principal component axis 1	24	15186.00	632.70	7.45	0.00000
Interactive principal component axis 2	22	4417.00	200.80	2.36	0.00064
Residuals	108	5758.00	53.30	0.63	0.99761
Error	330	28037.00	85		
Total	539	116917.00	216.90		

For the trait grade index, the variety G12 (K RK26R) still performed best across the locations and seasons (Table 6.14). Variety G8 (KE1 × AW3R) maintained the second position followed by variety G10 (AW3R × BAZR3-6-18) for grade index.

Table 6.14: Best four variety selections for grade index measured across seasons

Number	Environment	Mean	Score	1	2	3	4
1	**Centenary 2015/16**	56.09	-0.47	G12	G9	G8	G3
2	**Kutsaga 2015/16**	59.99	0.75	G12	G8	G10	G9
3	**Makoholi 2015/16**	53.98	4.13	G8	G12	G4	G6
4	**Matepatepa 2015/16**	63.56	1.12	G12	G8	G10	G9
5	**Nebo 2015/16**	60.69	-1.49	G3	G12	G9	G7
6	**Rusape 2015/16**	63.42	0.11	G12	G8	G10	G9
7	**Tengwe 2015/16**	63.59	-0.13	G12	G8	G9	G10
8	**Trelawney 2015/16**	53.68	0.58	G12	G8	G10	G9
9	**Centenary 2016/17**	52.74	-1.09	G9	G12	G3	G11
10	**Kutsaga 2016/17**	56.51	0.06	G12	G8	G9	G10
11	**Makoholi 2016/17**	59.42	-0.85	G12	G3	G9	G10
12	**Matepatepa 2016/17**	62.2	-0.36	G12	G10	G8	G4
13	**Rusape 2016/17**	53.18	0.20	G12	G8	G10	G9
14	**Tengwe 2016/17**	60.89	-1.89	G9	G12	G11	G3
15	**Trelawney 2016/17**	53.78	-0.67	G9	G12	G11	G3

Table 6.15 shows the best performing varieties for quality grade across the seasons and locations. Variety G12 (K RK26R) had the best quality grade across the environments and season. It was followed by variety G8 (KE1 × AW3R) and G11 (BAZR3-6-18 × XSR4-7-10).

Table 6.15: Best four variety selections for quality grade measured across environments

Number	Environment	Mean	Score	1	2	3	4
1	**Centenary 2015/16**	84.31	1.02	G12	G8	G6	G4
2	**Kutsaga 2015/16**	88.72	0.64	G12	G8	G4	G6
3	**Makoholi 2015/16**	77.03	4.07	G6	G8	G12	G4
4	**Matepatepa 2015/16**	91.48	0.43	G12	G8	G9	G6
5	**Nebo 2015/16**	89.56	-0.45	G12	G7	G11	G8
6	**Rusape 2015/16**	91.01	0.07	G12	G8	G11	G9
7	**Tengwe 2015/16**	83.23	0.53	G12	G8	G9	G6
8	**Trelawney 2015/16**	74.27	1.44	G12	G4	G8	G10
9	**Centenary 2016/17**	77.51	-0.34	G12	G8	G9	G11
10	**Kutsaga 2016/17**	87.04	0.11	G12	G8	G11	G9
11	**Makoholi 2016/17**	91.90	-0.43	G12	G7	G3	G11
12	**Matepatepa 2016/17**	92.59	-0.01	G12	G8	G11	G9
13	**Rusape 2016/17**	75.37	0.38	G12	G4	G3	G1
14	**Tengwe 2016/17**	55.41	-7.05	G9	G7	G11	G3
15	**Trelawney 2016/17**	72.42	-0.39	G9	G8	G11	G12

AMMI Anova Table 6.16 shows that there were significant ($P < 0.05$) differences among the varieties for the trait mass at untying. The same trend was recorded for the locations. There were significant genotype × environment interaction effects ($P < 0.05$) for the trait (Table 6.16). Only two IPCA is sufficient to describe the extent of the genotype x environment interaction.

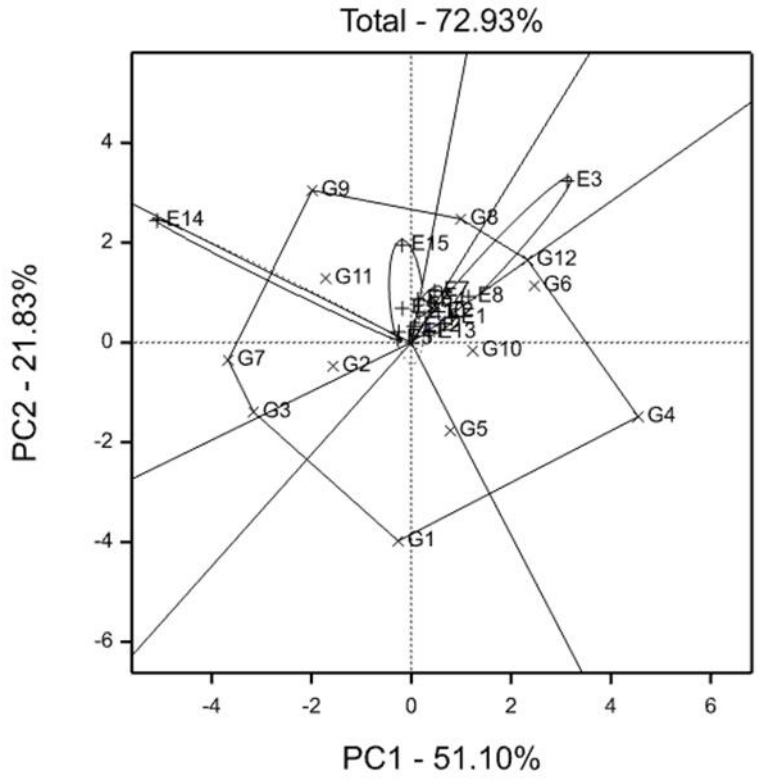


Figure 6.6: Which-won-where acrossite illustration for quality grade

The which-won-where biplot in Figure 6.7 shows that there were two mega-environments for the study season-location combination environments studied. The thee mega-environments converged at the centre of the biplot. One mega-environment involved Tengwe 2016/17 season-location combination with genotype 7 (K326R3-6-18 × K E1) being the best performer, the second mega-environment involved involved Trelawney 2016/17 season-location combination and genotype 9 (K326R3-6-18 × XZR2-2-4) was the best performing and the third had Makoholi 2015/16 season-location combination with genotype 8 (KE1 × AW3R) being the best.

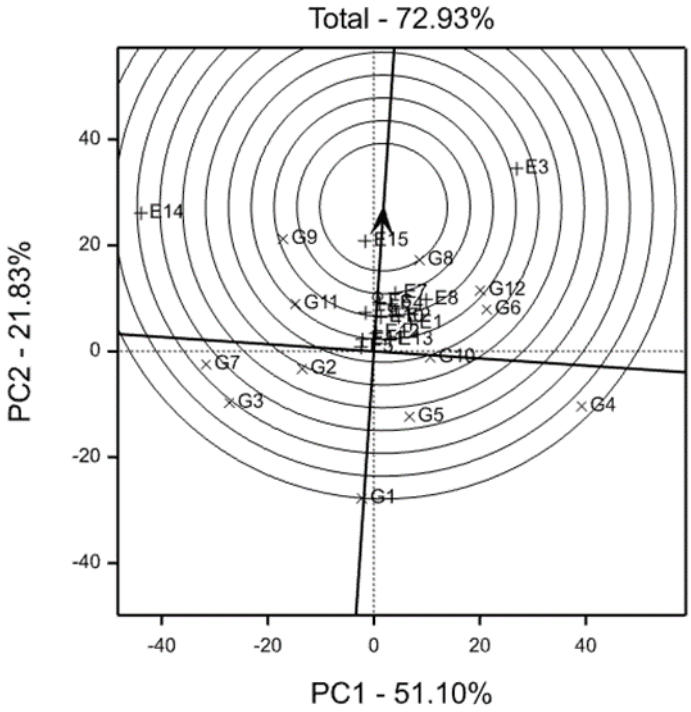


Figure 6.7: Acrossite comparison biplot for the trait quality grade

The acrossite comparison biplot for the trait quality grade shows that genotype 8 (KE1 × AW3R), followed by genotype 9 (K326R3-6-18 × XZR2-2-4) and genotype 11(BAZR3-6-18 × XSR4-7-10) are closest to the centre of the concentric rings (Figure 6.6). All the genotypes illustrated had their quality grade above average. Variety 4 (KE1 × XZR2-2-4) was the least stable on the acrossite comparison biplot.

6.5.1 Discussion

6.5.1.1 The 2015/16 season output

The significantly different output for all the traits measured was because the eight locations in which the varieties were tested lie in different altitudes of the country that receive different amounts of rainfall annually. Gauch and Zobel (1997) reported that differences in performance of genotypes grown across different environments can also be due to climatic factors. The test locations have widely different mean annual temperatures with Tengwe, Nebo Estates and Makoholi being among the hottest. Kutsaga Research Station and Rusape are situated on relatively high altitudes characterized by low temperatures while Nebo, Makoholi, Centenary and Tengwe are in the low altitude areas which are hot in summer. Trelawney and Matepatepa are medium altitude areas and experience moderate temperatures in summer. Apart from the climatic conditions, the areas also have different soil types with different capacities to support the tobacco genotypes. Yan and Tinker (2006) also reported that edaphic factors play a big role in determining the performance of a genotype in a specific environment as they influence the management approaches employed by a farmer. Although the soils in Centenary, Matepatepa, and Tengwe are gravelly, they are relatively fertile, deep and well drained which are good growing conditions for flue-cured tobacco. Nebo soils are relatively heavy with high clay content resulting in them giving thin tobacco which in most cases breaks down to give trashy leaves that have relatively high proportions of orange style tobacco. Trashy tobacco gives low grade indices and are rated low on the quality scale, hence, they also fetch low prices on the market.

The non-significant genotype × environment interaction effects for most of the traits except quality grade during the 2015/16 crop farming season contradicts results by Sadeghi *et al.* (2011) who reported the presence of genotype × environment interaction effects in flue-cured tobacco for cured leaf yield contrary to the results reported in this study. The findings of Sadeghi *et al.* (2011) contradict the findings of Farshadfar *et al.* (2012) who did not find significant genotype × environment interaction effects for the same trait.

Broad sense heritability values were also high for all the traits meaning that the conditions prevailing during the 2015/16 season favoured the passing on of the traits from parents to the offspring though slowly. The significant differences among the varieties for all the traits evaluated during the 2015/16 season (Table 6.2) could be because the varieties were constituted from different parentages with relatively different pedigrees, hence, they responded differently

to climate and management conditions presenting themselves in each location. Yield was relatively low during the season mainly because of the mid-season drought experienced in January of the 2015/16 season.

The significance of the trait quality grade during both the 2015/16 and 2016/17 seasons (Table 6.6) could be because the two seasons were normal with relatively high rainfall. Quality grade in flue-cured tobacco is dependent on a series of parameters which could be genetic or emanating from the handling of the crop from harvesting, curing to grading (Sadeghi *et al*., 2012). The ease of handling of a variety resulting in a high quality rating has being reported to be perculiar to certain genotypes by Ya-Nan *et al*. (2012).

The two mega-environments that emerged based on the environmental characteristics of the areas studied conform to Gauch and Zobel (1997) suggestions when they indicated that the basis of demarcation of mega-environments and indicated that mega-environments can be huge expanses of land cutting across countries or even continents. They showed that mega-environments can be huge uninterrupted growing areas or can be interrupted by different environments in some cases. In this study, two mega-environments emerged with one grouping Makoholi and Centenary together. Centenary was grouped with Makoholi since these locations are characterized by a prevalence of high temperatures and sandy, gravelly soils although Centenary gets relatively high rainfall annually of above 800 mm and Makoholi gets an average of 600 mm annually. Trelawney, Tengwe, Kutsaga, Rusape, Nebo and Matepatepa are high potential areas of Zimbabwe receiving relatively high rainfall. Hot environments usually result in difficulties in handling cured tobacco leaf resulting in defects like sun scalds and trashy tobacco leaf. However, it appears as if varieties with KE1 as one of the parents in a cross do well under the high temperature environments.

On Table 6.7, variety 12 (K RK26R) gave the best performance for quality as it won in five of the eight locations. It was the best in the environment grouping involving Centenary and Makoholi possibly because of the high temperatures and the mid-season droughts often experienced in these areas. It appears as though genotypes possessing K E1 as a parent do well in these locations as seen by genotypes 8 (KE1 × AW3R) and 6 (ONCR3-4-6 × KE1) coming out first in these areas.

6.5.1.2 Results for the 2016/17 season

Locations were significant during the 2016/17 season again showing the differences that exists among the locations, edaphically, and climatically. In this study genotype × environment interaction may have been due to differences caused by the differences in soil properties and characteristics in the areas studied which required that farmers in specific areas apply different crop management regimes resulting in different performances of the same crop variety when grown in other locations as pointed out by Gauch and Zobel (1997). In this study genotypes performed differently in each of the environments they were grown confirming the differences that exist among the genotypes for the traits studied.

The three mega-environments with the location Tengwe clustering alone, Matepatepa and Rusape in the second cluster and the other locations (Centenary, Kutsaga, Makoholi and Trelawney) constituting the third cluster differs from the tradition mega-environments used currently in tobacco farming in Zimbabwe which groupd Matepatepa and Trelawney together. Tengwe stood out from the rest possibly because of excessive rains received in the area during the season which severely reduced the tobacco cured leaf quality on this site. Most of the tobacco obtained at this location was very dark to the extent of falling into the "black" quality class (TIMB, 1999). Matepatepa and Rusape had relatively light tobacco with relative high proportions of tobacco leaf and the same scenario was with Centenary, Kutsaga, Makoholi and Trelawney. GGE biplot analysis by Yan (2001) allows for ranking of genotypes based on their performance in each and across the test environments. Genotypes 3(AW3R×KM10), 8(KE1×AW3R), 11(BAZR3-6-18×XSR4-7-10) and 12 (KRK26R) where the best performers across all the sites during the 2016/17 season maybe because they had parents sharing a common ancestor, K E1. Raeber and Smeeton (1970) reported that the line K E1 is heavy bodied and tended to increase weight in the varieties that contain it in ancestry or as one of the parents in a hybrid. The unrepeated which-won-where pattern for the traits under study show that varieties where not consistent in their performance and hence could not be conclusively recommended for any location.

6.5.1.3 Results of the across – site analysis

The main conclusions about the behavior of genotypes in this study are based on the across-site analysis. Across the seasons and locations, the environments showed significant differences for all the parameters under study showing that the environments were unique and supported the genotypes grown, differently (Table 6.10). Working with maize, Gauch and

Zobel (1997) emphasized that environments differ in their ability to support different genotypes agronomically. Significant genotype × environment interaction were also reported on dry leaf yield by Sadeghi *et al*. (2011). The significant differences among varieties for traits measured meant that the varieties responded differently to the different environments in which they were grown. Sadeghi *et al.* (2011) also pointed out that in tobacco production, genotypes that give high average yields across environments should be preferred which is in contrast to the general recommendations of Gauch and Zobel (1997) that cultivars have to be targeted to environments where they perform best agronomically, which they termed delineation of the environment. Maleki *et al*. (2011) also indicated that dry leaf yield in flue cured tobacco is highly influenced by the environment and that direct selection for the trait is highly unpredictable. The non-significance of genotype × environment interaction on dark cured leaf color means that the trait is not influenced by the locality where a genotype is grown. The traits with significant genotype × environment interaction mean that the parameters responded differently in the genotypes under different environment conditions.

The GCV and heritability estimates provide reliable estimates of the amount of GA to be expected through phenotypic selection (Mishra *et al*., 2015). When GCV and heritability values are high, as in this study, then more reliable information is provided than singular parameters (Baye *et al*., 2005). The high GCV values showed that the genotypes in this study had wide variation among themselves for morphological and genetic traits studied.

6.5.1.4 Mega environment analysis for quality grade

The polygon view of a GGE biplots in the results section of this study show the which-won-where pattern summarizing the genotype environment interaction pattern for the trait quality grade (Figures 6.2, 6.4 and 6.6). Just like what was illustrated by Yan (2002), the polygons were formed by joining vertex genotypes. Vertex genotypes have the longest vectors in their respective directions on the biplots meaning that they are the most responsive to the environments studied for the trait quality grade. All the other genotypes not on the vertices were contained within the polygons meaning that they were responsive in terms of the interaction with the environments occurring within the sectors they occur (Yan and Kang, 2003).

For each of the which-won-where biplots, lines were drawn from the biplot origin intersecting the sides of the each of the polygon at right angles dividing the biplots into sectors and also dividing the target environments into sub-regions or mega-environments. Sectors that have

one or more environments are called mega-environments (Gauch and Zobel, 1988). In this study, for the trait quality grade, and during the 2015/16 season, two mega environments were identified, in the 2016/17 season, three mega environments were also found, while in the across site biplot only genotypes, 9 (K326R3-6-18 × XZR2-2-4), 8 (KE1 × AW3R), 12 (K RK26R), 4 (KE1 × XZR2-2-4), 1 KE1 × K M10, 3 (AW3R × KM10), and 7 (K326R3-6-18 × K E1) were the vertex genotypes giving three mega environments.

6.4.6.5 Comparison analysis for quality grade

Figures 6.3, 6.5, and 6.7 showed the GGE-biplots based on the genotype-focused scaling for comparison of the genotypes studied with the ideal genotype for quality grade. According to Yan and Kang (2003) the ideal genotype is the one that has the highest mean performance and is absolutely stable. The duo stated that this variety is in the centre of the concentric circles and the arrow in the biplots point to it. The ideal genotype must also have the greatest vector length of the high quality grade genotypes in the study and also have the no genotype × environment interaction. It must have high average yield in all the environments studied and is graphically defined by the longest vector in PC2. And with no projections in PC2 (Yan, 2002).

The concentric circles in the Figures 6.3, 6.5. and 6.7 only serve to visualise the distances between each genotype and the ideal genotype. For a genotype to be regarded as having high quality grade and also highly stable for the trait, it must be located closer to the ideal genotype. So, during the 2015/16 season, genotypes 4 (KE1 × XZR2-2-4) and 8 (KE1 × AW3R) were close to ideal and during the 2016/17 season, genotype 3 (AW3R × KM10) approached ideal just like genotypes 8 (KE1 × AW3R) and 9 (K326R3-6-18 × XZR2-2-4) on the across site biplot in line with the recommendations of Rad *et al*. (2013).

6.6 Conclusions and recommendations

6.6.1 Conclusions

1. Genotype × environment interaction effects exist for quality grade, grade index and mass at untying but not for dark cured leaf colour in flue-cured tobacco farming in Zimbabwe.
2. Genotype 8 (KE1 × AW3R) was the most stable for quality grade while genotype 12 (K RK26R) was the most stable for the traits mass at untying and grade index.

6.6.2 Recommendations

1. Breeders need to factor in the availability of genotype × environment for the traits quality grade, grade index and mass at untying when developing flue-cured tobacco varieties and when recommending developed varieties for growing in specific regions of the country.
2. Breeders also need to understand the growing environment they intend to recommend a variety for them to appropriately match new varieties to appropriate growing environments to better exploit genotype × environment interaction for the benefit of farmers.
3. Agronomists and flue-cured tobacco breeders are recommended to utilize varieties with high stability like KE1 × AW3R across environments as they do not result in reduced performance across environments.

CHAPTER SEVEN

General Discussion

7.1 Introduction

This study entitled “Genetic characterization and analysis for yield-related traits in the Zimbabwean flue-cured tobacco germplasm” was an investigation into the genetic diversity, gene action for selected traits and the relative adaption of the available Zimbabwean flue-cured tobacco germplasm which is being used to develop currently grown varieties. The study included all the major accessions available in the country which are the progenitors of the presently grown local and regional flue-cured tobacco varieties. This was done to assess the extent of the genetic relationship among these accessions which originated from different parts of the world and assess the potential utility of the accessions in developing new varieties for the future. In addition, a study was conducted to ascertain the availability of genotype × environment interaction effects in flue-cured tobacco farming in Zimbabwe since tobacco growing areas in the country have been divided into different regions based on their ability to support flue-cured tobacco crop farming. This was done to ascertain whether the subdivision into growing regions has only commercial basis and not backed by empirical evidence on genotype × environment interaction effects for major traits in flue-cured tobacco production. The nature of inheritance of some traits of flue-cured tobacco were also studied in order to determine the nature and behaviour of genes controlling these important traits of the crop so that breeding schemes can be planned appropriately with relative predetermined success. Traits like growth rate, ripening rate and dark cured leaf colour are very important to growers of flue cured tobacco because they determine the acceptability of the crop on the market. This study was carried out with the following objectives:

1. To establish the genetic relatedness among elite Zimbabwean tobacco lines using SSR markers.
2. To determine the genetic diversity of Zimbabwean flue-cured tobacco germplasm by phenotypic characterization.
3. To investigate the mode of inheritance of growth rate, ripening rate and dark cured leaf colour in flue-cured tobacco germplasm.
4. To determine yield performance and stability of tobacco hybrid varieties across the three growing environments in Zimbabwe.

7.2 The study's major findings and their implications

Although flue-cured tobacco (*Nicotiana tabacum* L.) is a very popular foreign currency earning field crop in Zimbabwe among both the smallholder and large scale famers, very little scientific study has been done in the area of diversity among available germplasm, genetics of major traits of the crop and elite crop variety adaptation in the country. Many studies on tobacco diversity were, however, conducted in countries such as USA, China, Iran, India and Brazil involving both morphological and molecular approaches (Moon *et al*., 2009b; Ya-Nan *et al*., 2012; Shah *et al*., 2015; Parajuli *et al*., 2015). Results of these and several other studies showed that global elite flue-cured tobacco germplasm is closely related, especially at molecular level (Moon *et al*., 2009b; Lahiji *et al*., 2013; Maryan *et al*., 2012). This earlier finding has often been supported by the fact that elite tobacco germplasm was shared quite extensively across national and continental boundaries during the early years of popularization of the crop.

Although there is some slight resemblance, the findings in the above mentioned countries do not agree with the findings of this study especially showed that the Zimbabwean flue-cured tobacco germplasm shows some variability though little. The little variability available can be a result of germplasm inter-mating among and between germplasm originating from different parts of the world and the subsequent natural and artificial selection that takes place as the accessions adapted to the local growing conditions that could have happened in the country. Available similarities in the overall morphological analysis output can be attributed to the common traits like leaf shape, leaf length and fresh leaf color as well as plant height which are widespread in all global germplasm. Like the findings in the studies of Dengduangboripant *et al.* (2012) who also reported present and potentially useful diversity in their native flue-cured tobacco germplasm in Iran, this study also found some diversity within the Zimbabwean flue-cured tobacco germplasm. Studies by Fricano *et al.* (2012) using Simple Sequence Repeats (SSR) markers also reported the presence of exploitable variability in the flue-cured tobacco lines obtained from different parts of the world, which are now being kept as a worldwide collection of germplasm.

Chapter Four is an additional study of agro-morphological diversity involving a wider base of germplasm spanning more than a century which includes early introductions, elite lines used to constitute modern varieties and previously cultivated varieties which have lost popularity but are potential sources of valuable genes for different traits (obsolete varieties). The material also includes some of the key progenitors of elite cultivars of the flue-cured tobacco popular in

Zimbabwe. The general findings of the study in Chapter Four show that there is relatively wide diversity within the available pool of accessions in Zimbabwe, based on the TwinSpan clustering, from which new superior cultivars can be selected although there is a moderate degree of relatedness among most of the accessions available for breeding work in the country. Flue-cured tobacco cultivars grown in Zimbabwe and developed during the past ten years originate from similar ancestry constituted using germplasm mainly from the USA. These varieties lack diversity among themselves and are, therefore, vulnerable to common biotic and abiotic pressures prevalent in the country. Infusing some genes from the old varieties and abandoned promising lines like K ME and Kutsaga E1 can improve the diversity and the performance of the modern flue-cured tobacco varieties in the country. The results of this morphological diversity study, however, indicated that there is still diversity present among the accessions available and this diversity can be exploited especially in traits like leaf size, leaf duration, leaf number as well as leaf pigmentation which all contribute to improved cured leaf yield and quality.

The exclusion of widely variable flue-cured tobacco germplasm in the modern variety development programmes is a result of the fear of altering other important traits of flue cured tobacco such as quality, sugar nitrogen balance, handling properties and nornicotine content (Moon *et al*., 2009a; Sadeghi *et al*., 2012). Nornicotine is the chemical compound found in tobacco known for causing cancer in tobacco smokers while the sugar nitrogen balance determines the smoke flavor and impact of tobacco cigarettes . Quality grade and chemical composition are very important traits in flue-cured tobacco processing and utilizations as they have relatively high influence on the price of the crop on the market and the quality of the resulting cigarettes made from the flue-cured tobacco leaves (Ya-Nan *et al*., 2012). While it might be important to preserve such traits in the industry, there is the risk that varieties developed might not be able to achieve significant improvements in important traits to farmers like leaf yield which guarantees profitability of the crop.

Over the past ten years, varieties of flue-cured tobacco developed in Zimbabwe have failed to surpass the 4.5 t/ha yield mark. This could be a result of the continued recycling of the same germplasm in the development of cultivars. Exploitation of variation existing within germplasm of the same species brings about new gene combinations that may result in the expression of new advantageous traits in crops (Nduwumuremyi *et al*., 2013). For better yields of varieties to be produced in the present flue-cured tobacco industry, there is need to avoid the

use of closely related germplasm in the creation of new crosses. Crosses made must be dependent on the use of widely related material in order to broaden the diversity among the offspring produced from which the breeder selects useful lines (Fricano *et al*., 2012).

Results of Chapter Four also show clusters of accessions grouped depending on their genetic and morphological relationships. The results give an indication of the material to use in the creation of variability during foundation crop improvement work. The sets of accessions occurring in each of the clusters shown from Figures 4.1 to 4.2 in Chapter Four can be treated as possible heterotic groups from which lines for inter-crossing in complex crosses can be drawn.

The study in Chapter Five focused on the genetic parameters of the three major traits of flue-cured tobacco namely growth rate, ripening rate and flue-cured leaf color. Growth rate in flue-cured tobacco has often been associated with the yield potential of a variety as well as the ease with which it can be managed in the field. Farmers exploit this trait based on the location of their farms within the tobacco growing belt of the country. Farmers in the high temperature and often low altitude areas prefer slow growing varieties while those in the cool areas, which are high altitude areas, go for the relatively fast growing varieties. Also linked to growth rate, is the trait "ripening rate". Ripening rate is the speed with which a tobacco variety's leaves ripen. The standard acceptable ripening rate is two leaves per week. Varieties that ripen at a rate of more than two leaves per week are regarded as fast ripening while those that ripen at a rate less than two leaves per week are slow ripening. Similar to with growth rate, varieties that are fast ripening are recommended for the cool and cold areas while the slow ripening varieties are recommended for growing in the hot areas. In Zimbabwe, the hot areas are the low altitude areas such as the Hurungwe, Chinhoyi, Centenary, Masvingo and the northern parts of Mashonaland Central Province while the cool to cold areas are the high altitude areas such as Harare, Marondera and many parts of Manicaland and Mashonaland East and West provinces.

Flue-cured tobacco leaf color is also classified as either dark (orange or mahogany) or light (lemon). Early studies have linked flue-cured tobacco leaf color to the amount of nicotine present in the leaf and the rate of growth and ripening of the variety (Ahmed, 2014). Unconfirmed but widely accepted views in the tobacco farming industry in Zimbabwe claim that slow growing and ripening varieties throw dark style tobacco while fast growing and ripening varieties throw the light lemon cured varieties. Zimbabwe is known globally for producing the relatively dark style flue-cured tobacco which has relatively high nicotine

content, hence, farmers aim at producing more of this type although they also want the lemon styles especially for the Asian markets (Ya-Nan *et al*., 2012). Flue-cured tobacco breeders, therefore, always aim to have varieties that are capable of producing both dark and light cured styles on the market to cater for the preferences of all farmers, hence, the need to understand the genetic properties of the traits.

Growth rate, ripening rate and cured leaf color are very important traits bred for by flue-cured tobacco breeders in Zimbabwe. They determine the usefulness of a variety developed to both the growers of the variety and the buyers of the flue-cured leaf reaped from the variety which is overally refered to as the quality of the leaf (Ya-Na *et al*., 2012). Fast growing and fast ripening varieties produce light tobacco which breaks easily and gives trashy tobacco that fetches very low prices on the market because it is not attractive. Dark style flue-cured tobacco leaves are relatively firm and are preferred by the traditional markets of tobacco in the country. Understanding the nature of these traits at gene level helps breeders develop varieties with a combination of these traits with ease (Patel *et al*., 2012).

Results of the genetic studies on the three parameters of flue-cured tobacco in Chapter Five showed that dark flue-cured leaf colour is controlled by additive gene action. Additive gene action is associated with average effects of genes (Nduwumuremyi *et al*., 2013). Traits exhibiting average effects can be easily bred for with predictable success. For all the three traits, the SCA effects were not significant suggesting the absence of dominance effects. Dominance effects are associated with above average effects in crop improvement (Butorac *et al*., 2004). Traits that exhibit dominance gene action have been reported to have high heritability although environmental factors have been found to influence the expression of most quantitative traits (Korubin-Aleksoska, 2012). Results in this study suggests that it is possible for breeders to select for the trait with moderate heritability but those with low heritability might be difficult. Ya-Nan *et al*. (2012), however, warned that epistasis and genotype × environment interaction have to be considered when genetic studies are done in flue-cured tobacco.

In flue-cured hybrid seed production targeting the traits: growth rate, ripening rate and dark flue-cured leaf colour, it does not matter which parent is made male or female as this will not influence the expression of the trait being improved. In flue-cured tobacco hybrid development, it has been very difficult to make an interchange of parentals especially when the experimental female parent does not produce high flower numbers. The non-significance of the reciprocal

effects in this study confirms that parentals can be interchanged in hybrid seed production without reducing the performance of the original variety. The seed producer must be guided by the relative flower prolificacy of the parents used to determine the female or male parent. This result if also corroborated by Darvishzadeh *et al*. (2013) who reported that reciprocal effects are not important in flue-cured tobacco hybrid seed production.

The flue-cured tobacco farming belt in Zimbabwe has been divided into the fast, medium and slow growing areas since the 1960s (Schwepenhauser and Raeber, 1966). Even the germplasm used in the development of varieties grown by farmers and the cultivars developed using the germplasm have been grouped into three groups: fast, medium and slow ripening varieties. Results of studies on genotype × environment interaction in flue-cured tobacco in other countries like the United States of America and Brazil indicated that genotype × environment interaction was not important in tobacco farming. This was later contradicted by results of studies in Iran, China, India and Pakistan where genotype × environment interaction was significant (Ahmed, 2014). This study conducted in Zimbabwe over the period 2015-2017 also showed that genotype × environment interaction in flue-cured tobacco is significant (Table 6.4) for traits like quality grade, grade index, and mass at untying (ungraded leaf yield). This means that varieties of flue-cured tobacco perform differently for yield, quality grade and grade index depending on the area they have been grown in the flue-cured tobacco growing belt in Zimbabwe. To get maximum yield, varieties must be grown in the areas where they perform best. Before they are recommended for growing in the country, flue-cured tobacco varieties must be evaluated in more than one location so that best performers are identified and recommended to specific areas. Stable varieties can also be identified and recommended where appropriate.

The study in Chapter Six also showed that the tobacco growing environments in Zimbabwe are not homogeneous. They differ in their capacity to support different varieties of the crop, hence, different varieties yield differently and give different quality grades depending on the area they are grown within the tobacco growing belt of Zimbabwe. Varieties have been evaluated in each of the environments to determine those with either broad or specific adaptation. Those varieties with broad adaptation may be grown in all the environments. It must, however, be remembered that most varieties with broad adaption exhibit average performance for the traits of interest and may not be the best choice if the aim is to increase average national yields.

CHAPTER EIGHT

Conclusions and Recommendations

8.1 Conclusions

From the findings of the study, it can be concluded that:

1 Molecular analysis using SSR markers has shown that flue-cured tobacco germplasm in Zimbabwe has a narrow genetic diversity. There is a large proportion of common genes among accessions studied suggesting a high likelihood of common ancestry among them.

2 There is relatively narrow agro-morphological diversity among the flue–cured tobacco germplasm in Zimbabwe. The present agro-morphological diversity can be sufficiently grouped into ten clusters based on their agro-morphological similarity.

3 In flue-cured tobacco, dark cured leaf colour, days to 50% topping and days to 50% reaping exhibit partial dominance mode of inheritance while reciprocal and maternal effects were not significant for all the traits studied. Only the trait dark cured leaf colour showed additive gene action.

4 There is significant genotype x environment interaction for mass at untying, grade index and quality grade of flue-cured tobacco in Zimbabwe. BAZR3-6-18 × XSR4-7-10 has shown to be the most stable variety in the study.

8.2 Recommendations

1. Since there is narrow diversity among the flue cured tobacco germplasm genetically and morphologically, it is recommended that breeders of flue-cured tobacco work carefully in exploiting the little variation present in the development of new cultivars of flue cured tobacco with superior performance for both quality and yield. It is recommended that crosses are made between accessions from each of the identified ten clusters and not within a clusters. Making wide crosses of lines with wide morphological variation across observed groups may result in gene reshuffling and emergence of better varieties. Care must be taken, however, to preserve the chemical balance of tobacco varieties.

2. Yield, ripening rate and growth rate of flue-cured tobacco in Zimbabwe are controlled by additive gene action. This means that there is merit in selection for these traits in flue-cured tobacco breeding. Breeders of flue-cured tobacco are, therefore, recommended to combine genotypes like KE1 × AW3R, K326R1-1-28 × XZR2-2-4, and BAZR3-6-18 × XSR4-7-10 in their crosses for them to come up with new varieties that are relatively high yielding, growing and ripening relatively slowly to give relatively high leaf quality since these genotypes have shown broad adaption in the growing environments they are grown.

3. Stable genotypes like BAZR3-6-18 × XSR4-7-10 can be grown in any environment in the Zimbabwe flue-cured tobacco growing belt.

REFERENCES

Acquaah, G. 2012. *Principles of plant genetics and breeding*. 2nd Ed. Wiley-Blackwell, Oxford. 758 pp.

Ahmed, S., Mohammad, F., Ahmed, Q. and Khan, M. A. U. 2014. Assessing genetic variation for morpho-agronomic traits of some native and exotic FCV tobacco genotypes in Pakistan. *American-Eurasian J Agric & Environ Sci, 14*(5), pp.428-433.

Ahmed, Q., Mohammad, F., Rahman, H., Ahmed, S, and Fakharuddin. 2016. Heterotic studies in flue-cured tobacco across environments. *Sarhad Journal of Agriculture,* 32(2): pp. 112-120.

Ali, A. and Ali, N. 2010. Assessment of Inheritance Pattern of the Two Morphological Markers in Intraspecific Hybridization of *Brassica Napus* L. *Sarhad Journal of Agriculture,* 26(3), pp. 369-371.

Al-Khayri, J.M., Jain, S.M. and Johnson, D.V. eds., 2015. *Advances in plant breeding strategies: Breeding, biotechnology and molecular tools*. Springer International Publishing. 707 pp.

Allard, R. W. 1999. *Principles of plant breeding*. 2nd ed., Wiley, New York. 264 pp.

Alonso-Blanco, C., Aarts, M. G. M., Bentsink, L., Keurentjes, J. J. B., Reymond, M., Vreugdenhil, D. and Koornneef, M. 2009. What has natural variation taught us about plant development, physiology, and adaptation. *The Plant Cell Online,* 21, pp. 1877-1896.

Amiryousefii, A., Hyvönen, J. and Poczai, P. 2018. iMEC: Online marker efficiency calculator. *The Application in Plant Science,* 6, e 1159.

Argawal, M., Shrivastava, N. and Padh, H. 2008. Advances in molecular marker technologies and their application in plant science. *Plant Science Report,* 27, pp. 617-631.

Babu, N. N., Hittalmani S., Shivakumar, N. and Nandini, S. 2011. Effect of drought on yield potential and drought susceptibility index of promising aerobic rice (*Oryza sativa* L .) genotypes. *Electronic Journal of Plant Breeding,* 2(3), pp. 295-302.

Baye, B., Ravishankar, R. and Singh, H., 2005. Variability and association of tuber yield and related traits in potato (Solanum tuberosum L.). *Ethiopian Journal of Agricultural Sciences*, *18*(1), pp.103-121.

Bernardo, R. 2008. Molecular markers and selection for complex traits in plants: learning from the last 20 years. *Crop Science*, 48(5), pp.1649.

Bindler, G., Plieske, J., Bakaher, N., Gunduz, I., Ivanov, N., Van Der Hoeven, R., Ganal, M. and Donini, P. 2011. A high density genetic map of tobacco (*Nicotiana tabacum* L.) obtained from large scale microsatellite marker development. *Theoretical and Applied Genetics*, 123 (2), pp. 219-230.

Bonierbale, M. D., Plaisted, R. L., Tanksley, S. D. *(*1988*).* RFLP maps based on a common set of clones reveal modes of chromosomal evolution in potato and tomato. *Genetics,* 120*, pp.* 1095-1103.

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

Bourguet, D., Chaufaux, J., Micoud, A., Delos, M., Naibo, B., Bombarde, F., Marque, G., Eychenne N., Pagliari C. 2002. *Ostrinia nubilalis* parasitism and the field abundance of non-target insects in transgenic *Bacillus thuringiensis* corn (*Zea mays* L). *Environmental Biosafety Research,* 1, pp. 49-60.

Bradu, D. and Gabriel, K. R. 1978. The biplot as a diagnostic tool for models of two-way tables. *Technometrics*, 20, pp. 47-68.

Butorac, J., Beljo, J. and Gunjaca, J. 2004. Study of the inheritance of some agronomic and morphological traits in burley tobacco by graphic analysis of diallel cross. *Plant Soil Environment,* 50 (4), pp. 162-167.

Carvalho, B. L., Ramalho, M. P., Pulcinelli, C. And Bruzi, A. 2014. Genetic parameters estimates associated to conversion of nicotine to nornicotine in burley tobacco. *American Journal Of Plant Sciences*, 05(21), pp. 3380-3388.

Charcosset, A., Denis, J. B., Lefort-Buson, M., and Gallais, A. 1993. Modelling interaction from top-cross design data and prediction of F1 hybrid value. *Agronomie EDP Sciences*, 13 (7), pp. 597-608.

Chena, H., Chenb, H., Hua, L., Wanga, l. and Wanga, S. 2017. Genetic diversity and a population structure analysisof accessions in the Chinese cowpea [*Vigna unguiculata* (L.) Walp.] Germplasm Collection. *The Crop Journals*, 2, pp. 363-272.

Ciccarelli, S., 1989. Wide adaptation: How wide. *Euphytica*, 40, pp. 197-205.

Collard, B. C. Y. and Mackill, D. J. 2008. Marker Assisted Selection: An approach for precision plant breeding in the twenty-first century. *Philosophical Transactions of the Royal Society B*, pp. 557- 571.

Comstock, R. E. and Moll, R. H. 1963. Genotype × Environment Interactions. Symposium On Statistical Genetics and Plant Breeding. *National Academy Science National Research Council*. Washington D. C., pp. 164-196.

Cooper, M. and DeLacy, I. H. 1994. Relationships among analytic methods used to study genotypic variation and genotype-by-environment interaction in plant breeding Multi-environment experiments. *Theoretical and Applied Genetics*, 88, pp. 561-572.

Crossa, J. 1990. Statistical analysis of multi-location trials. *Advances in Agronomy*, 44, pp. 55-85.

Crossa, J., Campos, G., Perez, P., Gianola, D., Burgueno, J., Araus, J. L., Makumbi, D., Singh, R. P., Dreisigacker, S., Yan, J., Arief , V., Banziger, M. and Braun, H. 2010. Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers. *Genetics*, 186(2), pp. 713-724.

Dabholkar, A. R. 1999. *Elements of biometrical genetics*. Concept Publishing Company, New Delhi, India. 493 pp.

Dadras, A. R., Sabouri, H., Nejad, G. M., Sabouri, A, Shoai-Deylami, M. 2014. Association analysis, genetic diversity and structure analysis of tobacco based on AFLP markers. *Molecular Biology Reports*, 41(5), pp. 3317-3329.

Darvishzadeh, R., Mirzae, L, Maleki, H. H., Laurentin, H., Alavi, S. R. 2013. Genetic variation in oriental tobacco (*nicotiana tabacum* l) by agromorphological traits and simple sequence repeat markers. *Revista Ciencia Agronomica,* 44(2), pp. 347-355.

Davalieva, K., Maleva, I., Filiposki, K., Spiroski, O., and Efremov, G. D. 2010. Genetic variability of macedonian tobacco varieties determined by microsatellite marker analysis. *Diversity,* 2, pp. 439-449.

Davis, D. L. and Nielsen, M. T. 1999. Tobacco: Production Chemistry and Technology. Blackwell Science. USA. 467 pp.

Del Piano, L., Abet, M., Sorrentino, C., Acanfora, F., Cozzolino, E. and De Muro, A. 2000. Genetic variability in *Nicotiana tabacum* and *Nicotiana* species as revealed by RAPD markers. I. Development of the RAPD procedure. *Beitrage zur Tabakforschung International/Contribution to tobacco research*, 19, pp. 1-15.

Denduangboripant, J., Piteekan, T. and Nantharat, M. (2010). Genetic polymorphism between tobacco cultivar groups revealed by Amplified Fragment Length Polymorphism analysis. *Journal of Agricultural Science*, 2(2), pp. 41-48.

Ding, M., Tier B. and Yan W., 2008. Application of GGE Biplot Analysis to Evaluate Genotype (G), Environment (E) and GxE Interaction on {\em Pinus radiata}: A Case Study. *New Zealand Journal of Forestry Science*, 38(1), pp.132-142.

Djulgerski, Y. 2010. Studies on the inheritance of the number of leaves in the crosses of Virginia tobacco (*Nicotiana tabacum* L). *Tobacco*, 60(1-6): pp. 12-16.

Dobzhansky, T. H., and Levene, H. 1955. Genetics of natural populations. XXIV. *Genetics*, 40, pp. 797–808.

Doganlar, S., Frary, A., Daunay, M. C., Lester R. N., Tanksley, S. D. 2002. A comparative genetic linkage map of eggplant (*Solanum melongena*) and its implications for genome evolution in the Solanaceae. *Genetics*, 161, pp. 1697–1711.

Dolatabad, S. S., Hervan, R. C. E. M. and Dehghani, H. 2010. Multienvironment analysis of traits relation and hybrids comparison of maize based on the genotype by trait biplot. *American Journal of Agricultural and Biological Sciences*, 5(1), pp.107-113.

Earl, D. A. and von Holdt, B. M. 2012. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno Method. *Conservation of Genetics Resources*, 4, pp. 359-361.

Eberhart, S. A. and Russell W. A. 1966. Stability Parameters for Comparing Varieties. *Crop Science*, 6, pp. 36-40.

Ebdon, J. S. and Gauch, H. G. 2002. Additive Main Effect and Multiplicative Interaction Analysis of national turfgrass performance Trials: II. Cultivar recommendations. *Crop Science*, 42, pp. 497-506.

Edmeades, G. O., Bänziger, M., Beck, D., Bolaños, J.and Ortega. A. 1997. Development and per se performance of CIMMYT maize populations as drought-tolerant sources. In: Developing drought and low-N tolerant maize, ed. G.O. Edmeades, M. Bänziger, H.R. Mickelson, and C.B. Peña-Valdivia, CIMMYT: Mexico, pp. 254-262.

Eshghi, R., Abrahimpour, F., Ojaghi, J., Salayeva, J., Baraty, M. and Rahimi, M. 2012. Evaluation of genetic variability in naked barley (*Hordeum vulgare* L.). *International Journal of Agriculture and Crop Sciences*, 4 (16), pp. 1166-1179.

Evanno, G., Regnaut, S and Goudet, J. 2005. Detecting the Number of Clusters of Individuals Using the Software STRUCUTRE: A Simulation Study. *Molecular evolution*, 14, pp. 2611-2620.

Excoffier, L. and Lischer, H. E. L. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analysis in Linux and Windows. *Molecular Ecology Resources*, 10, pp. 564-567.

Falconer, D. S. and Mackay, T. F. C. 1996. *Introduction to quantitative genetics*. 4th Ed. Pearson Prentice Hall. Harlow, England. 480 pp.

Farshadfar, E. Mohammedi, R., Aghaee, M., Vaisi, Z. 2012. GGE Biplot analysis of genotype × environment interaction in wheat-barley disomic addition lines, *Australian Journal of Crop Science*, 6(6), pp. 1074-1079.

Finlay, K. W. and Wilkinson, G. N. 1963. The analysis of adaptation in a plant breeding programme. *Crop Pasture Science*, 14(6), pp. 742-754.

Forsberg, S. and Carlborg, O. 2017. On the relationship between epistasis and genetic variance heterogeneity. *Journal of Experimental Botany,* 68, pp. 5431-5438.

Fricano, A., Bakaher, N., Corvo, D. C., Piffanelli, P., Donini, P., Stella, A., Ivanov, N. V. and Pozzi, C. 2012. Molecular diversity, population structure, and linkage disequilibrium an a worldwide collection of tobacco (*Nicotiana tabacum* L.) germplasm. *BMC Genetics, 13*, pp. 18.

Gabriel, K. R. 1971. The biplot graphic display of matrices with application to principal component analysis. *Biometrika,* 58, pp. 453-467.

Gauch, H. G and Zobel, R. W. 1997. Identifying mega-environments and targeting genotypes. *Crop Science*, 37, pp. 31-326.

Gauch, H. G. 2006. Statistical analysis of yield trials by AMMI and GGE. *Crop Science, 46,* pp. 1488-1500.

Gauch, H. G. 1988. Model Selection and Validation for Yield Trials with Interaction. *Biometrics,* 44 (3), pp. 705-715.

Gauch, H.G., 1992. *Statistical analysis of regional yield trials: AMMI analysis of factorial designs*. Elsevier Science Publishers.

Gauch, H. G. 2006. Statistical Analysis of Yield Trials by AMMI and GGE. *Crop Science,* 46, pp. 1488-1500.

Gholizadeh, S., Darvishzadeh, R., Mandoulakani, B. A., Bernousi, I., Alavi, S. R. and Masouleh, A. K. A. 2012. Molecular characterization and similarity relationships among flue-cured tobacco (*Nicotiana tabacum* L.) genotypes using Simple Sequence Repeat markers. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca*, 40(1), pp. 247-253.

Gixhari, B. and Sulovari, H. 2010. Nature of inheritance and heterosis estimated on some morphological quantitative characters that influence the tobacco yield. *Universitatea "Vasile Alecsandri" din Bacău. Biologie*,18, 46-50.

Glaszmann, J. C., Kilian, B., Upadhyaya, H. D. and Varshney, R. K. 2010. Accessing genetic diversity for crop improvement. *Current Opinion in Plant Biology*, 13, pp.1-7.

Goncalves, G. M., Viana, A. P., do Amaral, A. T. and de Resende, M. D. V. 2014. Breeding New Sugarcane Clones by Mixed Models Under Genotype by Environmental Interaction. *Scientia Agricola,* 71 (1), pp. 66-71.

Govindaraj, M., Vetriventhan, M. and Srinivasan, M. 2015. Importance of genetic diversity assessment in crop plants and its recent advances: An overview of its analytical perspectives. *Genetics Research International*, pp. 431-487.

Griffing, B. 1956. Concept of general and specific combining ability in relation to diallel crossing systems. *Australian Journal of Biological Sciences,* 9, pp. 463-493.

Guddadamath, S., Mohankumar, H.D. and Salimath, P.M. 2011. Genetic analysis of segregating populations for yield in okra [*Abelmoschus esculentus* (L.) Moench]. *Karnataka Journal of Agricultural Sceence*, 24 (2), pp. 114-117.

Gul, S., Khan, N. U., Batool, S., Baloch, M. J., Munir, M., Sajid, M., Kakwoni, A. A., Ghaloo, S. H., Soomro, Z. A. and Kazni, S. F. 2014. Genotype by environment interaction and association of morphological yield variables in upland cotton. *The Journal of Animal and Plant Sciences*, 24 (1), pp. 262-271.

Hancock, W. G. and Lewis, R. S. 2017. Heterosis transmission genetics and selection for increased growth rate in *N. tabacum* × synthetic tobacco cross. *Molecular Breeding*, 37 (4), pp. 53.

Hallauer, A. R., Carena, M. J. and Filho, J. B. M. 2010. *Quantitative genetics in maize breeding*. 6th Ed. Springer, Iowa, USA. 664 pp.

Harding, S. and Payne, R. 2011. A guide to multivariate analysis in GenStat. VSN International, pp. 9-39.

Hassanpanah, D. 2011. Analysis of G × E interaction using the Additive Main Effects and Multiplicative Interaction (AMMI) in potato cultivars. *African Journal of Biotechnology*, 10(2), pp.154-158.

Haun, J. R. 1982. Early prediction of corn yields from daily weather data and single predetermined seasonal constants. *Agricultural Meteorology*, 27 (3), pp. 191-207.

Hayman, B. I. 1954a. The analysis of variance of diallel tables. *Biometrics*, 10, pp. 235-244.

Hayman, B. I. 1954b. The theory and analysis of diallel crosses. *Genetics*, 39, pp. 789-808.

Heslot, N., Akdemir, D., Sorrells, M. E., Jannik, J. L. 2014. Integrating environmental covariates and crop modelling into the genome selection framework to predict genotype by environment interactions. *Theory and Applied Genetics*, 127(2), pp. 463-480.

Hill, J., Becker, H. C. and Tigerstedt, P. M. A. 1998. Quantitative and *Ecological Aspects of Plant Breeding*. Chapman and Hall. London. UK. 275 pp.

Horner, T. W. and Frey, K. J. 1957. Methods for determining natural areas for oat varietal recommendations. *Agronomy Journal, 49* (6), pp. 313-315.

Hu, T. and Lee, A. H. 2015. Tobacco control and tobacco farming in African countries. *Journal of Public Health Policy*, 36: pp. 41-51.

Jia, J., Chen, X., Liu, H., Zhang, J., Xu, L., Zhu, D., Zhang, L. and Liu, Q. 2014. SSR analysis of genetic diversity of Ping'ou hybrid hazelnut. *China Agricultural Science Bulletin,* 30, pp. 25-28.

Jinks, J. L., and Hayman, B. I. 1953. The analysis of diallel crosses. *Maize Genetics Coop. Newsletter,* 27, pp. 48-54.

Johnson, M. T. 2007. Genotype-by-environment interactions leads to variable selection on life-history strategy in Common Evening Primrose (*Oenothera biennis*). *Journal of Evolutionary Biology*, 20, pp. 190-200.

Johnson, H.W., Robinson, H.F. and Comstock, R.E., 1955. Estimates of genetic and environmental variability in soybeans 1. *Agronomy journal*, *47*(7), pp.314-318.

Jonah, P. M., Bello, L. L., Lucky, O., Midau, A. and Moruppa, S. M., 2011. Review : The Importance of Molecular Markers in Plant breeding programmes. *Global Journal of Science Frontier Research,* 11, pp. 4-12.

Joshi, R. K. and Nayak, S. 2010. Gene Pyramiding-A Broad Spectrum Technique for Developing Durable Stress Resistance in Crops. *Biotechnology and Molecular Biology Review*, 5(3), pp. 51-60.

Joshi, S. P., Prabhakar, K., Ranjekar, P. K. and Gupta, V. S. 2011. Molecular Markers in Plant Genome Analysis. http/www.ias.ac.in/currsciju25/articles15.htm, pp. 1-19.

Julio, E., Denoyes, B., Verrier, J. L. and De Borne, F. D. 2006. Detection of QTLs linked to leaf and smoke properties in *Nicotiana tabacum* based on a study of 114 recombinant inbred lines. *Molecular Breeding*, 18 (1), pp. 69-91.

Karimizadeh, R., Mohammadi, M., Sabaghni, N., Mahmoodi, A. A., Roustami, B., Seyyedi, F., Akbari, F. 2013. GGE biplot analysis of yield stability in multi-environment trials of lentil genotypes under rainfed condition. *Notulae Scientia Biologicae*, 5(2), pp. 256-262.

Kearsey, M. J. and Jinks J. L. 1968. A general method of detecting additive, dominance and epistatic variation for metrical traits: I. Theory. *Heredity,* 23, pp. 403-409.

Kempton, R. A. 1984. The Use of Biplots in Interpreting Variety by Environment Interactions. *Journal of Agricultural Science* 103 (1), pp. 123-135.

Khalid, M., Farhatullah, Khan M. Y., Ali, A., Ali, N. 2010. Assessment of inheritance pattern of the two morphological markers in intraspecific hybridization of *Brassica napus* L. in Intraspecific Hybridization of *Brassica Napus* L. *Sarhad Journal of Agriculture*, 26 (3), pp. 369-371.

Khan, S. A., Ahmad, H., Khan, A., Saeed, M., Khan, S. M. and Ahmad B. 2009. Using line × tester analysis for earliness and plant height traits in sunflower (*Helianthus annus L.*). *Recent Research in Science and Technology,* 1, pp. 202-206.

Korubin – Aleksoska, A., Miceska, G., Gveroska, B., Dimitrieski, M. and Aleksoski, J. 2012. Stability of the yield in commercial tobacco varieties in Republic of Macedonia. *Turkish Journal of Agricultural and Natural Sciences. Special Issue, 2*. pp. 1391-1395.

Kroonenberg, P. M. 1995. Introduction to biplots for G ×E tables. Department of Mathematics, *Research Report 51*. University of Queensland, Australia. [Online] Available: http://three-mode.leidenuniv.nl/document/biplot.pdf.

Lahiji, H. S., Golfazani, M. M., Deylami, M. S. and Aalami, A. 2013. Assessing the genetic diversity of 89 flue-cured tobacco varieties using morphological traits and Inter-Simple Sequence Repeat markers. *Crop Breeding Journal* 3(2), pp. 79-85.

Lan, Q., Li, O., Zhang, J., Zhao, X., Zhu, Z., Chen, R., Xu, S., Wang, Y. and Guo Y. 2012. Construction of SSR-based molecular fingerprinting and analysis of genetic diversity for celery varieties from Tianjin. *Tianjin Agricultural Science,* 18, pp. 7-11.

Leitch, I. J., Hanson L., Lim K.Y., Kovarik A., Chase M.W., Clarkson J. J. and Leitch A. R. 2008. The ups and downs of genome size evolution in polyploid species of *Nicotiana* (Solanaceae). *Annals of Botany,* 101: 805-814.

Lewis, D. 1954. Gene-environment interaction: A relationship between dominance, heterosis, phenotypic stability and variability. *Heredity,* 8, pp. 333-356.

Linn, C. S. and Binns, M. R. 1988. A superiority measure of cultivar performance for cultivar × location data. *Canadian Journal of Plant Science,* 68, pp. 193-198.

Liu, X. Z., He, H. S., Yang, Y. M. and Zhang, H. Y. 2009. Genetic diversity among flue-cured tobacco cultivars on the basis of AFLP markers. *Czech Journal of Genetics and Plant Breeding,* 5, pp. 155-159.

Lv, J., Qi, J., Shi, Q., Shen, D., Zhang, S., Shao, G., Li, H., Sun, Z., Weng, Y., Shang, Y., Gu, X., Li, X., Zhu, X., Zhang, J., van Treuren, R., van Dooijeweert, W., Zhang, Z., Huang, S. 2012. Genetic diversity and population structure of cucumber (*Cucumis sativus* L.) T. Zhang, ed. *PLoS ONE,* 7(10), p.e46919. Available at: http://dx.plos.org/10.1371/journal.pone.0046919 [Accessed December 28, 2016].

Makumbi, D., Alvarado, G., Crossa, J. and Burgueno, J. 2018. SASHAYDIALL: A SAS Program for Hayman's Diallel Analysis. *Crop Science,* 58, pp. 1605-1615.

Maleki, H. H., Karimzadeh, G., Darvishzadeh, R. and Sarrafi, A. 2011. Correlation and sequential path analysis of some agronomic traits in tobacco. *Austraian Journal of Crop Science,* 5 (12), pp. 1644 - 1648.

Mandel, J. 1971. A new analysis of variance model for non-additive data. *Technometrics,* 13 (1), pp. 1-18.

Matzinger, D. F., Wernsman, E. A. and Ross, H. F. 1971. Diallel crosses among burley varieties of *Nicotiana tabacum* L. in the F_1 and F_2 generations. *Crop Science,* 11: pp. 275–279

Marani, A. and Sachs, Y. 1966. Heterosis and combining ability in a diallel cross among nine varieties of oriental tobacco. *Crop Science,* 6, pp. 19-22.

Maryan, K. E., Samizadeh, H. and Deylami, M. S. 2012. Assessing the genetic diversity of tobacco (*Nicotiana tabacum* L) varieties. *Crop Breeding Journal*, 2 (2), pp. 125-132.

Mather, K. J. and Jinks, J. L 1982. *Biometrical Genetics* - the study of continuous variation (3rd Ed). Chapman & Hall, London. 654 pp.

Masouleh, A. K. 2012. Molecular characterization and similarity relationships among flue-cured tobacco (*Nicotiana tabacum* L.) genotypes using Simple Sequence Repeat markers. *Not Bot Horti Agrobo*, 40, pp. 247-253.

Mbanga, T. 1991. Tobacco: A Century of Gold. ZIL Publications (Pvt) Ltd. 426 pp.

Mendes, F. F., Guimarães, L. J. M., Souza, J. C., Guimarães, P. E. O., Pacheco, C. A. P., Machado, J. R. de A., Meirelles, W. F., Silva, A. R., and da Parentoni, S. N. 2012. Adaptability and stability of maize varieties using mixed model methodology. *Crop Breeding and Applied Biotechnology, Londrina*, 12 (2), pp.111-117.

Miflin, B., 2000. Crop Improvement in the 21st Century. *Journal of experimental botany*, 51(342), pp.1-8.

Mishra, P.K., Ram, R.B. and Kumar, N., 2015. Genetic variability, heritability, and genetic advance in strawberry (Fragaria× ananassa Duch.). *Turkish Journal of Agriculture and Forestry*, *39*(3), pp.451-458.

Misra, R.C., Das, S. and Patnaik, M. C. 2009. AMMI model analysis of stability and adaptability of late duration finger millet (Eleusine coracana) genotypes. *World Applied Sciences Journal*, 6(12), pp. 1650-1654.

Mohamed, N. E. M. 2013. Genotype by environment interactions for yield in bread wheat (*Triticum aestivum* L.). *Journal of Plant Breeding and Crop Science*, 5(7), pp. 150-157.

Mohammadi, R., Haghparast, R., Amri, A., and Ceccarelli, S. 2010. Yield stability of rainfed durum wheat and gge biplot analysis of multi-environment trials. *Crop and Pasture Science*, 61, pp. 92-101.

Moon, H. S., Nicholson, J. S., Heineman, A., Lion K., Hoeven, R., Hayes, A. J., Lewis, R. S. 2009a. Changes in genetic diversity of u.s. flue-cured tobacco germplasm over seven decades of cultivar development. *Crop Science*, 49(2), pp. 498-508.

Moon, H. S., Nifong, J. M., Nicholson, J. S., Heineman, A., Lion, K., Van der Hoeven, R., Hayes, A. J., Lewis, R. S. 2009b. Microsatellite-based analysis of tobacco (*Nicotiana tabacum* L.) genetic resources. *Crop Science*, 49(6), pp.2149-2159.

Moose, S. P. and Mumm, R. H. 2008. Molecular plant breeding as the foundation for 21st century crop improvement. *Plant physiology*, 147(3), pp. 969-977.

Mosa, H. E. 2010. Estimation of combining ability of maize inbred lines using topcross mating design. *Journal of Agricultural Research. Kafer El-Sheikh University* 36: pp. 1-14.

Munafò, M. R., Durrant, C., Lewis, G., Flint, J. 2009. Gene × environment interactions at the serotonin transporter locus. *BPS*, 65(3), pp. 211-219.

Mwololo, J. K., Muturi, P. W., Mburu, M. W. K., Njeru, R. W., Kiarie, N., Munyua, J. K., Ateka, E. M., Muinga, R. W., Kapinga, R. E. and Lemaga, B. 2009. Additive Main Effects and Multiplicative Interaction analysis of genotype x environmental interaction among sweetpotato genotypes, *Journal of Animal and Plant Science*, 2(3), pp. 148-155.

Naik, V. R., Bentur, M. G. and Parameshwarappa, K. G. 2009. Impact of biparental mating on genetic variability and path analysis in safflower. *Karnataka Journal of Agricultural Sciences*, 22(1), pp. 44-46.

Nduwumuremyi, A., Tongoona, P. and Habimana, S. 2013. Mating designs: helpful tool for plant breeding analysis. *Journal of Plant Breeding and Genetics*, 3: pp. 117-129.

Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, 89, pp. 583-590.

Nunome, T., S. Negoro, I., Kono, H., Kanamori, K. and Miyatake, H. 2009. Development of SSR markers derived from SSR-enriched genomic library of eggplant. *Theoretical Applied Genetics*, 119, pp. 1143-1153.

Pandeya, R. S., Dirks, V. A. and Poushinsky, G. 1983. Quantitative genetic studies in flue-cured tobacco (*Nicotiana tabacum*). I. Agronomic characters. *Canadian Journal of Genetics and Cytology*, 25 (4), pp. 336-345.

Parajuli, S., Patel, D. A. and Delvadia, D. R. 2015. Genetic variability, correlation and path analysis for cured leaf yield and its components in Rustica Tobacco (*Nicotiana rustica*

L.). *International Journal of Tropical Agriculture. Serials Publications*, 33(3). pp. 2085-2092.

Patel, J. N., Patel, A. D., Chakraborty, S. and Sasidharan, N. 2012. Heterosis and relative heterosis in tobacco. (*Nicotiana tabaccum* L.). *International Journal of Plant, Animal and Environmental Sciences,* 2 (2), pp. 270-275.

Peakall, R. and Smouse, P.E. 2012. GenAlEx 6.5: Genetic analysis in excel. Population genetic software for teaching and research-an update. *Bioinformatics,* 28, pp. 2537-2539.

Pearson, K. 1901. On lines and planes of closest fit to systems of points in space. *Philos. Mag. Sixth Series,* 2, pp. 559 - 572.

Perrier, X. and Jacquemoud-Collet, J. P. 2006. DARwin software.

Piepho, H. P. 1998. Empirical best linear unbiased prediction in cultivar trials using factor – analytic variance-covariance structures. *Theory and Applied Genetics,* 97, 1-2, pp. 195-201.

Piyasatian, N. and Kinghorn, B. P. 2003. Balancing genetic diversity, genetic merit and population viability in conservation programmes. *Journal Animal Breeding and Genetics,* 120, pp. 137-149.

Plaisted, R. L. and Peterson, L. C. 1959. A technique for evaluating the ability of selections to yield consistently in different locations or seasons. *American Journal of Potato Research,* 36 (11), pp. 381-385.

Pritchard, J. K., Stephens, M., Donnelly, P. J. 2000. Inference of population structure using multilocus genotype data. *Genetics,* 155, pp. 945-955.

Rad, M.N., Kadir, M.A., Rafii, M.Y., Jaafar, H.Z., Naghavi, M.R. and Ahmadi, F., 2013. Genotype environment interaction by AMMI and GGE biplot analysis in three consecutive generations of wheat (Triticum aestivum) under normal and drought stress conditions. *Australian Journal of Crop Science*, *7*(7), p.956.

Raeber J. G. and Schwappenhauser, M. A. 1966. Breeding for resistance to root-knot nematode (*M. javanica*) from *Nicotiana tabacum* L. *Tobacco Research Board, Annual Report*, 1965-66.

Raeber J. G. and Smeeton, B. 1970. Breeding for resistance to Wildfire and TMV. *Tobacco Research Board, Annual Report*, 1969-70.

Rahman, R. M., Rahman, M. A., Hossain, A. and Rasul, G. 2011. Comparative study on morphological , physiological and molecular genetic diversity analysis in rice (*Oryza sativa* L.) using SSR markers. *Journal of Crop Improvement*, 2016(2). pp 244-257.

Raju, K. S., Sheshumadhav, M., Chandrasekhararao, C., Murthy, T. G. K. 2009. Molecular diversity in genus Nicotiana as revealed by randomly amplified polymorphic DNA. *Indian Journal of Biotechnology*, 8 (1). pp. 61-66.

Ramburan, S., Zhou, M., Labuschagne, M. 2011. Interpretation of genotype × environment interactions in sugarcane: Identifying significant environmental factors. *Field Crops Research*, 124 (3), pp. 392-399.

Ren, N. and Timko, M. P. 2001. AFLP analysis of genetic polymorphism and evolutionary relationships among cultivated and wild *Nicotiana* species. *Genome*, 441, pp. 559-571.

Resende, M. D. V. and Thompson, R. 2004. Factor analytic multiplicative mixed models in the analysis of multiple experiments. *Rev. Mat. Estat*., 22 (2), pp. 31-52.

Robinson, H. F., Mann, T. J. and Comstock, C. C. 1954. An analysis of quantitative variability in *Nicotaana tabacum*. *Heredity*, 8, pp. 365-376.

Röder, M. S., Korzun, V., Wendehake, K., Plaschke, J., Tixier, M. H., Leroy, P. and Ganal, M. W. 1998. A microsatellite map of wheat. *Genetics*, 149, pp. 2007-2023.

Rosielle, A. A. and Hamblin. J. 1981. Theoretical aspects of selection for yield in stress and non-stress environments. *Crop Science*, 21, pp. 943-946.

Rossi, L., Bindler, G., Pijnenburg, H., Isaac, P.G., Giraud-Henry, I., Mahe, M., Orvain, C., Gadani F. 2001. Potential of molecular marker analysis for variety identification in processed tobacco. *Plant Variety Seeds* 14, pp. 1863-1869.

Rukuni, M. and Eicher, C. K. 2006. *Zimbabwe's Agricultural Revolution Revisited*. University of Zimbabwe Publications Office. Harare. Zimbabwe. 728 pp.

Sabaghnia, N., Karimizadeh, R. and Mohammadi, M. 2012. Genotype by environment interaction and stability analysis for grain yield of lentil genotypes. Žemdirbystė. *Agriculture*, 99 (3), pp. 305-312.

Sabu, K. K., Abdullah, M. Z., Lim, L. S. and Wickneswari, R. 2009. Analysis of heritability and genetic variability of agronomically important traits in *Oryza sativa* x *O. rufipogon* cross. *Agronomy Research*, 7(1), pp. 97-102.

Saddoud, O., Baraket, G., Chatti, K., Trifi, M., Marrakchi, M., Mars, M., Salhi-Hannachi, A. 2012. Using Morphological characters and Simple Sequence Repeat (SSR) markers to characterize Tunisian fig (*Ficus Carica* L.) cultivars. *Acta Biologica Cracoviensia Series Botanica*, 53(2), pp. 7-14.

Sadeghi, S. M., Amiri, E. and Ashouri, M. 2012. An investigation of gene action on different traits of tobacco under irrigated and drought stress environment. *African Journal of Biotechnology* 11 (21), pp. 4740-4551.

Sajjad, M. and Khan, F. A. 2009. Genetic diversity among sugarcane cultivars in Pakistan. *American-Eurasian Journal of Agricultural & Environmental Science*, 6 (6), pp. 730-736.

Salmon, S. C. 1951. Analysis of variance and long-time variety tests of wheat. *Agronomy Journal*, 43 (11), pp. 562-570.

Sandison, A. and Bartlett, B. O. 1958. Comparison of varieties for yield. *Journal of National. Institute of Agricultural Botany,* 8, pp. 351-357.

Saroj, S. K., Singh, M. N. and Singh, T. 2014. Combining ability, heterosis and inbreeding depression analysis for using CMS lines in long duration pigeonpea. *International Journal of Research Studies in Biosciences*, 2(10), pp.7-15.

Shah, K. A., Shah, F. L., Ali, A., Ahmed, Q. and Zhou, L. 2015. Genetic variability and heritability studies for leaf and quality characters in flue-cured Virginia tobacco. *Academic Journal of Agricultural Research*, 3(3), pp. 044-048.

Seyyed-Nazari, R., Ghadimzadeh, M., Darvishzadeh, R. and Alavi, S. R. 2016. Diallel analysis for estimation of genetic parameters in oriental tobacco genotypes. *Genetika*, 48 (1), pp. 125-137.

Sharma, J. R. 2006. *Statistical and biometrical techniques in plant breeding*. 1st Ed. New Age International New Delhi. 432 pp.

Shuelke, M. 2000. An economic method for the fluorescent labeling of PCR. *Nature Biotechnology*, 18, pp. 233-234.

Singh, B. D. 2006. *Plant Breeding: Principles and methods*. Kalyani Publishers. India. 268 pp.

Singh, R. K. and Chaudhary, B. D. 2004. Biometrical *methods in quantitative genetic analysis*. Kalyani Publishers. New Delhi. India. 318 pp.

Singh, R. K. and Chaudhary, R. D. 1979. *Biometrical methods in quantitative genetic analysis.* New Delhi: Kalyani Publications. 304 pp.

Slana L. J., Stavely, J. R., Grosso, J. J. and Golden, A. M. 1976. Probable sources of *Meloidogyne incognita* resistance in tobacco as indicated by reactions to five *Meloidogyne* isolates. *Phytopathology,* 67, pp. 537-543.

Sprague, G. F. and Federer, W. T. 1951. A comparison of variance components in corn yield trials. ii. error, year × variety, location × variety, and variety components. *Agronomy Journal*, 43 (11), pp. 535-541.

Storck, L., Benin, G., Marchioro, V. S., Silva, R. R., Woyann, L. G. and Bornhofen, E. 2016. Strategy for grouping wheat genotypes according to environmental responses in multi-location trials. *Australian Journal of Crop Science*, 10(4), pp. 570-578.

Tanser, G. H. 1991. A *History of the Tobacco Research Board "The First 25 Years"*. Zimbabwe Agricultural Journal Technical Handbook No. 7. Modern Farming Publications. Harare. Zimbabwe.

Thakur, M. C., Agrawal, S., Patel, M., Khan, A. and Doshi, H. V. 2013. Genetic variability of *Nicotiana tabacum* (Linn.) using SSR marker. *World Journal of Agricultural Research*, 1 (6): 124-129.

Thermo Scientific, Pittsburgh, PA, USA. (https://www.thermalscientific.com/)

Thottapilly, G, Magonouna, H. D. and Omitogun, O. G. 2009. The Use of DNA Markers for Rapid Improvement of Crops in Africa. *African Crop Science Journal*, 8(1), pp. 99-108.

Tobacco Industry and Marketing Board 1999. The One industry approach to grading and presentation of flue – cured tobacco. Harare. Zimbabwe.

Tobacco Research Board Annual Report. 1961. Kutsaga Publications. Harare. Zimbabwe.

Tobacco Research Board Annual Report. 1962. Kutsaga Publications. Harare. Zimbabwe

Tobacco Research Board Annual Report. 1996. Kutsaga Publications. Harare. Zimbabwe.

Tobacco Research Board. 2012. *Flue-cured Tobacco Recommendations*. Kutsaga Publications. Harare. Zimbabwe.

Tobacco Research Board Annual Report. 2002. Kutsaga Publications. Harare. Zimbabwe.

Tobacco Research Board Annual Report. 2002. Kutsaga Publications. Harare. Zimbabwe.

Tobacco Research Board Annual Report. 2010. Kutsaga Publications. Harare. Zimbabwe.

Tobacco Research Board Annual Report. 1913. Kutsaga Publications. Harare. Zimbabwe.

Tobacco Research Board Annual Report. 2015. Kutsaga Publications. Harare. Zimbabwe.

Tong, Z., Xiao, B., Jiao, F., Fang, D., Zeng, J., Wu, X. and Chen, X. 2016. Large-scale development of SSR markers in tobacco and construction of a linkage. *Breeding Science,* pp. 381-390.

Ulukan, H. 2011. Plant genetic resources and breeding: current scenario and future prospects. *ISSN Online International Journal of Agricultural Biology*, 13, pp.1560-8530.

UPOV, T., 2002. 1/3 General Introduction to the Examination of Distinctness, Uniformity and Stability and the Development of Harmonized Descriptions of New Varieties of Plants.

Van Eeuwijk, F. A., Denis, J. B., Kang, M. S. 1996. Incorporating additional genotypes fnd environments in models for two-way genotype by environment tables. In: Kang, M. S. and Gauch, H, G, (eds). genotype-by-environment interaction. Boca Raton: CRC Press, pp 15-49.

Vargas, M, van Euwijk, F. A., Crossa, J., Ribaut, J. M. 2006. Mapping QTLs and QTL × environment interaction for CIMMYT maize drought stress programme using factorial regression and partial least squares methods. *Theoretical and Applied Genetics*, 112 (6), pp. 1009-1023.

Vencovsky, R., Ramalho M. A. P., Toledo F. H. R. B. 2012. Contribution and perspectives of quantitative genetics to plant breeding in Brazil. *Crop Breeding and Applied Biotechnoogy,* 2, pp. 7-14.

VSN International, 2017. Genstat *for Windows* 19th Edition. VSN International, Hemel Hempstead, UK. Web page: Genstat.co.uk

Wolf, J. B. and Wade, M. J. 2009. What are maternal effects (and what are they not)? Philosophical Transactions of the Royal Society of London. *Series B Biological Science,* 364**,** pp. 1107-1115.

Xiu-ping, L., Bing-guang, X., Yong-ping, L., Yi-jie, G., Yu, W. and Long-jiang, F. 2013. Diversity Arrays Technology (Dart) for studying the genetic polymorphism of flue-cured tobacco (*Nicotiana tabacum*). *Journal of Zhejiang University-Science B (Biomedicine & Biotechnology)*, 14 (7), pp. 570-577.

Yan, W. 2001. GGE biplot – a Windows application for graphical analysis of multi-environment trial data and other types of two-way data. *Agronomy Journal*, 93, pp. 1111-1118.

Yan, W. 2002. Singular value partitioning for biplot analysis of multi-environment trial data. *Agronomy Journal*, 94, pp. 990-996.

Yan, W., Cornelius, P. L., Crossa, J. and Hunt, L. A. 2001. Two types of GGE biplots for analyzing multi-environment trial data. *Crop Science*, 41, pp. 656-663.

Yan, W. and Hunt L. A. 2001. Interpretation of genotype x environment interaction for winter wheat yield in Ontario. *Crop Science*, 41, pp. 19-25.

Yan, W. and Hunt, L. A. 2002. Biplot analysis of diallel data. *Crop Science*, 42, pp. 21-30.

Yan, W, Hunt, L A, Sheng, Q, Szlavnics, Z. 2000. Cultivar evaluation and mega-environment investigation based on the GGE biplot. *Crop Science*, 40 (3), pp. 597-605.

Yan, W. and Kang, M. S. 2003. *GGE Biplot Analysis: A Graphical Tool for Breeders, Geneticists and Agronomists.* CRC Press, Boca Raton, FL. 288 pp.

Yan, W., Kang, M.S., Ma, B., Woods, S. and Cornelius, P.L., 2007. GGE biplot vs. AMMI analysis of genotype-by-environment data. *Crop science*, *47*(2), pp.643-653.

Yan, W. and Tinker, N. A. 2005a. An integrated biplot analysis system for displaying, interpreting, and exploring genotype-byenvironment interactions. *Crop Science*, 45, pp. 1004-1016.

Yan, W. and Tinker, N. A. 2005b. A biplot approach to the investigation of QTL-by-environment patterns. *Molecular Breeding*, 15, pp. 31-43.

Yan, W. and Tinker, N. A. 2006. Biplot analysis of multi-environment trial data: Principles and applications. *Canadian Journal of Plant Science*, 86, pp. 623-645.

Ya-Nan, M. *et al*. 2012. Developmental analysis on genetic behaviour of quality traits of flue-cured tobacco (*Nicotiana tabacum* L) in multiple environments. *Journal of Agriculture and Biology*, 14, pp 345-352.

Yang, B. C., Xiao, B. G., Chen, X. J., Shi, C. H. 2007. Assessing the genetic diversity of tobacco germplasm using Intersimple Sequence Repeat and Inter-Retrotransposon. *Annals of Applied Biology*, 150, pp. 393-401.

Yao, J. B., Ma, H.X. Ren, L. J., Zhang, P. P., Yang, X. M., Yao, G. C., Zhang, P. and Zhou, M. P. 2011. Genetic analysis of plant height and its components in diallel crosses of bread wheat (*Triticum aestivum* L.). *Australian Journal of Crop Science*, 5(11), pp. 1408-1418.

Yates, F. and Cochran, W. G. 1938. The analysis of groups of experiments. *Journal of Agricultural Science*, 28(4), pp. 556-580.

Ye, L., Xin, M., Du, J., Xie, C., Li, Y. and Zhou, Q. 2009. Genetic diversity revealed by SSR markers of the tobacco cultivars. *China Agricultural Science Bulletin*, 25, pp. 56-62.

Yong-Bi, F. 2015. Understanding crop genetic diversity under modern plant breeding. *Theoretical and Applied Genetics*, 128, pp. 2131-2142.

Zdravković, J., Pavlović, N., Girek, Z., Brdar-Jokanović, M., Savić, D., Zdravković, M. and Cvikić, D. 2011. Generation mean analysis of yield components and yield in tomato, *Pak. J. Bot,* 43(3), pp.1575-1580.

Zhang, H. Y., Yang, Y. M., Li, F. S, He, C. S. and Liu, X. Z. 2008. Screening and characterization of a RAPD marker of tobacco brown-spot resistant gene. *Journal of Biotechnology*, 7(15), pp. 2559-2561.

Zobel, R. W., Wright, M. J., Gauch, H. G. 1988. Statistical analysis of a yield trial. *Agronomy Journal,* 80(3), pp. 388-393.

Zobel, R.W. 1997. Identifying mega-environments and targeting genotypes, *Crop Science*, 37, pp. 311-326.

APPENDICES

Appendix 1: List of entries in the morphological diversity study

Entry No.	Entry Name	Entry No.	Entry Name
1	Station Bonanza	99	Speight g28
2	Jamaica Wrapper	100	TI 764
3	Willow Leaf	101	TI 799
4	Yellow Mammoth	102	TI 859
5	Gold Dollar	103	TI 933
6	Proctor's Special	104	TI 1372
7	Duquesne Johnson	105	Kutsaga 51E
8	Harrison's Special	106	Kutsaga Mammoth E
9	Delcrest	107	V572-MU
10	American 401	108	PY 10
11	American 402	109	LAFC 53
12	Coker's Certified 401	110	A56N (1-1-6)
13	Coker's Certified 402	111	Tl 33
14	Meadows Giant	112	Tl 38
15	Oxford 26	113	RKT (15-1-1)
16	Virginia Scheveiz	114	RKL (115-1-1)
17	American Joiner	115	RKR (143-1-1)
18	White Mammoth Norfolk	116	RKE (26-5-2-1-1)
19	Broad leaf	117	Kutsaga 110
20	Daines bright leaf	118	DETW 144
21	C7	119	Coker 347
22	Virginia Gold	120	Coker 411
23	Bottom special	121	Mcnair
24	Mammoth Gold	122	NC 13
25	Vamor 50	123	NC 79

Entry No.	Entry Name	Entry No.	Entry Name
26	Dixie Bright leaf 102	124	NC 89
27	Canadel	125	NC 5613
28	Hicks broad leaf	126	TB 22
29	Harrison’s special	127	Coker 48
30	Special 400	128	Speight g23
31	Oxford 26	129	WM (84-20-2)
32	Golden harvest	130	AE (48-18-5)
33	Oxford 402	131	TB 4
34	White stem Orinoco	132	PY (1-19-1)
35	Yellow Prior	133	G 27
36	Yellow special A	134	ETW 22
37	Silver dollar	135	Cash
38	Ruffle 1	136	ETW 124
39	Ruffle 3	137	W-1-3-1
40	Ruffle 4	138	KM 10
41	Golden cure	139	RL (1-2-1)
42	White Gold	140	RL (2-1-1)
43	Carolina bright	141	ST (1-2-8)
44	Virginia bright leaf	142	ST (3-2-6)
45	Golden wilt	143	TI (13-20
46	Yellow Gold	144	Speight g70
47	Delhi	145	Amaralinho
48	MK 94	146	Daetw 23
49	SCR	147	Ob 57
50	Golden harvest	148	ZZ 100
51	402	149	NC 744
52	Hickory Prior	150	New del
53	Mammoth Delcrest	151	Nordel
54	Warne	152	Delgold
55	Coker 187	153	Haploid no.2
56	Oxford 1	154	Jing chang

Entry No.	Entry Name	Entry No.	Entry Name
57	Oxford 1-181	155	C 319
58	Virginia 21	156	K326
59	SC 58	157	TI 1112
60	Chicoana	158	Ob 56
61	PD 611	159	AW(22-10-1-1-65)
62	NC 95	160	AW(22-10-1-1-3r)
63	Florida 22	161	Bel 921
64	Bel 61-10	162	STNCB (ss9-48-14-23)
65	Rhomos 7	163	KB 14/a
66	Kutsaga 51	164	STNCA (1-10-10-7-28-4)
67	Yellow Gold	165	STNCB (2-45-16-14-22-2)
68	Delhi 61	166	DSD (5-4-2r)
69	MK94	167	SR 1
70	SCR	168	BS40-1
71	Vinica	169	Victoria selection
72	Resistant Hicks	170	Queensland selection
73	Bel 61 -10	171	Coker 371 gold
74	Bel 61 – 12	172	A4
75	Kutsaga Mammoth	173	W2 (1-7-1-1)
76	WK35	174	NOD 7
77	WK 39	175	NC BMR 42
78	Coker 319	176	Flue-cured Kentucky 14
79	NC 2326	177	NC 27
80	NC 2512	178	NC 37
81	Speight G 10	179	MTRA 65
82	Speight G36	180	NK 149
83	Virginia 312	181	CR 401
84	Enshu	182	Candel
85	Kutsaga E1B	183	Deliot
86	Kutsaga 51B	184	Wanda

Entry No.	Entry Name	Entry No.	Entry Name
87	Kutsaga E1	185	Delfield
88	Kutsaga E2	186	Wisana
89	Delcrest 66	187	NK 346
90	Coker 254	188	ZAS 3-2-1-1-1-1
91	Hicks	189	NC 628
92	Delcrest 57	190	WA 169-7 d.h.
93	Mcnair	191	XA 219-5 D.H.
94	Kutsaga MB2	192	H170
95	PD121	193	F111
96	TI 1138	194	Dindigal cheroot
97	TI 804	195	CSIRO 40T
98	TI 814	196	SIRONE

Appendix 2: Similarity matrix group 1

A4	1	----								
Coker 319	2	68.6	----							
COKER 371 GOLD	3	54.3	65.7	----						
COKER 411	4	62.9	68.6	62.9	----					
DELGOLD	5	80.0	62.9	48.6	65.7	----				
F111	6	62.9	57.1	57.1	54.3	65.7	----			
Florida 22	7	60.0	60.0	62.9	57.1	57.1	60.0	----		
KUTSAGA 110	8	65.7	57.1	57.1	40.0	57.1	82.9	65.7	----	
LAFC 53	9	54.3	48.6	51.4	71.4	62.9	57.1	62.9	51.4	----
NC 27	10	60.0	51.4	71.4	62.9	62.9	54.3	62.9	54.3	62.9

NK 149	11	54.3	57.1	57.1	60.0	71.4	77.1	57.1	62.9	51.4
60.0										
OB 57	12	60.0	57.1	51.4	48.6	60.0	54.3	65.7	54.3	57.1
68.6										
PD 611	13	57.1	45.7	51.4	62.9	51.4	65.7	54.3	60.0	65.7
57.1										
TL 33	14	57.1	54.3	62.9	60.0	57.1	57.1	54.3	62.9	65.7
57.1										
W-1-3-1	15	60.0	68.6	60.0	54.3	62.9	48.6	60.0	54.3	51.4
65.7										

		1	2	3	4	5	6	7	8	9
10										
NK 149	11	----								
OB 57	12	54.3	----							
PD 611	13	48.6	45.7	----						
TL 33	14	62.9	42.9	54.3	----					
W-1-3-1	15	54.3	77.1	42.9	42.9	----				

	11	12	13	14	15

Appendix 3: Similarity matrix group 2

402	1	----							
Bel 61-10	2	68.6	----						
BEL 921	3	74.3	62.9	----					
Bottom special	4	68.6	74.3	57.1	----				
Cokers certified 402	5	71.4	71.4	62.9	71.4	----			
Cokers Certified 402	6	65.7	57.1	62.9	57.1	62.9	----		
CR 401	7	60.0	62.9	57.1	68.6	57.1	65.7	----	
Dixie Bright leaf 102	8	77.1	68.6	71.4	65.7	74.3	65.7	60.0	----
ETW 124	9	74.3	68.6	68.6	57.1	60.0	71.4	54.3	62.9

G 27	10	77.1	65.7	71.4	65.7	60.0	68.6	77.1	71.4
68.6									
Golden harvest	11	74.3	57.1	65.7	57.1	65.7	74.3	62.9	65.7
71.4									
Golden wilt	12	77.1	71.4	62.9	68.6	74.3	77.1	62.9	74.3
71.4									
HICKS	13	71.4	60.0	71.4	71.4	65.7	65.7	62.9	62.9
60.0									
Kutsaga 51	14	71.4	74.3	54.3	71.4	71.4	71.4	71.4	68.6
62.9									
KUTSAGA 51B	15	62.9	62.9	57.1	51.4	57.1	51.4	51.4	54.3
68.6									
Kutsaga E2	16	62.9	65.7	62.9	71.4	71.4	65.7	65.7	68.6
62.9									
Kutsaga mammoth	17	88.6	71.4	85.7	68.6	74.3	68.6	62.9	82.9
74.3									
MCNAIR	18	74.3	68.6	65.7	74.3	71.4	65.7	71.4	74.3
68.6									
MK 94	19	65.7	77.1	62.9	68.6	68.6	71.4	68.6	68.6
65.7									
NC 13	20	77.1	77.1	65.7	71.4	68.6	62.9	65.7	74.3
68.6									
NC 628	21	57.1	65.7	60.0	68.6	62.9	57.1	62.9	68.6
62.9									
NC 744	22	74.3	68.6	62.9	60.0	65.7	60.0	65.7	71.4
65.7									
NC 79	23	74.3	74.3	77.1	71.4	71.4	65.7	74.3	77.1
62.9									
NEW DEL	24	85.7	74.3	74.3	68.6	68.6	62.9	60.0	77.1
77.1									
NOD 7	25	65.7	71.4	60.0	68.6	57.1	57.1	71.4	62.9
60.0									
NORDEL	26	65.7	62.9	54.3	74.3	62.9	54.3	51.4	60.0
60.0									
Oxford 1	27	68.6	60.0	68.6	77.1	62.9	57.1	65.7	71.4
51.4									
Oxford 26	28	62.9	57.1	54.3	71.4	60.0	62.9	77.1	68.6
60.0									
Rhomos 7	29	74.3	74.3	57.1	68.6	71.4	65.7	74.3	65.7
65.7									
Rhomos 7	30	74.3	65.7	71.4	60.0	68.6	60.0	71.4	74.3
65.7									
RKT(15-1-1)	31	60.0	62.9	65.7	71.4	60.0	57.1	57.1	60.0
54.3									
Ruffle 3	32	62.9	62.9	57.1	68.6	65.7	45.7	57.1	68.6
57.1									
Ruffle 4	33	57.1	57.1	62.9	60.0	62.9	54.3	48.6	65.7
57.1									
Speight G36	34	71.4	74.3	74.3	71.4	74.3	65.7	57.1	80.0
62.9									
ST(1-2-8)	35	68.6	68.6	62.9	77.1	74.3	65.7	65.7	68.6
71.4									
ST(3-2-6)	36	60.0	62.9	57.1	65.7	65.7	42.9	51.4	48.6

62.9
STNCA(1-10-10-7-28-4) 37 65.7 60.0 57.1 65.7 68.6 62.9 82.9 65.7
54.3
TL (13-20 38 60.0 65.7 45.7 77.1 71.4 45.7 65.7 60.0
42.9
V572-MU 39 71.4 60.0 62.9 60.0 65.7 82.9 60.0 71.4
71.4
Vamor 50 40 71.4 60.0 68.6 62.9 68.6 65.7 54.3 62.9
60.0
Virginia 21 41 77.1 62.9 68.6 60.0 62.9 62.9 57.1 65.7
82.9
WA 169-7 d.h. 42 71.4 60.0 74.3 60.0 60.0 60.0 54.3 62.9
71.4
Wanda 43 74.3 57.1 68.6 54.3 65.7 62.9 51.4 68.6
68.6
White gold 44 65.7 65.7 65.7 65.7 62.9 62.9 54.3 54.3
71.4
Willow leaf 45 77.1 68.6 71.4 62.9 71.4 77.1 62.9 71.4
77.1
WK 39 46 62.9 62.9 57.1 80.0 65.7 57.1 68.6 62.9
48.6
Yellow gold 47 65.7 62.9 57.1 74.3 71.4 65.7 60.0 65.7
62.9

1 2 3 4 5 6 7 8
9

G 27 10 ----
Golden harvest 11 65.7 ----
Golden wilt 12 74.3 71.4 ----
HICKS 13 65.7 62.9 60.0 ----
Kutsaga 51 14 71.4 74.3 85.7 62.9 ----
KUTSAGA 51B 15 65.7 60.0 65.7 51.4 74.3 ----
Kutsaga E2 16 62.9 62.9 60.0 74.3 65.7 54.3 ----
Kutsaga mammoth 17 80.0 71.4 74.3 80.0 65.7 60.0 71.4 ----
MCNAIR 18 80.0 62.9 80.0 60.0 74.3 62.9 62.9 74.3

MK 94 19 71.4 68.6 68.6 65.7 74.3 60.0 68.6 74.3
62.9
NC 13 20 74.3 74.3 77.1 65.7 74.3 62.9 68.6 77.1
65.7
NC 628 21 65.7 48.6 71.4 60.0 62.9 57.1 68.6 68.6
71.4
NC 744 22 82.9 57.1 68.6 62.9 71.4 74.3 65.7 71.4
68.6
NC 79 23 88.6 71.4 77.1 71.4 74.3 60.0 68.6 85.7
74.3
NEW DEL 24 77.1 74.3 71.4 68.6 74.3 77.1 71.4 82.9
68.6
NOD 7 25 71.4 60.0 62.9 54.3 65.7 57.1 57.1 68.6
65.7
NORDEL 26 57.1 51.4 54.3 60.0 60.0 54.3 68.6 62.9
62.9

Oxford 1	27	71.4	51.4	57.1	80.0	60.0	48.6	65.7	74.3	68.6
Oxford 26	28	68.6	57.1	60.0	65.7	68.6	54.3	74.3	65.7	71.4
Rhomos 7	29	74.3	68.6	71.4	62.9	77.1	60.0	71.4	68.6	71.4
Rhomos 7	30	80.0	65.7	71.4	68.6	71.4	62.9	65.7	80.0	77.1
RKT(15-1-1)	31	65.7	54.3	68.6	74.3	62.9	54.3	57.1	65.7	60.0
Ruffle 3	32	48.6	51.4	62.9	65.7	62.9	57.1	71.4	62.9	68.6
Ruffle 4	33	48.6	51.4	60.0	60.0	54.3	54.3	62.9	65.7	57.1
Speight G36	34	65.7	62.9	71.4	68.6	68.6	54.3	71.4	77.1	68.6
ST(1-2-8)	35	60.0	65.7	74.3	60.0	71.4	54.3	68.6	68.6	77.1
ST(3-2-6)	36	51.4	51.4	60.0	68.6	60.0	60.0	62.9	60.0	54.3
STNCA(1-10-10-7-28-4)	37	71.4	62.9	74.3	65.7	77.1	54.3	65.7	62.9	71.4
TL (13-20	38	62.9	51.4	65.7	62.9	68.6	48.6	60.0	54.3	65.7
V572-MU	39	68.6	74.3	68.6	65.7	68.6	48.6	68.6	74.3	57.1
Vamor 50	40	62.9	85.7	65.7	71.4	74.3	65.7	62.9	68.6	62.9
Virginia 21	41	77.1	65.7	71.4	65.7	68.6	77.1	60.0	80.0	74.3
WA 169-7 d.h.	42	62.9	62.9	62.9	62.9	60.0	71.4	60.0	77.1	65.7
Wanda	43	71.4	60.0	80.0	54.3	65.7	68.6	45.7	68.6	71.4
White gold	44	62.9	65.7	62.9	71.4	60.0	57.1	74.3	68.6	51.4
Willow leaf	45	71.4	82.9	82.9	65.7	74.3	65.7	68.6	74.3	74.3
WK 39	46	62.9	51.4	62.9	65.7	65.7	51.4	68.6	62.9	71.4
Yellow gold	47	60.0	65.7	71.4	65.7	68.6	65.7	62.9	68.6	74.3

		10	11	12	13	14	15	16	17	18
MK 94	19	----								
NC 13	20	80.0	----							
NC 628	21	62.9	62.9	----						
NC 744	22	68.6	68.6	65.7	----					
NC 79	23	82.9	82.9	65.7	71.4	----				
NEW DEL	24	68.6	80.0	57.1	77.1	77.1	----			
NOD 7	25	74.3	65.7	74.3	68.6	71.4	65.7	----		

NORDEL	26	54.3	54.3	62.9	62.9	51.4	65.7	54.3	----
Oxford 1	27	65.7	68.6	65.7	71.4	71.4	65.7	65.7	60.0

Oxford 26	28	62.9	60.0	68.6	68.6	62.9	65.7	65.7	60.0
80.0									
Rhomos 7	29	74.3	80.0	57.1	71.4	71.4	71.4	65.7	60.0
62.9									
Rhomos 7	30	71.4	68.6	77.1	74.3	82.9	68.6	68.6	57.1
71.4									
RKT(15-1-1)	31	62.9	74.3	57.1	54.3	74.3	62.9	51.4	45.7
71.4									
Ruffle 3	32	51.4	62.9	71.4	57.1	54.3	68.6	60.0	60.0
65.7									
Ruffle 4	33	51.4	60.0	57.1	48.6	57.1	57.1	40.0	62.9
57.1									
Speight G36	34	65.7	71.4	60.0	65.7	77.1	77.1	54.3	65.7
71.4									
ST(1-2-8)	35	68.6	65.7	71.4	60.0	65.7	60.0	62.9	65.7
65.7									
ST(3-2-6)	36	57.1	60.0	62.9	60.0	60.0	57.1	51.4	60.0
62.9									
STNCA(1-10-10-7-28-4)	37	62.9	60.0	68.6	68.6	71.4	60.0	65.7	54.3
62.9									
TL (13-20	38	62.9	68.6	57.1	65.7	68.6	57.1	57.1	57.1
68.6									
V572-MU	39	82.9	68.6	57.1	65.7	71.4	68.6	60.0	60.0
62.9									
Vamor 50	40	65.7	65.7	51.4	60.0	68.6	77.1	60.0	57.1
60.0									
Virginia 21	41	60.0	68.6	60.0	68.6	71.4	80.0	57.1	54.3
60.0									
WA 169-7 d.h.	42	62.9	71.4	54.3	57.1	68.6	77.1	57.1	51.4
62.9									
Wanda	43	57.1	68.6	65.7	77.1	65.7	65.7	60.0	51.4
65.7									
White gold	44	65.7	74.3	62.9	62.9	68.6	71.4	57.1	65.7
57.1									
Willow leaf	45	60.0	68.6	60.0	68.6	71.4	80.0	54.3	60.0
51.4									
WK 39	46	57.1	65.7	68.6	60.0	62.9	62.9	68.6	62.9
80.0									
Yellow gold	47	60.0	65.7	71.4	57.1	60.0	68.6	60.0	62.9
65.7									

		19	20	21	22	23	24	25	26
27									
Oxford 26	28	----							
Rhomos 7	29	68.6	----						
Rhomos 7	30	71.4	68.6	----					
RKT(15-1-1)	31	54.3	62.9	57.1	----				
Ruffle 3	32	68.6	54.3	60.0	60.0	----			
Ruffle 4	33	54.3	48.6	57.1	60.0	65.7	----		

Speight G36	34	62.9	62.9	65.7	71.4	62.9	65.7	----	
ST(1-2-8)	35	74.3	71.4	71.4	60.0	60.0	62.9	71.4	----
ST(3-2-6)	36	57.1	54.3	68.6	65.7	65.7	68.6	57.1	68.6

STNCA(1-10-10-7-28-4)	37	71.4	68.6	77.1	54.3	68.6	57.1	57.1	68.6
62.9									
TL (13-20	38	60.0	62.9	60.0	68.6	68.6	54.3	60.0	62.9
68.6									
V572-MU	39	65.7	68.6	68.6	51.4	45.7	54.3	65.7	71.4
51.4									
Vamor 50	40	54.3	65.7	65.7	57.1	54.3	45.7	65.7	60.0
51.4									
Virginia 21	41	62.9	62.9	68.6	62.9	57.1	57.1	68.6	62.9
57.1									
WA 169-7 d.h.	42	62.9	60.0	62.9	65.7	57.1	62.9	71.4	68.6
57.1									
Wanda	43	57.1	62.9	71.4	65.7	54.3	57.1	65.7	71.4
62.9									
White gold	44	54.3	65.7	60.0	65.7	54.3	54.3	68.6	62.9
71.4									
Willow leaf	45	57.1	65.7	65.7	60.0	62.9	60.0	74.3	65.7
57.1									
WK 39	46	74.3	62.9	62.9	68.6	74.3	60.0	68.6	65.7
60.0									
Yellow gold	47	62.9	62.9	62.9	62.9	71.4	62.9	60.0	62.9
57.1									

		28	29	30	31	32	33	34	35
36									
STNCA(1-10-10-7-28-4)	37	----							
TL (13-20	38	74.3	----						
V572-MU	39	62.9	51.4	----					
Vamor 50	40	62.9	54.3	65.7	----				
Virginia 21	41	54.3	45.7	60.0	60.0	----			
WA 169-7 d.h.	42	45.7	45.7	57.1	62.9	77.1	----		
Wanda	43	62.9	60.0	60.0	60.0	68.6	71.4	----	
White gold	44	51.4	57.1	65.7	65.7	60.0	68.6	62.9	----
Willow leaf	45	68.6	57.1	65.7	74.3	74.3	60.0	65.7	65.7

WK 39	46	65.7	71.4	48.6	57.1	57.1	62.9	62.9	60.0
57.1									
Yellow gold	47	60.0	60.0	54.3	68.6	62.9	60.0	60.0	54.3
71.4									

		37	38	39	40	41	42	43	44
45									
WK 39	46	----							
Yellow gold	47	74.3	----						

		46	47

Appendix 4: Similarity matrix group 3

```
AW(22-10-1-1-65)  1    ----
Bel 61 - 12       2    54.3  ----
C7                3    60.0  62.9  ----
JING CHANG        4    57.1  60.0  51.4  ----
KB 14/a           5    48.6  54.3  57.1  54.3  ----
Mammoth Delcrest  6    57.1  65.7  60.0  45.7  54.3  ----
PD121             7    54.3  65.7  62.9  62.9  60.0  68.6  ----
Ruffle 1          8    62.9  65.7  60.0  68.6  57.1  60.0  74.3  ----
SR 1              9    45.7  60.0  57.1  57.1  65.7  62.9  68.6  68.6  ----
XA 219-5d.h      10    57.1  54.3  54.3  74.3  42.9  42.9  60.0  65.7  42.9

                          1     2     3     4     5     6     7     8     9

XA 219-5d.h      10    ----

                        10
```

Appendix 5: Similarity matrix group 4

```
AMARALINHO        1    ----
HAPLOID No.2      2    68.6  ----
KM 10             3    45.7  68.6  ----
Kutsaga mammoth   4    60.0  54.3  57.1  ----
Kutsaga mammoth E 5    60.0  60.0  51.4  71.4  ----
NC 744            6    51.4  51.4  40.0  60.0  60.0  ----

                          1     2     3     4     5     6
```

Appendix 6: Similarity matrix group 5

```
Meadows giant 1    ----
A56N (1-1-6)  2    57.1  ----
COKER 254     3    51.4  71.4  ----
NC 37         4    77.1  62.9  65.7  ----
TI 1372       5    54.3  54.3  80.0  51.4  ----

                      1     2     3     4     5
```

Appendix 7: Similarity matrix group 6

```
401                  1    ----
American 401         2    65.7  ----
American Joiner      3    74.3  85.7  ----
```

Bel 61 -10	4	65.7	68.6	60.0	----				
Broad leaf	5	71.4	74.3	68.6	62.9	----			
C 319	6	77.1	80.0	80.0	74.3	74.3	----		
CANDEL	7	68.6	68.6	68.6	71.4	65.7	74.3	----	
Carolina bright	8	71.4	68.6	71.4	68.6	71.4	68.6	68.6	----
Coker 48	9	62.9	65.7	65.7	65.7	54.3	68.6	57.1	62.9

Cokers Certified 401	10	80.0	71.4	74.3	74.3	71.4	77.1	74.3	62.9
54.3									
CSIRO 40T	11	62.9	65.7	68.6	60.0	65.7	62.9	60.0	62.9
51.4									
Daines bright leaf	12	77.1	71.4	77.1	68.6	68.6	77.1	77.1	77.1
57.1									
Delhi	13	71.4	77.1	80.0	62.9	65.7	74.3	74.3	68.6
54.3									
ENSHU	14	68.6	71.4	74.3	68.6	62.9	80.0	71.4	68.6
60.0									
H170	15	71.4	68.6	71.4	68.6	68.6	74.3	77.1	68.6
68.6									
Hickory Prior	16	65.7	71.4	74.3	68.6	71.4	74.3	77.1	54.3
51.4									
Hicks broad leaf	17	62.9	68.6	62.9	60.0	74.3	62.9	74.3	62.9
51.4									
Jamaican Wrapper	18	60.0	62.9	71.4	48.6	65.7	57.1	54.3	62.9
60.0									
K326	19	57.1	62.9	62.9	62.9	60.0	62.9	65.7	65.7
51.4									
Mammoth Gold	20	60.0	65.7	68.6	60.0	54.3	60.0	65.7	60.0
57.1									
MCNAIR	21	65.7	68.6	71.4	65.7	62.9	71.4	77.1	71.4
45.7									
MK94	22	45.7	65.7	65.7	51.4	54.3	60.0	57.1	54.3
62.9									
MK94	23	57.1	71.4	65.7	68.6	60.0	71.4	77.1	60.0
62.9									
NC 95	24	65.7	77.1	77.1	71.4	65.7	77.1	74.3	62.9
65.7									
NK 346	25	65.7	71.4	77.1	62.9	62.9	71.4	74.3	62.9
54.3									
NORDEL	26	57.1	65.7	71.4	57.1	48.6	71.4	51.4	60.0
68.6									
Oxford 1-181	27	68.6	62.9	68.6	60.0	54.3	68.6	77.1	68.6
54.3									
Oxford 26	28	71.4	74.3	77.1	71.4	68.6	85.7	71.4	68.6
68.6									
Proctors Special	29	42.9	54.3	60.0	42.9	57.1	48.6	45.7	54.3
54.3									
Queensland selection	30	65.7	62.9	62.9	57.1	60.0	65.7	77.1	71.4
71.4									
RKL (115-1-1)	31	65.7	74.3	77.1	68.6	57.1	74.3	71.4	62.9
60.0									
SC 58	32	65.7	68.6	80.0	57.1	54.3	71.4	57.1	62.9
71.4									

Special 400	33	65.7	74.3	71.4	71.4	74.3	80.0	74.3	65.7
57.1									
Speight G 10	34	71.4	65.7	65.7	60.0	68.6	68.6	77.1	68.6
57.1									
SPEIGHT G28	35	62.9	74.3	77.1	68.6	62.9	80.0	62.9	65.7
60.0									
SPEIGHT G70	36	68.6	68.6	68.6	77.1	65.7	80.0	94.3	68.6
62.9									
Station Bonanza	37	62.9	77.1	68.6	68.6	65.7	71.4	71.4	65.7
48.6									
TI 764	38	57.1	60.0	65.7	57.1	60.0	57.1	60.0	68.6
51.4									
Virginia Bright LIFT	39	74.3	60.0	60.0	54.3	65.7	65.7	71.4	68.6
54.3									
W2(1-7-1-1)	40	71.4	68.6	74.3	65.7	71.4	68.6	71.4	60.0
54.3									
White stem Orinoco	41	62.9	60.0	71.4	48.6	54.3	60.0	51.4	51.4
71.4									
WK 39	42	54.3	62.9	65.7	45.7	51.4	60.0	60.0	62.9
62.9									
Yellow Mammoth	43	57.1	74.3	82.9	54.3	62.9	68.6	62.9	62.9
60.0									
Yellow Prior	44	60.0	57.1	65.7	62.9	60.0	54.3	65.7	77.1
51.4									
ZZ 100	45	65.7	68.6	77.1	60.0	60.0	77.1	74.3	65.7
68.6									

		1	2	3	4	5	6	7	8
9									
Cokers Certified 401	10	----							
CSIRO 40T	11	65.7	----						
Daines bright leaf	12	74.3	71.4	----					
Delhi	13	68.6	68.6	77.1	----				
ENSHU	14	71.4	62.9	71.4	65.7	----			
H170	15	80.0	62.9	77.1	65.7	68.6	----		
Hickory Prior	16	77.1	62.9	71.4	71.4	74.3	68.6	----	
Hicks broad leaf	17	65.7	54.3	68.6	71.4	65.7	74.3	71.4	----
Jamaican Wrapper	18	54.3	48.6	57.1	57.1	54.3	65.7	54.3	62.9

K326	19	62.9	62.9	65.7	60.0	68.6	65.7	68.6	68.6
54.3									
Mammoth Gold	20	71.4	60.0	68.6	68.6	62.9	80.0	60.0	74.3
65.7									
MCNAIR	21	74.3	60.0	68.6	68.6	82.9	71.4	74.3	74.3
62.9									
MK94	22	48.6	45.7	54.3	68.6	54.3	60.0	60.0	68.6
71.4									
MK94	23	62.9	57.1	65.7	77.1	68.6	71.4	74.3	74.3
57.1									
NC 95	24	68.6	60.0	68.6	71.4	80.0	68.6	74.3	68.6
65.7									
NK 346	25	80.0	65.7	77.1	74.3	80.0	71.4	80.0	68.6

57.1
NORDEL 26 62.9 51.4 62.9 54.3 60.0 71.4 54.3 54.3
74.3
Oxford 1-181 27 74.3 54.3 77.1 71.4 65.7 74.3 62.9 65.7
57.1
Oxford 26 28 71.4 71.4 77.1 71.4 80.0 77.1 74.3 62.9
62.9
Proctors Special 29 51.4 40.0 48.6 48.6 51.4 54.3 57.1 65.7
68.6
Queensland selection 30 60.0 51.4 68.6 71.4 68.6 77.1 60.0 74.3
62.9
RKL (115-1-1) 31 80.0 68.6 74.3 71.4 80.0 71.4 80.0 60.0
54.3
SC 58 32 62.9 57.1 65.7 60.0 62.9 62.9 62.9 54.3
65.7
Special 400 33 85.7 57.1 68.6 65.7 77.1 77.1 80.0 68.6
60.0
Speight G 10 34 71.4 54.3 71.4 68.6 68.6 77.1 57.1 85.7
60.0
SPEIGHT G28 35 68.6 74.3 74.3 68.6 82.9 71.4 65.7 57.1
57.1
SPEIGHT G70 36 74.3 60.0 80.0 68.6 77.1 77.1 80.0 68.6
54.3
Station Bonanza 37 71.4 54.3 68.6 77.1 68.6 62.9 71.4 74.3
60.0
TI 764 38 62.9 54.3 62.9 57.1 51.4 60.0 42.9 60.0
71.4
Virginia Bright LIFT 39 65.7 60.0 60.0 62.9 71.4 62.9 60.0 60.0
51.4
W2(1-7-1-1) 40 77.1 74.3 74.3 80.0 65.7 68.6 71.4 65.7
60.0
White stem Orinoco 41 60.0 51.4 54.3 65.7 62.9 62.9 62.9 57.1
65.7
WK 39 42 54.3 45.7 51.4 62.9 57.1 65.7 54.3 60.0
71.4
Yellow Mammoth 43 62.9 54.3 60.0 62.9 71.4 71.4 62.9 68.6
77.1
Yellow Prior 44 62.9 57.1 65.7 62.9 65.7 57.1 68.6 65.7
62.9
ZZ 100 45 60.0 60.0 74.3 71.4 62.9 62.9 62.9 60.0
60.0

		10	11	12	13	14	15	16	17
18									
K326	19	----							
Mammoth Gold	20	62.9	----						
MCNAIR	21	71.4	74.3	----					
MK94	22	57.1	57.1	57.1	----				
MK94	23	68.6	68.6	65.7	77.1	----			
NC 95	24	62.9	65.7	71.4	65.7	74.3	----		
NK 346	25	68.6	74.3	80.0	57.1	65.7	74.3	----	
NORDEL	26	51.4	68.6	62.9	65.7	57.1	60.0	60.0	----

Oxford 1-181	27	60.0	80.0	77.1	57.1	68.6	62.9	74.3	62.9

Oxford 26	28	60.0	68.6	74.3	57.1	68.6	77.1	77.1	74.3
65.7									
Proctors Special	29	60.0	68.6	60.0	68.6	60.0	54.3	51.4	65.7
57.1									
Queensland selection	30	60.0	74.3	68.6	65.7	80.0	65.7	65.7	57.1
77.1									
RKL (115-1-1)	31	62.9	80.0	77.1	51.4	65.7	77.1	88.6	65.7
74.3									
SC 58	32	68.6	60.0	60.0	62.9	57.1	65.7	60.0	77.1
60.0									
Special 400	33	60.0	68.6	82.9	57.1	65.7	71.4	85.7	65.7
71.4									
Speight G 10	34	71.4	77.1	71.4	57.1	68.6	68.6	65.7	54.3
74.3									
SPEIGHT G28	35	62.9	65.7	71.4	54.3	65.7	71.4	77.1	65.7
60.0									
SPEIGHT G70	36	71.4	60.0	71.4	57.1	77.1	80.0	74.3	54.3
71.4									
Station Bonanza	37	60.0	71.4	85.7	62.9	65.7	65.7	74.3	62.9
74.3									
TI 764	38	51.4	65.7	65.7	57.1	48.6	57.1	60.0	65.7
62.9									
Virginia Bright LIFT	39	54.3	57.1	71.4	45.7	60.0	65.7	62.9	45.7
65.7									
W2(1-7-1-1)	40	57.1	68.6	62.9	51.4	60.0	74.3	80.0	54.3
65.7									
White stem Orinoco	41	57.1	57.1	48.6	71.4	68.6	65.7	54.3	60.0
51.4									
WK 39	42	51.4	65.7	65.7	68.6	65.7	62.9	54.3	74.3
62.9									
Yellow Mammoth	43	62.9	68.6	68.6	68.6	62.9	77.1	65.7	71.4
57.1									
Yellow Prior	44	71.4	65.7	77.1	51.4	57.1	68.6	74.3	51.4
65.7									
ZZ 100	45	57.1	54.3	60.0	71.4	65.7	71.4	62.9	65.7
65.7									

		19	20	21	22	23	24	25	26
27									
Oxford 26	28	----							
Proctors Special	29	48.6	----						
Queensland selection	30	65.7	60.0	----					
RKL (115-1-1)	31	82.9	51.4	65.7	----				
SC 58	32	68.6	74.3	54.3	62.9	----			
Special 400	33	80.0	54.3	68.6	82.9	54.3	----		
Speight G 10	34	60.0	57.1	80.0	62.9	57.1	65.7	----	
SPEIGHT G28	35	85.7	45.7	60.0	77.1	62.9	74.3	60.0	----
SPEIGHT G70	36	77.1	45.7	71.4	74.3	62.9	74.3	71.4	68.6

Station Bonanza	37	71.4	57.1	62.9	74.3	57.1	80.0	65.7	65.7

```
65.7
TI 764               38    60.0  57.1  57.1  57.1  57.1  62.9  68.6  62.9
54.3
Virginia Bright LIFT 39    65.7  45.7  74.3  65.7  48.6  71.4  71.4  62.9
65.7
W2(1-7-1-1)          40    77.1  40.0  57.1  77.1  57.1  68.6  65.7  68.6
71.4
White stem Orinoco   41    57.1  65.7  68.6  57.1  68.6  51.4  57.1  54.3
57.1
WK 39                42    62.9  68.6  68.6  57.1  68.6  60.0  57.1  51.4
54.3
Yellow Mammoth       43    65.7  65.7  62.9  65.7  74.3  65.7  71.4  71.4
62.9
Yellow Prior         44    60.0  65.7  62.9  68.6  65.7  65.7  60.0  54.3
65.7
ZZ 100               45    74.3  54.3  62.9  62.9  74.3  60.0  62.9  62.9
74.3

                           28    29    30    31    32    33    34    35
36

Station Bonanza      37    ----
TI 764               38    62.9  ----
Virginia Bright LIFT 39    62.9  51.4  ----
W2(1-7-1-1)          40    68.6  62.9  57.1  ----
White stem Orinoco   41    45.7  42.9  51.4  57.1  ----
WK 39                42    60.0  54.3  62.9  45.7  65.7  ----
Yellow Mammoth       43    57.1  71.4  54.3  62.9  65.7  68.6  ----
Yellow Prior         44    65.7  60.0  60.0  65.7  54.3  60.0  62.9  ----
ZZ 100               45    60.0  68.6  57.1  65.7  57.1  62.9  68.6  54.3
----

                           37    38    39    40    41    42    43    44
45
```

Appendix 8: Similarity matrix group 7

```
AE (48-18-5)         1    ----
Delcrest 66          2    74.3  ----
Delhi 61             3    74.3  62.9  ----
Dingigal cheroot     4    57.1  62.9  48.6  ----
DSD(5-4-2R)          5    65.7  85.7  62.9  71.4  ----
Harrisons special    6    71.4  74.3  62.9  68.6  68.6  ----
Harrisons special    7    62.9  71.4  77.1  57.1  77.1  77.1  ----
Oxford 26            8    71.4  60.0  60.0  60.0  60.0  65.7  62.9  ----
RL (1-2-1)           9    68.6  62.9  57.1  68.6  62.9  74.3  60.0  62.9  ---
-
TB 22               10    65.7  68.6  60.0  74.3  77.1  68.6  65.7  57.1
62.9
Vinica              11    60.0  54.3  57.1  51.4  51.4  51.4  51.4  60.0
51.4
```

ZAS 3-2-1-1-1-1	12	74.3	71.4	57.1	60.0	62.9	68.6	57.1	68.6
65.7									

		1	2	3	4	5	6	7	8
9									

TB 22	10	----							
Vinica	11	48.6	----						
ZAS 3-2-1-1-1-1	12	54.3	74.3	----					

	10	11	12

Appendix 9: Similarity matrix group 8

BS40-1	1	----							
Chicoana	2	71.4	----						
CR 401	3	65.7	62.9	----					
DAETW 23	4	62.9	68.6	74.3	----				
Delfield	5	60.0	68.6	77.1	60.0	----			
DELIOT	6	60.0	60.0	71.4	60.0	74.3	----		
DETW 144	7	68.6	74.3	68.6	71.4	65.7	68.6	----	
Flue-cured kentucky 14	8	51.4	60.0	54.3	65.7	57.1	71.4	62.9	----
Golden harvest	9	51.4	62.9	65.7	65.7	71.4	54.3	60.0	62.9
Kutsaga 51E	10	74.3	65.7	65.7	68.6	68.6	68.6	74.3	62.9
Kutsaga E1	11	68.6	68.6	74.3	65.7	60.0	60.0	60.0	57.1
MTRA 65	12	65.7	57.1	65.7	71.4	60.0	62.9	57.1	68.6
NC 2326	13	71.4	74.3	65.7	62.9	77.1	71.4	68.6	71.4
NC 95	14	54.3	54.3	74.3	62.9	65.7	77.1	57.1	74.3
NEW DEL	15	68.6	68.6	71.4	82.9	54.3	57.1	68.6	57.1
NOD 7	16	51.4	48.6	74.3	65.7	57.1	71.4	57.1	65.7
OB 56	17	74.3	74.3	71.4	71.4	71.4	65.7	74.3	54.3
Resistant Hicks	18	65.7	74.3	68.6	65.7	80.0	82.9	74.3	74.3
RKE (26-5-2-1-1)	19	54.3	57.1	65.7	68.6	51.4	65.7	60.0	71.4
RL (1-2-1)	20	62.9	71.4	60.0	54.3	57.1	62.9	62.9	54.3
Ruffle 3	21	62.9	60.0	77.1	74.3	68.6	68.6	71.4	62.9
STB (SS9-48-14-23)	22	71.4	74.3	71.4	65.7	68.6	68.6	82.9	60.0
STNCB(2-45-16-14-22-2)	23	57.1	68.6	65.7	65.7	71.4	57.1	60.0	57.1
Warne	24	51.4	57.1	68.6	65.7	57.1	68.6	62.9	62.9

		1	2	3	4	5	6	7	8
Golden harvest	9	----							
Kutsaga 51E	10	57.1	----						
Kutsaga E1	11	62.9	65.7	----					
MTRA 65	12	62.9	57.1	71.4	----				
NC 2326	13	77.1	74.3	74.3	62.9	----			
NC 95	14	54.3	57.1	65.7	65.7	57.1	----		
NEW DEL	15	51.4	71.4	60.0	60.0	57.1	60.0	----	
NOD 7	16	60.0	65.7	60.0	62.9	60.0	68.6	57.1	----
OB 56	17	54.3	71.4	68.6	65.7	68.6	62.9	71.4	57.1
Resistant Hicks	18	65.7	77.1	65.7	65.7	80.0	74.3	62.9	62.9
RKE (26-5-2-1-1)	19	57.1	57.1	65.7	68.6	60.0	71.4	65.7	60.0

RL (1-2-1)	20	54.3	60.0	62.9	57.1	60.0	62.9	57.1	60.0
Ruffle 3	21	62.9	68.6	65.7	57.1	62.9	77.1	68.6	62.9
STB (SS9-48-14-23)	22	62.9	74.3	62.9	60.0	71.4	60.0	71.4	54.3
STNCB(2-45-16-14-22-2)	23	74.3	60.0	74.3	65.7	74.3	62.9	57.1	51.4
Warne	24	60.0	60.0	57.1	65.7	57.1	65.7	68.6	60.0

	9	10	11	12	13	14	15	16
OB 56	17	----						
Resistant Hicks	18	74.3	----					
RKE (26-5-2-1-1)	19	54.3	65.7	----				
RL (1-2-1)	20	68.6	62.9	60.0	----			
Ruffle 3	21	65.7	74.3	65.7	51.4	----		
STB (SS9-48-14-23)	22	68.6	80.0	65.7	60.0	65.7	----	
STNCB(2-45-16-14-22-2)	23	74.3	68.6	62.9	62.9	62.9	57.1	----
Warne	24	54.3	65.7	88.6	62.9	68.6	68.6	57.1

	17	18	19	20	21	22	23	24

Appendix 10: Similarity matrix group 9

Duquesne Johnson	1	----							
American 402	2	51.4	----						
AW(22-10-1-1-3R)	3	62.9	65.7	----					
Cooker 347	4	68.6	60.0	62.9	----				
Delcrest	5	60.0	65.7	60.0	60.0	----			
Golden cure	6	60.0	62.9	68.6	60.0	60.0	----		
KUTSAGA E1B	7	62.9	62.9	51.4	60.0	57.1	71.4	----	
KUTSAGA MB2	8	62.9	62.9	54.3	68.6	57.1	65.7	57.1	----
NC 13	9	77.1	65.7	65.7	71.4	68.6	51.4	65.7	68.6
NC 5613	10	40.0	57.1	51.4	54.3	62.9	54.3	57.1	48.6
Oxford 402	11	57.1	80.0	51.4	54.3	71.4	62.9	60.0	62.9
PY (1-19-1)	12	65.7	62.9	65.7	57.1	62.9	51.4	54.3	62.9
PY 10	13	57.1	62.9	57.1	71.4	57.1	54.3	74.3	51.4
SCR	14	65.7	60.0	60.0	62.9	65.7	68.6	57.1	68.6
SCR	15	62.9	71.4	54.3	54.3	68.6	65.7	54.3	71.4
SIRONE	16	62.9	85.7	62.9	60.0	71.4	62.9	65.7	57.1
TI 799	17	68.6	54.3	54.3	68.6	54.3	54.3	51.4	65.7
TI 859	18	62.9	77.1	68.6	62.9	71.4	60.0	57.1	57.1
TI 933	19	51.4	74.3	54.3	62.9	57.1	62.9	57.1	54.3
TL 38	20	71.4	54.3	65.7	80.0	54.3	62.9	60.0	71.4

Virginia 312	21	62.9	62.9	71.4	71.4	60.0	62.9	48.6	60.0
68.6									
Virginia scheveiz	22	71.4	62.9	65.7	68.6	65.7	71.4	51.4	57.1
60.0									
WANDA	23	51.4	60.0	48.6	57.1	37.1	57.1	60.0	62.9
51.4									
Yellow Gold	24	65.7	57.1	51.4	62.9	54.3	62.9	65.7	65.7
71.4									
Yellow special A	25	71.4	62.9	57.1	74.3	62.9	54.3	65.7	71.4
88.6									
ZZ 100	26	51.4	68.6	57.1	45.7	68.6	54.3	48.6	48.6
62.9									

		1	2	3	4	5	6	7	8
9									
NC 5613	10	----							
Oxford 402	11	51.4	----						
PY (1-19-1)	12	54.3	62.9	----					
PY 10	13	51.4	60.0	57.1	----				
SCR	14	45.7	60.0	57.1	60.0	----			
SCR	15	54.3	77.1	60.0	42.9	65.7	----		
SIRONE	16	54.3	80.0	71.4	65.7	62.9	62.9	----	
TI 799	17	45.7	57.1	62.9	54.3	57.1	48.6	62.9	----
TI 859	18	57.1	77.1	65.7	60.0	68.6	65.7	82.9	54.3

TI 933	19	54.3	60.0	57.1	60.0	60.0	57.1	62.9	57.1
62.9									
TL 38	20	51.4	54.3	60.0	74.3	65.7	57.1	57.1	68.6
60.0									
Virginia 312	21	54.3	62.9	71.4	68.6	68.6	62.9	65.7	62.9
71.4									
Virginia scheveiz	22	45.7	71.4	54.3	62.9	74.3	65.7	68.6	68.6
71.4									
WANDA	23	42.9	51.4	40.0	60.0	48.6	48.6	57.1	42.9
51.4									
Yellow Gold	24	42.9	60.0	60.0	57.1	51.4	51.4	65.7	51.4
57.1									
Yellow special A	25	51.4	65.7	74.3	68.6	60.0	62.9	68.6	57.1
68.6									
ZZ 100	26	51.4	74.3	65.7	60.0	51.4	62.9	68.6	51.4
62.9									

		10	11	12	13	14	15	16	17
18									
TI 933	19	----							
TL 38	20	51.4	----						
Virginia 312	21	60.0	77.1	----					
Virginia scheveiz	22	65.7	71.4	77.1	----				
WANDA	23	45.7	60.0	40.0	48.6	----			
Yellow Gold	24	51.4	57.1	48.6	54.3	65.7	----		
Yellow special A	25	51.4	68.6	65.7	57.1	57.1	77.1	----	

```
ZZ 100            26    54.3  54.3  62.9  60.0  45.7  51.4  57.1  ----

                          19    20    21    22    23    24    25    26
```

Appendix 11: Similarity matrix group 10

```
TI 1112 1    ----
TI 764  2    40.0  ----
TI 804  3    40.0  97.1  ----
TI 814  4    51.4  74.3  74.3  ----

                1     2     3     4
```

Appendix 12: Similarity matrix - all 10 groups

```
1     ----
2     14.3  ----
3     14.3  11.4  ----
4     17.1  14.3  20.0  ----
5     22.9  17.1  14.3  22.9  ----
6     20.0  17.1  14.3  17.1  20.0  ----
7     14.3  14.3  20.0  22.9  17.1  17.1  ----
8     14.3  17.1  14.3  17.1  17.1  20.0  22.9  ----
9     14.3  11.4  14.3  17.1  17.1  17.1  14.3  14.3  ----
10    14.3  11.4  11.4  14.3  14.3  14.3  14.3  11.4  14.3  ----

          1     2     3     4     5     6     7     8     9    10
```

Appendix 13: Publications