GENETIC DIVERSITY ANALYSIS AND GENOTYPE X ENVIRONMENT INTERACTION IN ETHIOPIAN MUSTARD

(*Brassica carinata* A. BRAUN)

BY

TSIGE GENET KASSA

Submitted in the fulfilment of the requirements for the degree of Philosophiae Doctor, in the Department of Plant Sciences (Plant Breeding) Faculty of Natural and Agricultural Sciences

UNIVERSITY OF THE FREE STATE
BLOEMFONTEIN, SOUTH AFRICA
DECEMBER 2002

SUPERVISOR: PROF. M.T. LABUSCHAGNE
CO-SUPERVISOR: DR. C.D. VILJOEN

DECLARATION

I hereby declare that this dissertation, prepared for the degree Philosophiae Doctor, which was submitted by me to the University of the Free State, is my own original work and has not previously in its entirety or in part been submitted to any other University. All sources of materials and financial assistance used for the study have been duly acknowledged. I also agree that the University of the Free State has the sole right to the publication of this dissertation.

Signed on 5th of December 2002 at the University of Free State, Bloemfontein, South Africa.

Signature: ----------

Name: Tsige Genet Kassa

ACKNOWLEDGEMENTS

I wish to extend my sincere appreciation to the following people and organizations for their contribution for the success of this study:

1. Prof. M.T. Labuschagne (major-supervisor) and Dr. C.D. Viljoen (co-supervisor) for their interest in my work, supervision, and encouragement throughout the course of the study.

2. Elizma Koen for her excellent help at the molecular biology laboratory.

3. Prof. C.S. Deventer for his stimulating discussions and encouragement and Mrs Sadie Geldenhuys for her support in all administrative matters.

4. Dr. Arno Hugo and Eileen Roodt for their assistance rendered for determination of oil content and fatty acid composition.

5. The Agricultural Research and Training Project of the Ethiopian Agricultural Research Organization for the financial support.

6. The Biodiversity Conservation and Research Institute of Ethiopia for the provision of the germplasm.

7. Senderos Demeke, Seleshi Genet, Demeke Nigussie for their assistance in some statistical analysis.

8. All graduate students of Plant Breeding and Genetics for their useful discussions, and encouragement.

9. Alamine Atanaw, Melkie Niberet, Elias Menebere, and Minyechel Alamenehe for their assistance in field data collection.

DEDICATION

This piece of work is dedicated to my parents:

Enat Antenehe

Genet Kassa

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

AMMI	Additive Main effects and Multiplicative Interaction
AFLP	Amplified Fragment Length Polymorphism
ALA	Alpha Linolenic Acid
ANOVA	Analysis of Variance
bp	Base pairs
CGC	Capillary Gas Chromatography
CTAB	Cetyl trimethyl ammonium bromide
CV	Coefficient of variation
DR	Desaturation Ratio
DNA	Deoxyribonuleotide triphosphate
DNTP	Deoxynucleotide
EARO	Ethiopian Agricultural Research organization
EPA	Eicosapentaenoic acid
GC	Gas Chromatography
EDTA	Ethylene Diamine Teteracetic Acid
GLA	Gamma Linolenic Acid
LC	Leaf Colour
LS	Leaf Size
SC	Stem Colour
FC	Flower Colour
SEC	Seed Colour
AB	Angle of Branching
ER	Elongation Ratio
e.g.	*exempli gratia (*for example*)*
EFA	Essential Fatty Acids
et al	Et alli (and others)
EtOH	Ethanol
FA	Fatty acid
g	Gram
GDE	Genetic Distance Estimate
GE	Genotype x Environment

ha	Hectare
kg	Kilogram
Kg/ha	Kilogram per hectare
LDR	Linoleic Desaturation Ratio
ODR	Oleic Desaturation Ratio
ml	Mililiter
mm	Milimeter
mM	Milimolar
MUFA	Monounsaturated Fatty Acid
NaCl	Sodium chloride
ng	Nanogram
PCR	Polmerase Chain Reaction
PUFA	Polyunsaturated Fatty Acid
RAPD	Random Amplified Polymorphic DNA
RFLP	Restriction Fragment Length Polymorphism
UPGMA	Unweighted Pair Group with Arithmetic Average
rpm	Revolution per minute
SAS	Statistical Analysis System
SFA	Saturated Fatty Acid
SSR	Simple Sequence Repeat
TAE	Tris-acetate-EDTA
Taq	*Thermus aquaticus*
Tris-HCl	(Tris[hydroxymethyl]aminomethane hydrochloric acid)
UV	Ultraviolet
W/V	Weight per volume
W/W	Weight per weight
V/V	Volume per volume
°C	Degree Celsius
µg	Microgram
µl	Microliter
µM	Micromolar

CHAPTER I

INTRODUCTION

The oilseeds in Ethiopia constitute a considerable number and diversity of crop plants (Seegler, 1983). Ethiopian or Abyssinian mustard (*Brassica carinata* A. Braun.) is one of the major oil crops next to noug or nigerseed (*Guizotia abyssinica* Cass.) and linseed (*Linum usitatissimum L.*). In terms of area (25000 hectares) and production (28000 tonnes), it is the third most important oilseed crop (CSA, 1998). In Ethiopia mustard is cultivated primarily in the 500-1200 mm annual rainfall belt of the mid- and high-altitude (1700-2800 meters above sea level (masl)) areas in the central, south eastern and north-western plateaus.

Cultivation of Ethiopian or Abyssinian mustard (*B. carinata*) in Ethiopia is an old practice, which is believed to date back to the 4^{th} or 5^{th} Millennia BC (Simmonds, 1979). Traditional utilization of the crop in the country encompasses an array of purposes. Ground seeds are used to grease traditional bread-making clay pans (oven); this is possible owing to its erucic acid content, and is the reason erucic acid is used in plastic industry (Alemayehu, 2001). Ground seeds are used also to cure certain ailments, or stomach upsets, and to prepare beverages. Boiled leaves of young plants are excellent sources of vegetable relish. It is also a security crop, especially for the small farmers since the time of its vegetative stage coincides with periods of grain shortage, which occur in the middle of the rainy season when cereals are still in the booting stage. It is often grown in rotation with cereals to prevent the build-up of diseases, since it is immune to cereal diseases (Angus *et al.*, 1991; Alemayehu, 2001).

Brassica carinata (n=17, BBCC) has evolved in the Ethiopian plateau and the adjoining portion of East Africa and the Mediterranean coast through natural hybridisation of black mustard (*B. nigra (L.) Koch* n=8, BB) with cabbage (*B. oleraceae* L. n=9, CC) and followed by the chromosome doubling of the hybrid plant (UN, 1935; Mizushima, 1980; Hemingway, 1995; Gomez-Campo and Prakash, 1999). Wild forms of the species have not been reported (Shigesaburo-Tsunoda, 1980), but there are diverse eco-types with a range of morphological and agronomic differences (Abebe *et al.*, 1992).

Brassica carinata is the most adapted oil crop in Ethiopia and is a commercial crop only in Ethiopia. Outside Ethiopia, it has been tested in Canada (Getinet *et al.*, 1996; Falk, 1999), India (Malik, 1990), Spain (Fereres *et al.*, 1983), California (Cohen and Knowles, 1983) and Zambia (Mnzava and Olsson, 1990). In California it was found to be very slow maturing and low yielding and had a lower harvest index than *B. napus* and *B. juncea* (Cohen and Knowles, 1983).

In Ethiopia, it is more heat and drought tolerant, resistant to diseases and pod shatter and higher yielding when compared to *B. napus* (Knowles *et al.*, 1981; Fereres *et al.*, 1983; Malik, 1990). *B. carinata* can also serve as an important source of resistance genes, which are rare in other oilseed *Brassicas*. In order to use such important genes, the genetic diversity and geographic pattern of the variability should be studied.

The oil of *B. carinata*, like those of other oil seed *Brassicas*, contains between 35 and 44 % erucic acid (C22: 1). But as opposed to most other vegetable oils, contains lower amounts of the fatty acids with C16 and C18 carbon atoms (Röbbelen and Thies, 1980; Westphal and Marquard, 1980; Röbbelen, 1981; Downey, 1990; Becker *et al.*, 1999). Feeding experiments of 1940's and 1950's suggested that erucic acid was associated with problems of poor digestibility, weight loss, mycocardial lipidosis and death (Sauer and Kramer, 1983). Therefore the present erucic acid level in *B. carinata* cultivars is significantly beyond the level acceptable for human nutrition (Sauer and Kramer, 1983).

Crop breeders and agronomists have been striving to develop improved genotypes that are superior in seed yield and contain other desirable agronomic characteristics over a wide range of environmental conditions. However, a significant genotype x environment interaction can seriously limit progress in selection. Testing of materials over sites and years to ensure that the forthcoming cultivars have stable performance over a wide range of environments is a universal practice (Yau, 1995; Yan and Hunt, 1998). Consequently, many plant breeders use estimates of various stability parameters to assist them in identifying superior genotypes in the presence of genotype x environment interactions.

Objectives of this study therefore, were to:

1) Study the geographical pattern of morphological variation in the germplasm of *B. carinata* collections.

2) Assess the variation of oil content and fatty acid compositions of *B. carinata* germplasm collections.

3) Analyse the amount of genetic diversity in *B. carinata* and identify genotypes that are genetically different to be used for the crossing program.

4) Compare various statistical procedures for assessing genotype x environment interaction and yield stability of Ethiopian mustard.

CHAPTER II

LITERATURE REVIEW

2.1. Phenotypic diversity analysis

The foundation of the Ethiopian farming comprises the traditional crops and landraces (crop plant populations that have not been bred as varieties but have been adapted through years of natural or artificial selections under which they are cultivated) which farmers have adapted over centuries of selection and which is used to meet dynamic and changing needs (Worede, 1993). Ethiopian farmers are also instrumental in conserving germplasm as they control the bulk of the country's genetic resources. Peasant farmers retain some seed stock for security unless circumstances dictate otherwise. Even when forced to temporarily leave their farms because of severe drought or other threats, farmers have often stored small quantities of seed for later use (Worede *et al.*, 2000).

The broad range of genetic diversity existing in Ethiopia is presently subject to serious genetic erosion and irreversible losses. This threat results from the interaction of several factors that is progressing at an alarming rate. The most crucial factors include the displacement of indigenous landraces by new genetically uniform crop cultivars, changes in development agriculture or land use, destruction of habitats and ecosystems, and drought (Worede *et al.*, 2000).

Two conservation strategies are generally distinguished: *in situ* and *ex situ.* At the global level, genetic erosion has been addressed by efforts to conserve plant genetic resources in off-farm or *ex situ* gene banks, both as seeds and as living plants. To date, nearly all such efforts have focused on conserving crop genetic resources in formal gene banks that are part of the international institutional network. *Ex situ* conservation has limits. Gene banks are limited in what they can store. They have collected only a fraction of the existing genetic diversity and the size of the sample varies and depends on the crop. The materials kept in the gene banks are not accessible by the primary users and the original custodian of the materials. It also terminates the enhancement of the material through the process of natural evolution (Demissie, 1999).

The other conservation strategy is *in situ* conservation. The primary objectives of *in situ* conservation are to conserve the biodiversity of traditional crop varieties on the farm with the help of farmer's knowledge and traditional practices. *In situ* or on-farm conservation of agrobiodiversity is conservation in a dynamic agroecosystem, ideally one that is self-supporting and favouring evolutionary processes. Thus, it allows ongoing host-parasite co-evolution, which is likely to provide material resistant to diseases and pests (Demissie, 1999). Maintaining this dynamic process is especially significant in drought-prone regions of the country.

A wider genetic base of germplasm is a prerequisite to the success of a plant-breeding programme and to cope with unforeseen breeding challenges in a changing environment. Landraces of crop species have been the priority targets of collection since Vavilov's expedition to the various parts of the globe in the 1920's (Bechere *et al.*, 1996). Although what and how much to conserve has been controversial, there has been general agreement that landraces should be conserved, either *in situ* or *ex situ (*Bechere *et al.*, 1996).

Quantitative estimates of phenotypic diversity with respect to geographical origin and altitude class are a prerequisite for a sound genetic conservation strategy. Diversity analysis of world collections of germplasm of several crop species and wild relatives using the Shannon Weaver diversity index (H') have revealed considerable variability for a wide range of characters. Several authors have used this index extensively to estimate the phenotypic diversity in crop germplasm collections (Perry and McIntosh 1991; Yang *et al.*, 1991; Ayana and Bekele, 1999; Polignano *et al.*, 1999; Yoon *et al.*, 2000; Kebebew *et al.*, 2002; Upadhyaya *et al.*, 2002).

For nominal scale data there is no mean or median to serve as a reference for discussion of dispersion. Instead the concept of *diversity* is used to study the distribution of observations among categories. Observations distributed evenly among categories result in high diversity, whereas a set of observations where the bulk of the data occurs in a few of the categories is the one exhibiting low diversity (ZAR, 1984).

If a set of nominal scale data may be considered to be a random sample, then a quantitative expression appropriate as a measure of diversity is that of Shannon

(1948). For n independent information, whose probabilities of choice are $p_1, p_2, ..., p_n$ then the actual expression for the information is:

$$H' = -[p_1 \log p_1 + p_2 \log p_2 + ... + p_n \log p_n], \qquad (1)$$

Any probability is a number less than or equal to one, and the logarithms of numbers less than one are themselves negative. Thus the minus sign is necessary in order that H' be positive (Shannon and Weaver, 1949).

$$H' = -\sum_{i=1}^{k} p_i \log p_i , \qquad (2)$$

The diversity index is often referred to as Shannon-Wiener diversity index or the Shannon-Weaver index. Here, k is the number of categories and p_i is the proportion of the observations found in category i. Denoting n to be the sample size, and f_i to be the number of observations in category i, then $p_i = f_{i/n}$. Some mathematical manipulation arrives at the equivalent function:

$$H' = \frac{n \log n - \sum_{i=1}^{k} f_i \log f_i}{n}, \qquad (3)$$

This formula is easier to use than equation (2) because it eliminates the necessity of calculating the proportions (p_i). Published tables of $n \log n$ and $f_i \log f_i$ are available (Lloyd *et al.*, 1968; Brower and Zar, 1977). Any logarithmic base may be used to compute H'; bases 10, e and 2 are most frequently encountered in that order of commonness. A value of H' calculated using one logarithmic base may be converted to that of another base. According to Bowman *et al.* (1971), H' is known to be an underestimate of the diversity in the sampled population, however, this bias decreases with increasing sample size. The magnitude of H' is not only affected by the distribution of the data but also by the number of categories, for theoretically, the maximum possible diversity for a set of data consisting k categories is

$$H'_{max} = \log k . \qquad (4)$$

Therefore, many users of Shannon's index prefer to calculate

$$J' = \frac{H'}{H'_{max}} \qquad (5)$$

instead of, or in addition to, H', thus expressing the observed diversity as a proportion of the maximum possible diversity. The quantity J' has been termed *evenness* (Pielou, 1966) and may also referred to as *homogeneity* or *relative diversity.*

H', is a measure of the uncertainty, or of the choice associated with a frequency distribution (vector) ρ; p_i is the probability of the different events i, or the relative frequency (on a scale 0 to 1) of each state i in the descriptor-vector ρ. Shannon recognized that this equation was similar to the equation of entropy that physicist Boltzmann had published in 1898 as a quantitative formulation of the second law of thermodynamics, about the degree of disorganization in a closed system. He thus concludes that H' corresponds to the entropy of an information system (Legendre and Legendre, 1983).

2.2. Morphological characters as markers

The use of morphological characters dates back to breeding and selection itself (Koebner *et al.,* 1994). The application of morphological characters has been used for different purposes. It has been used as a powerful tool in the classification of lines, to study taxonomic status, identification, determination of genetic variation and correlation of characters with agronomic potential (Millan and Cubero, 1995).

A basic prerequisite of any breeding programme is the presence of a genetic variation, from which selection can be made (Dudley and Moll, 1969). The careful selection of parental genotypes is therefore critical. This requires organization of the germplasm so that only genotypes that are different for the traits under consideration are employed. Genetic distance estimates might also be useful for identifying heterotic groups for crops in which this information is required, but not currently available for *B. carinata.*

Estimation of genetic diversity and the relationships between gemplasm collections are very useful for facilitating efficient gemplasm collection and management. Many tools are now available for studying the variability and the relationships among accessions including total seed proteins, isozymes, and various types of molecular

markers. However, morphological characterization is the first step in the description and classification of germplasm (Smith and Smith, 1989). Various numerical taxonomic traits have been successfully used to classify and measure the patterns of phenotypic diversity in the relationships of species and germplasm collections of a variety of crops (Gomez-Campo and Tortosa, 1974; Takahata and Hinata, 1986; Gupta *et al.*, 1991; Perry and MacIntosh, 1991; Dias *et al.*, 1993; Amurrio *et al.*, 1995; Li *et al.*, 1995; Revilla and Tracy, 1995; Smith *et al.*, 1995; Tatineni *et al.*, 1996; Rabbani *et al.*, 1998).

2.3. Lipid and fatty acids composition

Lipids are a heterogeneous class of compounds whose general solubility in organic solvents and insolubility in water distinguishes them from other cellular constituents such as proteins, carbohydrates, and nucleic acids (Hitchcock, 1975). The bulk of the world lipids are produced by plants and, of these, acyl lipids form the largest part (Harwood, 1996). As in most eucaryotic organisms, plant lipids have three main functions (1) they are basic components of cellular membranes, (2) acyl lipids (almost always as tricylglycerols) are an important energy store (Murphy, 1994), (3) many plant lipids or their metabolic derivatives have acute biological activity, (4) in plants lipids have a fourth major function as constituents of the surface layers (Harwood, 1996). This layer, which includes epicuticular wax, cutin, and suberin, serves as a vital barrier between the plant tissues and external environment (Harwood, 1996).

Vegetable oils are predominantly (92-98%) tricylglycerols, the most important of the remaining components being polar lipids (phospholipids and galactolipids), mono- and diacylglicerols, free fatty acids, and polyisoprenoid lipids (Åppelqvist, 1989).

Triacylglycerols are glycerol molecules containing one fatty acid esterified to each of the three-hydroxyl groups. The stereochemical positions of the three fatty acids in the glycerol molecules are designed $sn-1, sn-2$ and $sn-3$. Both the relative amount of fatty acids that are present in the oil and their distribution in triacylglycerol molecular species determine the physical, chemical, physiological, and nutritional properties of vegetable oils (Padley *et al.*, 1994).

Each oil has a characteristic pattern of tricylglycerols, which depends on the available fatty acids and the specificity of the biosynthetic enzymes (Fernández-Moya *et al.*, 2000). The fatty acids are not distributed randomly between the different *sn*-carbon atoms of the tricylglycerol molecule. As a general rule, saturated fatty acids are confined to positions *sn-1* and *sn*-3, whereas polyunsaturated fatty acids are located mainly at the *sn*-2 position (Stymne and Stobart, 1987). One exception is the palm oil, which contains higher levels of saturated fatty acids at the *sn*-2 position. This has been suggested to have negative biological effects and to be involved in the atherogenic process (Renaud *et al.*, 1995). In rapeseed oil, erucic acid is excluded from the *sn*-2 position, which results in a theoretical breeding limit for increasing the concentration of this fatty acid of 66% of the total fatty acids (Taylor *et al.*, 1994).

Fatty acids differ in their number of carbon atoms and/or number and position of the carbon chain of double bonds and functional groups (hydroxy, epoxy, etc.). Depending on the presence or absence of double bonds, the fatty acids are divided into *saturated*, which do not contain double bonds at all, and *unsaturated*, which contain at least one double bond. The unsaturated fatty acids are divided into *monounsaturated,* those with one bond and *poly-unsaturated* fatty acids those with two or more double bonds. Unsaturated fatty acids usually exist in the *cis* form (Charley and Weaver, 1998). As an example of the nomenclature commonly used to indicate the three parameters, linoleic acid is presented as 18:2 (n-6), which expresses that this fatty acid consist of a chain with 18 carbon atoms, two double bonds, with the six carbon from the methyl end being the first unsaturated one. The distance between the methyl end of the carbon chain and the first double bond is of utmost importance for the nutritional and pharmaceutical properties of fatty acids (Åppelqvist, 1989). The unsaturated fatty acids can also have two possible configurations, *cis* and *trans*, depending on the relative position of the alkyl groups. This is of great relevance from nutritional point of view, since trans fatty acids have a detrimental effect on human beings (Willet and Ascherio, 1994). Most naturally occurring unsaturated fatty acids have the *cis*- orientation, although several common industrial processes such as hydrogenation, which is applied for example for margarine production, induce cis-trans isomeration (Tatum and Chow, 1992).

All fatty acids are linked in the biosynthetic pathway through modifications such as elongation and desaturation (Figure 2.1). This fact determines that the alteration of any of the biosynthetic steps influences the whole fatty acid profile. A comprehensive review of the description of the biosynthesis of the various fatty acids and their enzymatic relationships have been made by Harwood (1996).

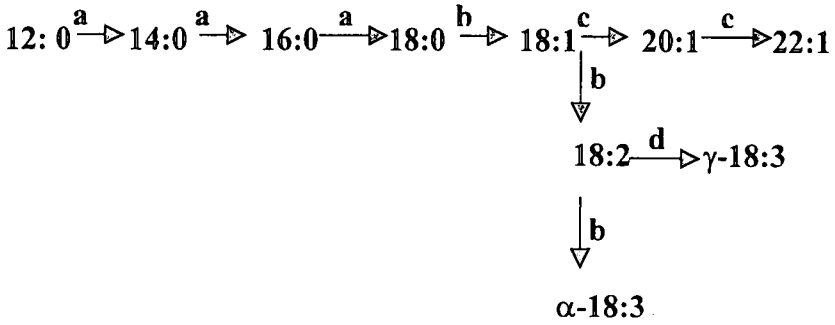


Figure. 2.1. Schematic representation of the biosynthetic pathway of the principal fatty acids. a=elongation of saturated fatty acids; b=desaturation pathway to α-linolenic acid, typical of linseed; c=elongation of monounsaturated fatty acids, typical of the *Brassicaceae* family; d=desaturation pathway to γ-linolenic acid, typical of borage and evening primose (adapted from Velasco and Fernández-Martinez, 2002).

The overall scheme for the biosynthetic pathways of fatty acids in *Brassica* is given in Figure 2.2.

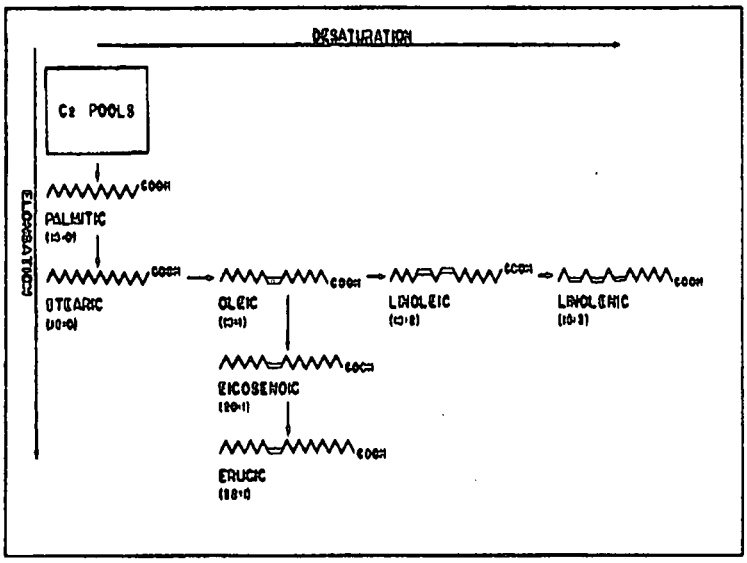


Figure 2.2. Biosynthetic pathways of the major fatty acids in oilseed brassicas (adapted from Downey, 1990; Thies, 1998).

In general, saturated fatty acids have a hypercholesterolemic effect, whereas unsaturated fatty acids act by lowering serum cholesterol. The exception to this rule is stearic acid, which exhibits a neutral effect (Mensink *et al.*, 1994). Linoleic acid, α-linolenic, and γ-linolenic acids are three of the essential fatty acids, i.e., they must be included in the diet because the human body is not able to manufacture them (Horrobin, 1992). These fatty acids are polyunsaturated, i.e., they contain more than one double bond in the carbon chain. Polyunsaturated fatty acids are more susceptible to autoxidation than monounsaturated or saturated fatty acids. The double bonds react rapidly with oxygen in the air in a process involving the production of free radicals, which are implicated in a number of diseases, tissue injuries and in the process of aging (Shahidi, 1996). Furthermore, the breakdown products of fatty acid autoxidation are the major source of off-flavours in oils, which reduce their shelf life (Tatum and Chow, 1992). Therefore, although polyunsaturated fatty acids such as α-linolenic are beneficial *per se*, their susceptibility to autoxidation make them undesirable at high levels in vegetable oils. Oleic acid is considered as the preferred fatty acid for edible purposes, as it combines a hypocholesterolemic effect and a high oxidative stability (Mensink and Katan, 1989; Yodice, 1990).

Solely *de novo* metabolic processes within animal tissues cannot meet requirements for polyunsaturated acyl chains. Animals are absolutely dependent on plants for the two major precursors of (n-6) and (n-3) fatty acids, linoleic and linolenic acids. In animal tissue these acyl chains can be converted to fatty acids containing 3-6 double bonds (Cook, 1991).

Severe effects observed in experimental animals and in humans in the absence of these dietary acids include a dramatic decrease in weight, dermatosis and increased skin permeability to water, enlarged kidneys and reduced adrenal and thyroid glands, cholesterol accumulation, and ultimate death. The four (n-6) acids in the sequence from 18:2 (n-6) to 20:4 (n-6) individually have similar potency in reversing these effects of deficiency, whereas the activity of 18:3 (n-3) alone is much lower. Thus the term "essential fatty acid" clearly applies at least to the two major (n-6) acids (Cook, 1991). A function of 18:2 (n-6), in addition to its role as precursor of 20:4 (n-6), seems likely (Mead, 1984).

The relationships among fatty acids in metabolic conversions can be evaluated by considering groups or families of fatty acids. The predominant fatty acid families are the (n-6) acids derived from 18:2 (n-6), the (n-3) acids derived from 18:3 (n-3), the (n-9) acids derived from 18:1 (n-9), and the (n-7) acids derived from 16:1 (n-7) (Cook, 1991).

Oil quality is a relative concept that depends on the end use of the oil. Vegetable oils are intended for food and non-food applications. The former includes salad and cooking oils as well as oils for the food industry (margarines, shortening, etc.). The latter comprises countless industrial sectors such as lubricants, surfactants, surface coatings, cosmetics, plastics, etc. (Velasco and Fernández-Martínez, 2002). In general, the oil characteristics undesirable for a particular application are required for others. Therefore breeding for improved oil quality is in some ways a continuous exercise of divergent selection.

2.4. DNA-based molecular markers (DNA fingerprinting)

Molecular markers include proteins and nucleic acids that are detectably different, i.e., polymorphic among individuals or populations. Markers that reveal polymorphisms at the protein level are known as biochemical markers, while DNA markers reveal polymorphisms at the DNA level. Biochemical markers are proteins produced as a result of gene expression, which can be separated by electrophoresis to identify the alleles. The most commonly used protein markers are isozymes, which are variant forms of the same enzyme (Vodenicharova, 1989). Other biochemical characteristics, such as lipids and sugars are also considered as markers (Winter and Kahl, 1995). Protein markers reveal differences in the gene sequence and function as co-dominant markers. However, their use is limited due to their limited number in any crop species (Kumar, 1999).

Molecular markers provide a quick and reliable method for estimating genetic relationships among genotypes. They can facilitate rapid screening of large numbers of genotypes for polymorphic loci (Thormann *et al.*, 1994). The most appropriate markers should be those that are (1) heritable, (2) discriminate between individuals or populations, (3) are easy to measure and evaluate and (4) provide results that can be compared with similar studies (Westman and Kresovich, 1997).

Molecular markers offer numerous advantages over conventional phenotypic alternatives as they: (1) are stable and detectable in all tissues regardless of growth, differentiation, development, or defence status of the plant cell; (2) are unaffected by the environment, and (3) generally lack pleiotropic and epistatic effects (Caetano-Anollés and Trigiano, 1997).

Applications of molecular markers in the plant system involve improvements in the efficiency of conventional plant breeding by carrying out indirect selection through molecular markers linked to the trait of interest both simple and quantitative traits (QTL), because these markers are not influenced by the environment and can be scored at all stages of plant growth. In addition DNA markers can also be used in plant systems for germplasm characterization, genetic diagnostics, characterization of

transformants, study of genome organization, phylogenetic analysis etc. (Rafalski *et al.*, 1996). Although each marker system associated with some advantages and disadvantages, the choice of marker system is dictated to a large extent by the intended applications, convenience and the cost involved. Gupta *et al.* (1999) broadly classified these molecular markers in the following groups: (1) Hybridization-based DNA markers such as restriction fragment length polymorphism (RFLPs) (Beckman and Soller, 1983) and oligonucleotide fingerprinting (2) polymerase chain reaction (PCR) (Mullis *et al.*, 1986) based DNA markers such as random amplified polymorphic DNAs (RAPDs) (Williams *et al.*, 1990) which can also be converted into sequence characterized amplified regions (SCARs), simple sequence repeats (SSRs) or microsatellites (Tautz, 1989) sequence–tagged sites (STS), amplified fragment length polymorphisms (AFLPs) (Vos *et al.*, 1995), inter-simple sequence repeat amplification (ISA), cleaved amplified polmorphic sequences (CAPS) and amplicon length polymorphisms (ALPs). (3) DNA chip and sequencing-based DNA markers such as single nucleotide polymorphisms (SNPs).

4.1. Amplified Fragment Length Polymorphism (AFLP)

Amplified Fragment Length Polymorphism (AFLP) is a PCR based technology for marker-assisted breeding and genotyping. AFLP represents a significant breakthrough compared to the currently available methods in terms of facility, precision, flexibility, speed and cost. AFLP enables the generation of thousands of DNA markers from a genome of any complexity and without prior knowledge of the genome's structure or sequence.

AFLP involves the amplification of small restriction fragments, obtained by cleaving genomic DNA with restriction enzymes, to produce high-resolution DNA "fingerprinting" patterns on denaturing polyacrylamide gels. The rational of the AFLP technique is based on the use of specifically designed PCR primers which selectively amplify a small subset of restriction fragments, or markers out of the complex mixture comprising as many as several million different fragments. The product of the reaction can be visualized by conventional DNA staining or DNA labelling procedures using either radioactive or non-radioactive methods.

AFLP is an extremely flexible technology, which offers multiple applications in the fields of crop breeding and plant genome analysis, especially in the fields of genotyping, marker assisted breeding and plant genome analysis (Thottappilly *et al.*, 2000).

The AFLP techniques can be used for DNA samples of any origin or complexity. Small sequence variations can be detected using only small quantities of genomic DNA (0.05-0.5μg). The capacity to reveal many polymorphic bands in one lane is a major advantage of AFLP markers. The numerous bands on a gel are analysed simultaneously making AFLP an extremely efficient technique. AFLP has the capacity to inspect a much greater number of loci for polymorphism than other currently available PCR-based techniques, such that the number of polymorphisms detected per reaction is much higher. AFLP is superior in terms of the number of sequences amplified per reaction and their reproducibility. The markers produced are reliable and reproducible within and between laboratories, and are relatively easy and inexpensive to generate. A virtually unlimited number of markers can be generated by simply varying the restriction enzymes, and the nature and number of selective nucleotides (Blears *et al.*, 1998).

2.5. Genotype x environment interactions and stability analyses of Ethiopian Mustard

2.5.1. Genotype x environment interaction

Genotype x environment (GE) interactions are an important issue facing plant breeders and agronomists. A significant GE interaction for a quantitative trait such as grain yield can seriously limit progress in selection. Testing of selected materials over sites and years to ensure that forthcoming cultivars have stable performance over a range of environments is a universal practice (Yau, 1995).

An understanding of environmental and genotypic causes of GE interaction is important at all stages of plant breeding, including ideotype design, parent selection, selection based on traits, and selection based on yield (Jackson *et al.*, 1996; Yan and Hunt, 1998). Understanding of the cause of GE interaction can be used to establish breeding objectives, identify ideal test conditions, and formulate recommendations for areas of optimal cultivar adaptation (Yan and Hunt, 2001). The basic cause of differences between genotypes in their yield stability is the wide occurrence of GE interactions. These interactions of genotypes with environments can be partly understood as a result of a differential reaction to environmental stress factors like drought or diseases, and other factors (Becker and Léon, 1988).

Data collected in multilocation trials are intrinsically complex, having three fundamental aspects: (a) structural patterns; (b) non-structural noise; and (c) relationships among genotypes, environments, and genotypes and environments considered jointly (Crossa, 1990). A pattern implies that a number of genotypes respond to certain environments in a systematic, significant, and interpretable manner, whereas noise suggests that the responses are unpredictable and un-interpretable (Crossa, 1990). The function of experimental design and statistical analyses of multilocation trials is thus to eliminate and discard as much of this unexplainable noise as possible (Crossa, 1990).

2.5.2. Reducing genotype x environment interaction

There are three possible strategies for a plant breeder to develop varieties that show a low GE interaction:

(1) The subdivision of a heterogeneous area into smaller and more homogeneous regions (Tai, 1971). This stratification of environments usually is based on macro-environmental differences such as temperature gradients, rainfall distribution, and soil types (Eberhart and Russell, 1966). However, even with this technique, the interaction of genotypes with locations in a sub region, and with environments encountered at the same location in different years frequently remains large

(2) Increase the number of cultivars

(3) The introduction of varieties that show a high degree of stability in performance over a wide range of environments (Tai, 1971). According to Eberhart and Russell (1966), the use of genetic mixtures rather than homogeneous or pure line varieties has been suggested as a means to reduce GE interaction. Allard and Bradshaw (1964) suggested that heterozygous and heterogeneous populations offer the best opportunity to produce varieties, which show small GE interactions. The latter authors use the term individual buffering and population buffering. Individual buffering is used when the individual members of the population are well buffered such that each member of the population is well adapted to a range of environments. Thus, a heterozygous or homozygous genotype may have individual buffering (Eberhart and Russell, 1966). On the other hand population buffering occurs when the population consists of a number of genotypes each adapted to a somewhat different range of environments. Thus, according to Eberhart and Russell (1966), heterogeneous populations may have population buffering.

2.5.3. Concepts of stability

In the presence of GE interaction, the use of genotypic means across environments as criteria for selecting superior genotypes is not valid. This leads to the concept of stability of performance (Kang, 1990). Lin *et al.* (1986) pointed out that the concept of stability could be defined in many ways, depending on how the scientist wishes to

look at the problem. There is also no consensus on which stability concept would be most useful for applications by plant breeders (Kang, 1990).

Lin *et al.* (1986) have classified stability into three types as follows:
Type 1: A genotype is considered to be stable if its among-environment variance is small. This kind of stability is useful when the environments considered are not very diverse and is equivalent to the static concept of stability (Becker and Lèon, 1988). A genotype showing this type of stability would not respond to a high level of inputs such as fertilizers (Kang, 1990). This type of stability would be of little use to a farmer if the stability related to yield and the cultivars in question were low yielding (Kang, 1990), but would be useful if the stability related to quality traits, to disease resistance, or to stress characters like winter hardiness (Becker and Léon, 1988). These, latter characters usually are controlled by one or a few genes, and their variation from environment to environment is often negligible. Coefficient of variation (CV_i) and genotypic variances across environment (S_i^2) can be used to describe this type of stability.

Type 2: A genotype is considered to be stable if its response to environments is parallel to the mean response of all genotypes in the trial. Type 2 stability is equivalent to the dynamic concept, in which a stable genotype has no deviations from the general response to environments (Becker and Léon, 1988). For each environment the performance of the stable genotype corresponds completely to the estimated level or prediction. A regression coefficient (Finlay and Wilkinson, 1963) stability variance can be used for measuring type 2 stability.

Type 3: A genotype is said to be stable if the residual mean square (MS) from the regression model on an environmental index is small. The environmental index in this context is the mean yield of all the genotypes in each location minus the grand mean of all genotypes in all locations. The use of type 3 stability is difficult to justify unless the environmental index can be replaced by actual environmental factors such as temperature, rainfall, etc. (Lin *et al.*, 1986). Type 3 is also dynamic. The methods of Eberhart and Russell (1966) and Tai (1971) can be used for estimating Type 3 stability.

Becker and Léon (1988) also distinguished between two different concepts of stability, termed static stability and dynamic stability. Static stability is defined as a stable genotype possessing unchanged performance regardless of any variation of the environments, thus implying that its variance among environments is zero. This is equivalent to the biological concepts of stability and similar to Type 1 stability of Lin *et al.* (1986). Dynamic stability is defined as a genotype having a predictable response to environments and thus has no deviations from this response to environments. Becker (1981) termed this type of stability the agronomic concept to distinguish it from the biological or static concept. Becker and Léon (1988) stated that all stability procedures based on quantifying GE interaction effects belong to the dynamic stability concept. This includes the procedures for partitioning the GE interactions of Wricke's (1962) ecovalence and Shukla's (1972) stability of variance, procedures using the regression approach such as proposed by Finlay and Wilkinson (1963), Eberhart and Russell (1966) and Perkins and Jinks (1968), as well as non-parametric stability statistics.

Lin *et al.* (1986) defined four groups of stability statistics. Group A is based on deviations from average genotype effect (DG), group B on GE interaction term (GEI), and group C and D on either DG or GEI. Group A and B formulas represent sums of squares, and those of groups C and D represent regression coefficient or regression deviation. They integrated type 1, type 2, and type 3 stabilities with the four groups: group A was regarded as type 1, groups B and C as type 2, and group D as type 3 stability. In type 1 stability, genotype is regarded as stable if its among-environment variance is small; in type 2, a genotype is regarded as stable if its response to environment is parallel to the mean response of all genotypes in a test; and in type 3 stability, a genotype is regarded as stable if the residual mean square from regression model on environmental index is small (Lin *et al.*, 1986). Lin and Binns (1988) proposed type 4 stability concepts on the basis of predictable and unpredictable non genetic variation; the predictable component related to locations and the unpredictable component related to years. Lin and Binns (1988) suggested the use of a regression approach for the predictable portion and the mean square for years-within-locations for each genotype as a measure of the unpredictable variation. The latter was called type 4 stability statistics.

2.5.4. Statistical methods for measuring GE interactions

A combined analysis of variance procedure is the most common method used to identify the existence of GE interaction from replicated trials over a series of environments. If the GE interaction variance is found to be significant, one or more of the various methods for measuring the stability of genotypes can be used to identify stable genotype(s). The statistics, which can be used to identify stable genotypes, are classified into parametric and nonparametric. Parametric statistics are more useful when the data are continuous, nonparametric when the data are discontinuous. Nonparametric data analysis has the potential to reduce complex data into intuitive measures of stability.

Lin *et al.* (1986) have described nine parametric stability statistics: (1) the variance of a genotype across environments (S_i^2); (2) coefficient of variability (CV_i); (3) Plaisted and Peterson's (1959) mean variance component for pairwise GE interaction $(\bar{\theta}_i)$; (4) Plaisted's (1960) variance component for GE interaction $(\theta_{(i)})$; (5) Wricke's (1962) ecovalence (W_i); (6) Shukla's (1972) stability variance $(\sigma^2{}_i)$; (7) Finlay and Wilkinson's (1963) regression coefficient (b_i); (8) Perkins and Jink's (1968) regression coefficient (β_i); (9) Eberhart and Russell's (1966) deviation parameter $(s^2 d_i)$.

According to Becker and Léon (1988) the parametric approach described above gives only the individual aspects of the stability but can not provide an over all picture of the response. The basic reason for this apparent difficulty is that a genotype's response to environment is multivariate yet the multivariate approach tries to transform it into a univariate problem, via a stability index. To avoid this problem, a different line of thought has emerged, namely to cluster genotypes according to their response structure (i.e. non-parametric method).

Although the parametric approach to stability is relatively simple, it does not provide information for the resolution of any conflicting Type 1 and Type 2 inferences. Under these circumstances, quantitative mathematical characterization should be considered

as well as qualitative descriptions of genotypes, as like or unlike genotypes; i.e., to adopt a non-parametric clustering procedure (Lin *et al.*, 1986)

2.5.4.1. Analysis of variance

In a conventional variety trial the yield of G genotypes is measured in E environments each with R replicates. The classic model for analysing the total yield variation contained in GER observations is the analysis of variance (Fisher, 1918; 1925). After removing the replicate effect when combining the data, the GE observations are partitioned into two sources: (a) additive main effects for genotypes and environments and (b) nonadditive effects due to GE interaction. The analysis of variance of the combined data expresses the observed (Y_{ij}) mean yield of the i^{th} genotype at the j^{th} environment as:

$$Y_{ij} = \mu + G_i + E_j + GE_{ij} + e_{ij} \qquad (1)$$

Where μ is the general mean; G_i, E_j, and GE_{ij} represent the effect of the genotype, environment, and genotype-environment interaction, respectively; and e_{ij} is the average of the random errors associated with the r^{th} plot that receives the i^{th} genotype in the j^{th} environment. The nonadditivity interaction as defined in (1) implies that the expected value of the i^{th} genotype in j^{th} environment (Y_{ij}) depends not only on the levels of G and E separately but also on the particular combination of levels of G and E (Crossa, 1990).

The major limitation in this analysis is that the error variances over environments should be homogeneous to test for genotypic differences. If error variances are heterogeneous, this analysis is open to criticism as the F-test of the G x E interaction mean squares against the pooled error variances is biased. A correct test of significance, by weighting each genotype mean by the inverse of its estimated variance, has been used by Yates and Cochran (1938) and Cochran and Cox (1957). The weighted analysis gives less weight to environments that have a high residual mean square. The disadvantage of weighted analysis is that weights may be correlated to environment yield responses (with high yielding environments showing higher

error variance and low yielding sites presenting lower error variances). This would mask the true performance of some genotypes in certain environments (Crossa, 1990).

One of the main deficiencies of the combined analysis of variance of multilocation yield trials is that it does not explore any underlying structure within the observed nonadditivity (genotype-environment interaction) (Crossa, 1990). The analysis of variance fails to determine the pattern of response of genotypes and environments. The valuable information contained in (G-1) (E-1) degrees of freedom is particularly wasted if no further analysis is done. Since the nonadditive structure of data matrix has a non-random (pattern) and random (noise) component, the advantage of the additive model are lost if the pattern component of the nonadditive structure is not further partitioned into functions of one variable each (Crossa, 1990).

Analysis of variance of multilocation trials is useful for estimating variance components related to different sources of variation, including genotypes and genotype-environment interaction. In general variance component methodology is important in multilocation trials, since errors in measuring the yield performance of a genotype arise largely from GE interaction. Therefore, knowledge of the size of this interaction is required to (a) obtain efficient estimates of genotype effects and (b) determine optimum resource allocations, that is the number of plots and locations to be included in future trials. In a breeding program, variance component methodology is used to measure genetic variability and to estimate the heritability and predicted gain of a trait under selection (Crossa, 1990).

2.5.4.2. Regression coefficient (b_i) and deviation mean square $(s^2 d_i)$

Joint linear regression (JLR) is extensively used method for analysing and interpreting the non-additive GE interaction of two-way classification data. The GE interaction is partitioned into a component due to the linear regression (b_i) of the i^{th} genotype on environment mean, and a deviation (d_{ij}):

$$(GE)_{ij} = b_i E_j + d_{ij} \qquad (2)$$

and thus $$Y_{ij} = \mu + G_i + E_j + (b_i E_j + d_{ij}) + e_{ij} \qquad (3)$$

This model was first proposed by Yates and Cochran (1938) in their evaluations of barley yield trials. The method divides the $(G-1)(E-1)$ df for interaction into $G-1$ df for heterogeneity among genotype regression and the remainder $(G-1)(E-2)$ for deviation. Further details about interaction are obtained by regressing the performance of each genotype on the environmental means. Finlay and Wilkinson (1963) determined the regression coefficient by regressing variety mean on environmental mean, and plotting the obtained genotype regression coefficients against the genotype mean yield. Figure 2.3 is a generalized interpretation of the genotype pattern obtained when genotype regression coefficients are plotted against genotype mean yields.

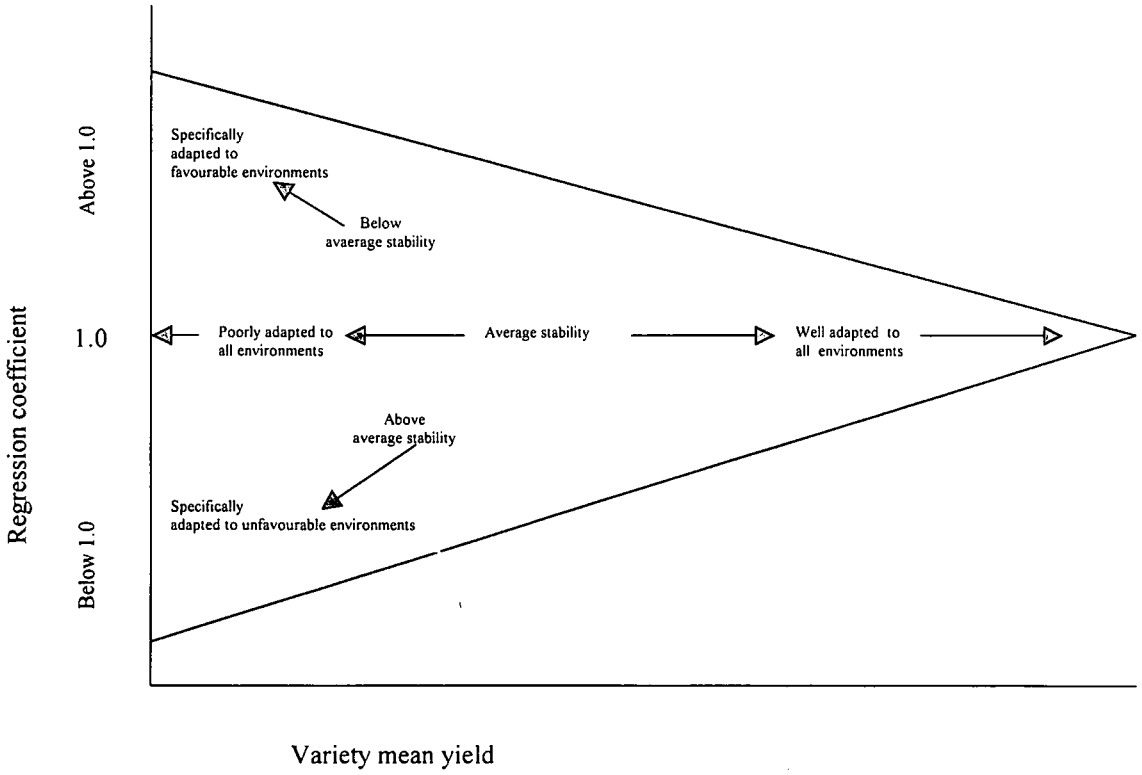


Figure 2.3. A generalized interpretation of the variety population pattern obtained when variety regression coefficients are plotted against variety mean, according to Finlay and Wilkinson (1963).

In Finlay and Wilkinson's (1963) study, the adaptation of the whole population of varieties was facilitated by the use of a two-dimensional plot (Scatter diagram), with mean yield and regression coefficient as coordinates of each variety. Here, an absolute phenotypic diversity would be expressed by regression coefficient $(b_i) = 0$. Though wide variation was evident in both mean yield and sensitivity to environment as characterized by the regression coefficient, the variation in sensitivity was proportionally less among varieties with higher mean yield. One of the most interesting features of their study is that the variability in phenotypic stability (regression coefficient) is inversely proportional to the mean yield. Perkins and Jinks (1968) regression coefficient is similar to Finlay and Wilkinson's (1963) except that the observed values are adjusted for location effects before the regression. While, for the Eberhart and Russell (1966) deviation parameter, the residual mean square of deviation from the above mentioned regression is the measure of stability for each genotype. They defined a genotype with regression coefficient $(b_i) = 1$ to be stable.

Eberhart and Russell (1966) proposed pooling the sum of squares for environments and GE interactions and subdividing it into a linear effect between environments (with 1 df), a linear effect for GE (with G-1 df), and a deviation from regression for each genotype (with E-2 df). In effect the residual mean square from the regression model across environments is used as an index of stability, and a stable genotype is one in which the deviation from regression mean square $(s^2 d_i)$ is small:

$$s^2 d_i = \frac{1}{E-2}\left[E_j(X_{ij} - \bar{X}_{i\cdot} - \bar{X}_{\cdot j} + \bar{X}_{\cdot\cdot})^2 - (b_i - 1)^2 E_j(\bar{X}_{\cdot j} - \bar{X}_{\cdot\cdot})^2\right] \qquad (4)$$

It was not until this era that the problem of solving the intractable GE interaction problem could be solved. Subsequently Freeman (1973), Hill (1975) and Westcott (1986) have reviewed the regression approach to study GE interaction extensively.

Freeman (1990) reported that the stability method which attempted to analyse GE interactions as opposed to simply recognizing their existence was that of joint regression analysis. Many developments have followed the use of joint regression, including several definitions of the stability of a genotype. Although joint regression

analysis has been criticized as giving biased results it has continued to be used. However, despite the advantage of certain meaningful interpretations of GE interaction through joint regression, several statistical and biological limitations and criticisms of this method should be noted.

The first statistical criticism of regression analysis is that the genotype mean is not independent of the marginal means of the environment. Regressing one set of variables on another that is not independent violates one of the assumptions of regression analysis (Freeman and Perkins, 1971; Freeman, 1973).

The second statistical limitation is that errors associated with the slopes of the genotypes are not statistically independent, because the sum of squares for deviation, with (G-1) (E-2) df, can not be subdivided orthogonally among the G genotypes (Crossa, 1990).

The third statistical problem is that it assumes a linear relationship between interaction and environmental means. When this assumption is violated, the effectiveness of the analysis is reduced, and results may be misleading (Mungomery *et al.*, 1974).

A major biological problem with regressing genotype means on environmental means arises when only a few low or very high yielding sites are included in the analysis. The genotype fit may be determined largely by its performance in a few extreme environments, which in turn generates misleading results (Westcott, 1986). Regression analysis should be used with caution when the data set includes results from only a few extremely high or low yielding locations (Crossa, 1990).

Becker and Léon (1988) in their review noted that the regression approach is of little use if the regression coefficient b_i is included in the definition of "stability". For this reason b_i generally viewed by authors not as a measure of stability, but rather as additional information on the average response of a genotype to advantageous environmental conditions. This is schematically presented in Figure 2.4. (Becker and Lèon, 1988).

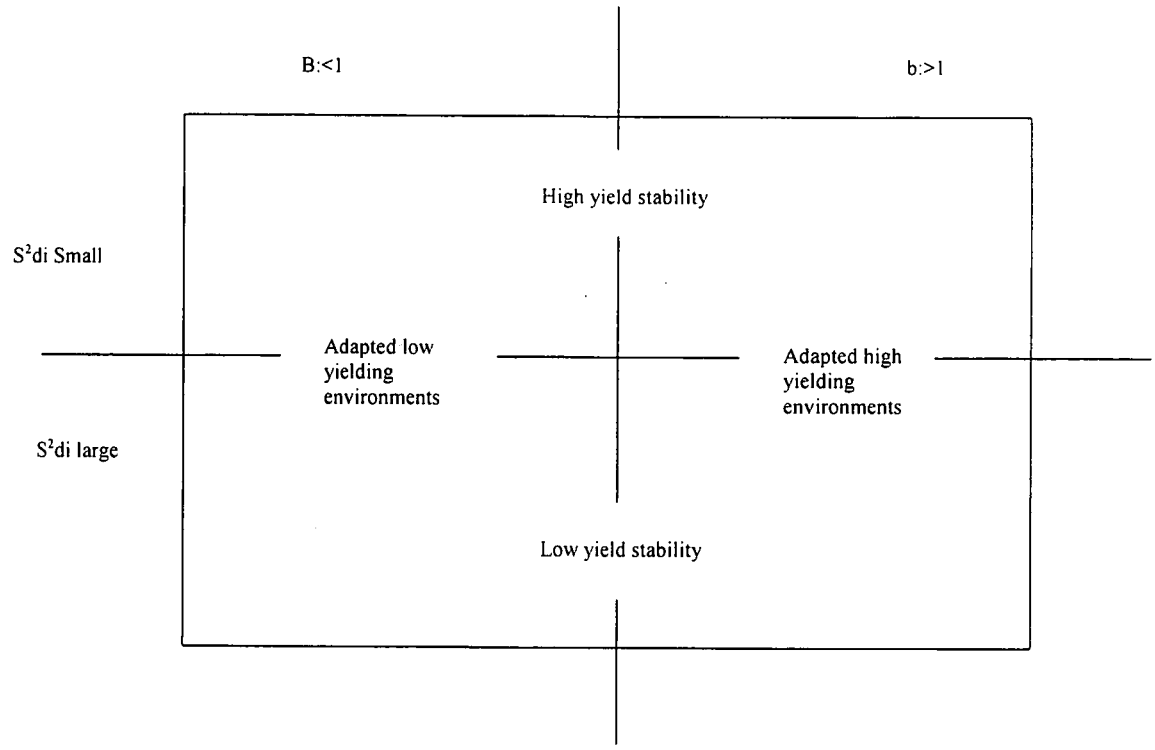


Figure 2.4. Interpretation of the parameters b_i and $s^2 d_i$ for the regression approach (Becker and Léon, 1988).

Alternative methods of determining genotype stability based on the GE interaction are also available. The most important and frequently used ones are discussed as follow.

2.5.4.3. Coefficient of determination (r_i^2)

Pinthus (1973) proposed to use the coefficient of determination (r_i^2) instead of deviation mean squares to estimate stability of genotypes, because r_i^2 is strongly related to s^2_{di} (Becker, 1981):

Coefficient of determination: $r_i^2 = 1 - \frac{s^2_{di}}{s^2_{xi}}$ (5)

The application of r_i^2 and b_i has the advantage that both statistics are independent of the units of measurement.

2.5.4.4. Ecovalence (W_i)

Wricke (1962) proposed using the contribution of each genotype to the GE interaction sum of squares as a stability measure and defined this concept or statistic as ecovalence (W_i). Ecovalence is simple to calculate and is expressed as:

$$W_i = \sum_j [Y_{ij} - \bar{Y}_{i\cdot} - \bar{Y}_{\cdot j} + \bar{Y}_{\cdot\cdot}]^2, \quad (6)$$

Where, Y_{ij} is the mean performance of genotype i in the j^{th} environment and $Y_{i\cdot}$ and $Y_{\cdot j}$ are the genotype and environment mean deviations respectively, and $Y_{\cdot\cdot}$ is the overall mean. For this reason, genotypes with a low W_i value have smaller deviations from the mean across environments and are thus more stable. According to Becker and Léon (1988) ecovalence measures the contribution of a genotype to the GE interaction, a genotype with zero ecovalence is regarded as stable. According to the meaning of the word ecovalence, this stable genotype possesses a high ecovalence (low values of W_i = high ecovalence).

Becker and Lèon (1988) illustrated ecovalence by using a numerical example of plot yields of genotype i in various environments against the respective mean of environments (Figure 2.5).

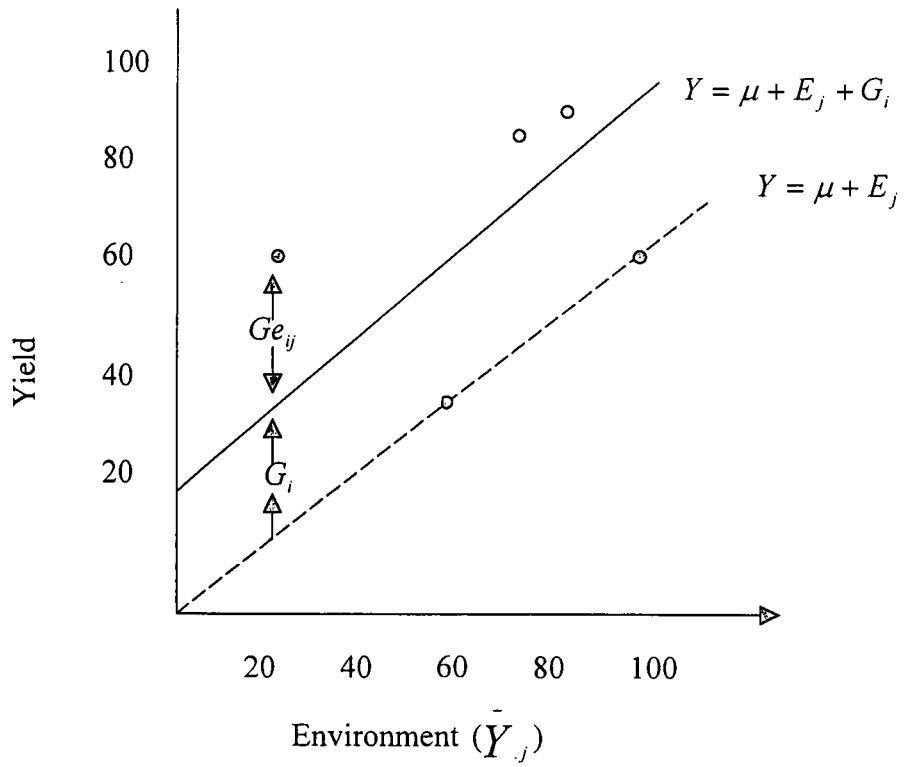


Figure 2.5. Graphical representation of GE interactions: the stability statistic ecovalence (W_i) is the sum of squares of deviations from the upper straight line.

The lower straight line estimates the average yield of all genotypes simply using information about the general mean (μ) and the environmental effects (E_j), while the upper line takes into account the genotype effect (G_i) and therefore estimates the yield of genotype i. Deviations of yield from the upper straight line are the GE interaction effects of genotype i and are summed and squared across environments and constitute ecovalence.

2.5.4.5. Shukla stability variance parameters (σ^2_i and s^2_i)

Shukla (1972) used the stability variance of a genotype across environments for discrimination of stability. According to Lin *et al.* (1986), Shukla's (1972) stability-variance (σ^2_i) is considered as Type 2 stability. That is, the method is the relative measure, which depends on the cultivars in the test, and thus the results of the test must be restricted to only those genotypes in the test and should not be generalized (Lin *et al.*, 1986). Therefore, a genotype is considered stable only with respect to the other genotypes in the test without any assurance that it will remain stable if it is compared with other sets of genotypes (Lin *et al.*, 1986). Moreover, the use of this method depends also on the range of environments. If the range is very large, then this method can be useful. In this method, Shukla (1972) considered the partitioning of the GE interaction sum of squares into components which are given each genotype separately by considering the stability variance (σ^2_i) of the i^{th} genotype defined as the variance over environments of $(g_{ij} + \bar{e}_{ij})$ in the equation below.

$$Y_{ijk} = \mu + a_i + E_j + g_{ij} + e_{ijk} \qquad (7)$$

Where, Y_{ijk} is the yield of the i^{th} genotype in the k^{th} replicate of the j^{th} environment, μ is the overall mean, a_i is the effect of i^{th} genotypes, E_j is the effect of j^{th} environment, g_{ij} is the interaction of i^{th} genotype in the j^{th} environment, and e_{ijk} is the random error, e_{ij} is the mean of e_{ijk} over replicate.

For t genotypes in s environments, the σ^2_{ij} is calculated as:

$$\sigma^2_i = \frac{1}{(s-1)(t-1)(t-2)}\left[t(t-1)\sum_j (Y_{ij} - \bar{Y}_{i\cdot} - \bar{Y}_{\cdot j} + \bar{Y}_{\cdot\cdot})^2 - \sum_i \sum_j (Y_{ij} - \bar{Y}_{i\cdot} - \bar{Y}_{\cdot j} + \bar{Y}_{\cdot\cdot})^2\right] \qquad (8)$$

Where, Y_{ij} is the mean yield of the i^{th} genotype in the j^{th} environment, $\bar{Y}_{i.}$ is the mean of genotype i in all environments, $\bar{Y}_{.j}$ is the mean of all genotypes in j^{th} environments and $\bar{Y}_{..}$ is the mean of all genotypes and all environments. The ratio of $\sigma^2{}_i$ to the pooled error mean square ($\sigma^2{}_e$) is used to approximate the F-test.

Where, $\sigma^2{}_e = \dfrac{\sum_i \sum_j \sum_k (Y_{ijk} - \bar{Y}_{ij})^2}{str(r-1)}$

r is the number of replications.

Once some of the genotypes are found unstable (i.e., significant value of $\sigma^2{}_i$), further progress in interpretation can be made by using a covariate (Shukla, 1972). In this case, an environmental index can be used as a covariate to remove its linear effects from the GE interaction and $s^2{}_i$ values are assigned to each genotype. An environmental index can be defined as the effect of differential fertility, cultural practices, incidence of diseases, and insect infestations, because these factors are not uniform across sites.

The covariate can be calculated for the j^{th} location as: $Z_j = \bar{Y}_{.j} - \bar{Y}_{..}$

$$s_i^2 = \frac{t}{(t-2)(s-2)}\left[S_i - \sum_i \frac{s_i}{t(t-1)}\right] \qquad (9)$$

Where $S_i = \sum_{j=1}^{s} (u_{ij} - \bar{u}_{i.} - b_i Z_j)^2$.

$u_{ij} = \bar{Y}_{ij} - \bar{Y}_{.j}$ and $\bar{u}_{i.} = \dfrac{\sum_j u_{ij}}{s}$

$b_i = \sum_j \left[(u_{ij} - \bar{u}_{i.})Z_j / \sum_j Z_j^2\right]$

Where, b_i is the regression coefficient of the i^{th} genotype, and Z_j a covariate for the j[th] environment

According to the equation for $s^2{}_i$ and b_i above, the GE interaction sum of squares is divided into two components, heterogeneity (non-additivity), and balance (residual). The residual is the remainder of the GE interaction variance after removing the linear

effect of the covariate, and it is partitioned into components (s^2_i) assignable to each cultivar. Thus stability variance parameters (σ^2_i and s^2_i) will be used in deciding whatever a genotype is stable or not based on F-test, which is the ratio of σ^2_i or s^2_i to the pooled error mean square (σ^2_e) with $(s-1)$ and $st(r-1)$ degrees of freedom. Relatively large values of σ^2_i and s^2_i show more un-stability in a genotype.

2.5.4.6. Cultivar superiority measure (P_i)

Lin and Binns (1988) defined the superiority measure (P_i) of the i^{th} genotype as the mean square of distance between the i^{th} genotype and the genotype with maximum response as:

$$P_i = \frac{\left[n(Y_i. - M..)^2 + (Y_{ij} - Y_i. + M_j. + M..)^2\right]}{2n} \qquad (10)$$

Where, Y_{ij} is the average response of the i^{th} genotype in the j^{th} environment, $Y_i.$ is the mean deviation of genotype i, M_j is the genotype with maximum response among all genotypes in the j^{th} locations, and n is the number of locations. The smaller the value of P_i, the less its distance to the genotype with maximum yield and the better the genotype. A pairwise GE interaction mean square between the maximum and each genotype is also calculated and it is similar to the method used by Plaisted and Peterson (1959), except that (a) the stability statistics are based on both the average genotypic effects and GE interaction effects and (b) each genotype is compared only with the one maximum response at each environment (Crossa, 1990).

2.5.5. Nonparametric techniques for stability analysis

Nonparametric statistics for GE interaction based on ranks provide a useful alternative to parametric approaches currently used, which are based on absolute data. Some advantages of nonparametric statistics compared to parametric ones are: reduction or even avoidance of the bias caused by outliers, no assumptions are needed about the

distribution of the analysed values, homogeneity of variances, and additivity (linearity) of effects are not necessary requirements (Hühn, 1996).

Further advantages are that nonparametric stability statistics are expected to be less sensitive to errors of measurement than parametric estimates and the addition or deletion of one or a few observations is not likely to cause great variation in the estimate as would be the case for parametric stability statistics (Nassar and Hühn, 1987).

When GE interactions are present, the effects of genotypes and environments are statistically nonadditive, which means that differences between genotypes depend on the environment. Existing GE interactions may, but not necessarily, lead to different rank orders of genotypes in different environments (Hühn, 1996). From a practical application point of view the experimenter is not interested in the knowledge of the numerical amount of GE interaction *per se*, but in the existence (or non existence) of GE interactions insofar as they led to different orderings of genotypes in different environment (Hühn, 1996). For two genotypes A and B, and two environments X and Y, the basic types of relationships between GE interactions and changes of rank orders are demonstrated schematically in Figure 2.6.

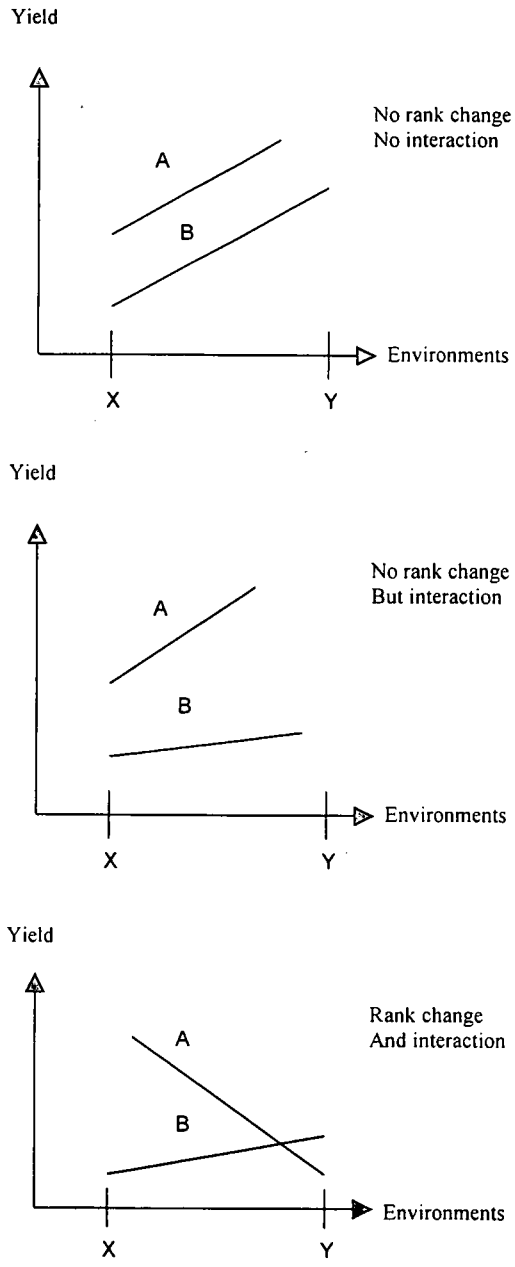


Figure 2.6. Genotype x environment interactions and changes of rank orders of different type of relationships (for two environments X and Y and two genotypes A and B modified from Wricke, 1965).

The crossover (qualitative) interactions are more important than non-cross over (quantitative) interactions in agricultural productions. For non-crossover interactions, the true treatment differences vary in magnitude but not in direction, where as for crossover interactions, the direction of true treatment differences varies (Gail and Simon, 1985; Baker, 1988; Crossa, 1990; Hühn, 1996). If significant qualitative interactions occur, subsets of genotypes are to be recommended only for certain

environments, whereas with quantitative interactions the genotypes with superior means can be used in all environments. Therefore testing of crossover interactions is important (Baker, 1988). Hühn (1996) has further studied the nonparametric analysis in detail and concluded that the procedure by Kroon and Van der Laan (1981) appears to be the best for plant breeding and selection.

Baker (1988) and Hühn (1990) discussed tests for rank interactions. Hühn (1990) also discussed non-parametric methods for grouping environments and genotypes. Two environments, regardless of their yield level, may be considered alike for selection purposes if both rank genotypes similarly. Therefore, the similarity of any two environments may be estimated by Spearman's Rank correlations Coefficient (Romagosa and Fox, 1993).

2.5.6. Multivariate stability analysis techniques

Multivariate techniques are widely applied in stability analysis to provide further information on real multivariate response of genotypes to environments. According to Becker and Léon (1988) multivariate analysis has three main purposes: (1) to eliminate noise from the data pattern, (2) to summarize the data, (3) to reveal the structure in the data. Through multivariate analysis, genotypes with similar responses can be clustered, hypothesized, and later tested, and their data can be easily summarized and analysed (Crossa, 1990; Gauch, 1982b; Hohls, 1995).

Multivariate analyses are appropriate for analysing two-way matrices of G genotypes and E environments. The response of any genotype in E environments may be conceived as a pattern in E-dimensional space, with the coordinate of an individual axis being the yield or other metric of the genotype in one environment (Crossa, 1990). Crossa (1990) has also distinguished two groups of multivariate techniques used to elucidate the internal structure of GE interactions:

(1) Ordination techniques, such as principal component analysis, principal coordinate analysis, and factor analysis, assumes that data is continuous. These techniques attempt to represent GE interrelationships as faithfully as possible in a low dimensional space. A graphical output displays similar genotypes or environments

near each other and dissimilar items are farther apart (Crossa, 1990). Ordination is effective for showing relationships and reducing noise (Gauch, 1982a; 1982b).

(2) Classification techniques, such as cluster analysis and discriminant analysis, seek discontinuities in the data. These methods involve grouping similar entities in clusters and are effective for summarizing redundancy in the data.

Williams (1976) recommended pattern analysis for describing GE interactions, defining it as the joint use of classification and ordination methods. This method groups genotypes and environments according to either their similarity (ordination methods) or dissimilarity (classification methods). Similarly measures such as Pearson's coefficient, which are larger for genotypes that are more similar for a set of environments and dissimilarity measures such as Euclidean distance, which are larger the more the genotypes are different (DeLacy *et al.*, 1996). However, the pattern analysis methods have been criticized on the grounds of their inability to distinguish between real pattern and background noise.

2.5.6.1. Principal component analysis

Principal component analysis is one of the most frequently used multivariate methods. Its aim is to transform the data from one set of coordinate axes to another, which preserves, as much as possible, the original configuration of the set of points and concentrates most of the data structure in the first principal components axes (Gower, 1966; Crossa, 1990)

Principal components analysis assumes that the original variables define a Euclidean space in which similarity between items is measured as Euclidean distance. This analysis can effectively reduce the structure of a two-way GE data matrix of G (genotypes) points in E (environments) dimension in a subspace of fewer dimensions. The matrix can also be conceptualised as E points in G dimensions.

The model is written as:

$$Y_{ij} = \mu + \sum_{n=1}^{h} k_n v_{ni} s_{nj} \qquad (11)$$

Where, k_n is the singular value of the n^{th} axis, v_{ni} is the eginvector of the i^{th} genotype for the n^{th} axis, s_{nj} is the eginvector of the j^{th} environment for the n^{th} axis, and $\sum_{n=1}^{h} v_{ni} = \sum_{n=1}^{h} s_{nj} = 1$. This result links the analysis of variance with the principal component analysis.

The principal component analysis was found to be efficient in describing GE interactions. Cruz (1992) showed that the principal component analysis was more efficient than regression model when he analysed a set of maize (*Zea mays* L.) data. Perkins (1972) also reported that principal component analysis was not useful for studying the adaptation of a group of inbred lines of tobacco (*Nicotiana tabacum* L.). Principal component analysis combined with cluster analysis was effective forming subgroups among 29 populations of faba bean (*Vicia faba* L.), which differed in mean performance of and response across environments (Polignano *et al.*, 1989).

Ordination techniques such as principal component analysis may have some limitations, e.g., in reducing dimensionality of multivariate data distortion may occur. If the percentage of variance accounted for by the first principal components axis is small, individuals that are really far apart may be represented by points that are close together. However, principal component analysis has an obvious advantage as compared with the linear regression methods. The regression analysis uses only one statistic, the regression coefficient, to describe the pattern of response of pattern of a genotype across environments, and most of the information is wasted in accounting for deviations. Principal component analysis overcomes this difficulty by providing the scores on the principal component axes to describe the response pattern of genotypes (Crossa, 1990). These scores allow depicting GE interactions into two dimensions (biplot) and identifying the factor responsible for the interaction.

2.5.6.2. Principal coordinate analysis

Principal coordinate analysis (Gower, 1966) is a generalization of principal components analysis in which any measure of similarity between individuals can be used. Its objectives and limitations are similar to those of principal component

analysis (Crossa, 1990). Principal coordinates analysis was used in combination with cluster analysis (pattern analysis) to study the adaptation of soybean lines evaluated across environments in Australia (Mungomery *et al.*, 1974; Shorter *et al.*, 1977). Westcott (1987) and Crossa (1990) also advocated the use of principal coordinates analysis.

2.5.6.3. Factor analysis

Factor analysis is an ordination technique related to principal components analysis, the factors of the former being similar to the principal component of the latter. Variation is explained in terms of general factors common to all variables and in terms of factors unique to each variable (Crossa, 1990).

2.5.6.4. Cluster analysis

Cluster analysis is a numerical classification technique that defines groups or clusters of individuals. Two types of classification can be distinguished. The first is nonhierarchical classification, which assigns each item to a class. The second type is hierarchical classification, which groups individuals into clusters and arranges these into a hierarchy for the purpose of studying relationships in the data (Crossa, 1990). Comprehensive reviews of the applications of cluster analysis to study GE interactions can be found in Lin *et al.* (1986) and Westcott (1987).

2.5.6.5. Additive Main Effects and Multiplicative Interaction (AMMI)

According to Zobel *et al.* (1988), considering the three traditional models, analysis of variance (ANOVA) fails to detect a significant interaction component, principal component analysis (PCA) fails to identify and separate the significant genotype and environment main effects, and linear regression models account for only a small portion of the interaction sum of squares. But, AMMI analysis reveals a highly significant interaction component that has clear agronomic meaning and it has no specific design requirements, except for a two-way data structure. In recent years, the AMMI biplot analysis received wide attention and is considered to be an effective tool to diagnose the GE interaction pattern. Van Eeuwijk (1995) has discussed the

theoretical aspects of about the linear and bilinear models for analysing GE interaction, and reported that the AMMI model could be treated as a bilinear model. In AMMI, the additive main effect portion is separated from interaction by ANOVA model. Then the principal component analysis (PCA), that provides a multiplicative model (Gabriel, 1971; Zobel *et al.*, 1988) is applied to analyse the interaction effect from the additive ANOVA model.

The advantages of the AMMI model or its variants are that, they use overall fitting, impose no restrictions on the multiplicative terms and result in least square fit (Freeman, 1990). Within limits, any model may be expected to fit the data from which it was derived. However, AMMI model has a good chance of being able to predict for new sites and new years, thus contributing a real advance (Gauch, 1988).

The principal components analysis of AMMI partitions GE interactions into several orthogonal axes, the interaction principal component analyses (IPCA). Gauch and Zobel (1996) showed that AMMI 1 with IPCA 1 and AMMI 2 with IPCA 1 and IPCA 2 are usually selected and the graphical representation of axes, either as IPCA 1 or IPCA 2 against main effects or IPCA 1 against IPCA 2 is generally informative. When AMMI 3 and higher models are presented for agricultural data, the third and higher IPCA axes are dominated by noise and have no predictive value (Van Eeuwijk, 1995)

Additive Main Effects and Multiplicative Interaction (AMMI) model combines analysis of variance for the genotype and environment main effects with principal components analysis of the genotype-environment interaction. It has proven useful for understanding complex GE interactions. The results can be graphed in a very useful biplot that shows both main and interaction effects for both genotypes and environments.

AMMI combines analysis of variance (ANOVA) and principal component analysis (PCA) into a single model with additive and multiplicative parameters.
The model equation is:

$$Y_{ij} = \mu + g_i + l_j + \sum_{k=1}^{n} \lambda_k \alpha_{ik} \gamma_{jk} + E_{ij}, \qquad (12)$$

Where, Y_{ij} is the yield of the i^{th} genotype in the j^{th} environment; μ is the grand mean; g_i and l_j are the genotype and location deviations from the grand mean, respectively; λ_k is the egien value of the principal component analysis axis k; α_{ik} and λ_{jk} are the genotype and location principal component scores for axis k; n is the number of principal components retained in the model; and E_{ij} is the error term.

Here, the interaction is explained in the form of biplot display where, PCA scores are plotted against each other and it provides visual inspection and interpretation of the GE interaction components. Integrating biplot display and genotypic stability statistics enables genotypes to be grouped based on similarity of performance across diverse environments.

REFERENCES

ABEBE D., D. TADESSE, G. MULAT and D. BEYENE, 1992. Ethiopia's oilseed genetic resources. *In*: Oilseeds research and development in Ethiopia. First National Oilseeds Workshop, 3-5 December 1991, Addis Ababa. Institute of Agricultural Research, pp. 13-20.

ALEMAYEHU, N., 2001. Germplasm diversity and genetics of quality and agronomic traits in Ethiopian mustard. Ph.D. Thesis, University of Göttingen.

ALLARD, R.W., and A.D. BRADSHAW, 1964. Implication of genotype-environment interactions in plant breeding. *Crop Sci.* 4: 503-508.

AMURRIO, J.M., A.A. DE RON, and A.C. ZVEN, 1995. Numerical taxonomy of Liberian pea land races based on quantitative and qualitative characters. *Euphytica* 82: 195-205.

ANGUS, J.F., A.F. van HERWAADEN, and G.N. HOWE, 1991. Productivity and break crop effect of winter growing oilseeds. *Aust. J. Experimental Agriculture* 31: 669-677.

ÅPPELQVIST, L.A., 1989. The chemical nature of vegetable oils. *In:* R.K. Downey. G. Röbbelen and A. Ashri (eds.). Oil Crops of the World. New York: McGraw-Hill, pp. 22-37.

AYANA, A., and E. BEKELE, 1999. Geographical patterns of morphological variation in Sorghum (*Sorghum bicolour* (L) Moench.) germplasm from Ethiopian and Eritrea. Qualitative characters. *Hereditas* 129: 195-205.

BAKER, R.J., 1988. Test for crossover genotype-environment interactions. *Can. J. Plant Sci.* 68: 405-410.

BECHERE, E., G. BELAY, D. MITIKU, and A. MERKER, 1996. Phenotypic diversity of tetraploid wheat landraces from northern and north-central regions of Ethiopia. *Hereditas* 124: 165-172.

BECKER, H.C., 1981. Correlations among some statistical measures of phenotypic stability. *Euphytica* 30: 835-840.

BECKER, H.C., and J. LÉON, 1988. Stability analysis in plant breeding. *Plant Breeding* 101: 1-23.

BECKER, H.C., H. LÖPTIEN, and G. RÖBBELEN, 1999. Breeding: an overview. *In*: C. Gomez-Campo (ed.). Biology of *Brassica* Coeno-species. 413-460. Elsevier, Science BV, Amsterdam.

BECKMAN, J.S., and M. SOLLER, 1983. Restriction fragment length polymorphism in plant genetic improvement: methodologies, mapping and costs. *Theor. Appl. Genet.* 67: 35-43.

BLEARS, M.J., S.A. De GRANDIS, H. LEE, and J.T. TREVORS, 1998. Amplified fragment length polymorphism (AFLP): a review of the procedure and its applications. *Journal of Industrial and Biotechnology* 21: 99-114.

BOWMAN, K.O., K. HUTCHESON, P. ODUM, and L.R. SHENTON, 1971. Comments on the distribution of indices of diversity, *Statistical Ecology* 3: 315-366.

BROWER, J.E., and J.H. ZAR, 1977. Field and laboratory methods for general ecology. Wm. C. Brown, Dubuque, Iowa. 194 pp.

CAETANO-ANOLLÉS, G., and R.N. TRIGIANO, 1997. Nucleic acid markers in agricultural biotechnology. *AgBiotech News and Information* 9(10): 235N-242N.

CENTRAL STATISTICAL AUTHORITY OF ETHIOPIA (CSA), 1998. Agricultural sample survey for 1997/98: Area and production of belg season crops for private peasant holdings. Volume 5 Addis Ababa, Ethiopia.

CHARLEY, H., and C. WEAVER, 1998. Foods: A scientific approach. 3rd edition. Prentice-Hall Inc.

COCHRAN, W.G., and G.M. COX, 1957. Experimental designs. 2nd Ed. Wiley, New York.

COHEN, B.M., and R.P. KNOWLES, 1983. Evaluation of *Brassica* species in California. Proc. of 6th International Rapeseed Congress, Paris, pp. 282-285.

COOK, H.W., 1991. Fatty acid desaturation and chain elongation in eucaryotes. *In:* VANCE, D.E., AND J. VANCE (Eds.), Biochemistry of lipids, Lipoproteins and Membranes. Elsevier Science Publishers. pp. 141-169.

CROSSA, J. 1990. Statistical analyses of multilocation trials. *Advances in Agronomy* 44: 55-85.

CRUZ, R.M., 1992. More about the multiplicative model for the analysis of genotype-environment interaction. *Heredity* 68: 135-140.

DeLACY, I.H., M. COOPER, and K.E. BASFORD, 1996. Relationships among analytical methods used to study genotype-environment interactions and evaluation of their impact on response to selection. *In*: M.S. Kang and H.G. Zobel Jr (Eds.) Genotype-by-environment interaction, CRC Press; Boca Raton, New York. pp. 51-84.

DEMISSIE, A., 1999. *In situ* conservation: the Ethiopian experience. *ILEIA* Newsletter No. 15(3&4): 30-31.

DIAS, J.S., A.A. MONTERIO, and M.B. LIMA, 1993. Numerical taxonomy of Portuguese Tronchuda cabbage and Galega kale land races using morphological characters. *Euphytica* 69: 51-68.

DOWNEY, R.K., 1990. Canola: A quality Brassica oilseed. *In*: Janick, J. and SIMON, J.E. (eds.). Advances in new crops. Timber Press, Portland OR, pp 211-217.

DUDLEY, J.W., and R.H. MOLL, 1969. Interpretation and use of estimates of genetic variance in plant breeding. *Crop Sci.* 9: 257-262.

EBERHART, S.A., and W.A. RUSSELL, 1966. Stability parameters for comparing varieties. *Crop Sci.* 6: 36-40.

FALK, K., 1999. Development of early maturing *Brassica carinata* for western Canada. *In*: *Proceedings of the 10th International Rapeseed Congress*, Canberra, Australia, 26-29 September 1999.

FERERES, E., J. FERNÁNDEZ-MARTÍNEZ, I. MINGUEZ, and J. DOMÍNGUEZ, 1983. Productivity of *B. juncea* and *B. carinata* in relation to rapeseed. *In:* Int. Rapeseed Conf., 6th, Paris, France, 17-19 May 1983. GCIRC, Paris France. pp. 293-298

FERNÁNDEZ-MOYA, V., E. MARTÍNE-FORCE, and R. GARCÉS, 2000. Identification of triacylglycerol species from high-saturated sunflower (*Helianthus annuus*) mutants. *J. Agric. Food Chem.* 48:764-769.

FINLAY, K.W., and G.N. WILKINSON, 1963. The analysis of adaptation in plant breeding programme. *Aust. J. Agric. Res.* 14: 742-754.

FISHER, R.A., 1918. The correlation between relatives on the supposition of Mendelian inheritance. *Trans. R. Soc. Edinburgh* 52: 399-433.

FISHER, R.A., 1925. Statistical methods for the analysis of genotype environment interaction. *Heredity* 31: 339-354.

FREEMAN, G.H., 1973. Statistical methods for the analysis of genotype-environment interactions. *Heredity* 31: 339-354.

FREEMAN, G.H., 1990. Modern statistical methods for analysing genotype x environment interactions. *In*: Kang, M.S. (ed.). Genotype-by-environment interaction and plant breeding. Louisiana State University Agricultural Centre, Baton Rouge, La., pp. 118-125.

FREEMAN, G.H., and J.M. PERKINS, 1971. Environmental and genotype–environmental components of variability. VIII. Relation between genotypes grown in different environments and measures of these environments. *Heredity* 27: 15-23.

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

GAIL, M., and R. SIMON, 1985. Testing for qualitative interactions between treatment effects and patient subsets. *Biometrics* 41: 361-372.

GAUCH, H.G., 1982a. Noise reduction by eigenvector ordinations. *Ecology* 63: 1643-1649.

GAUCH, H.G., 1982b. Multivariate analysis in community ecology, 1st ed. Cambridge Univ. Press, London and New York.

GAUCH, H.G., 1988. Model selection and validation for yield trials with interaction. *Biometrics* 44: 705-715.

GAUCH, H.G., and R.W. ZOBEL, 1996. AMMI analysis of yield trials. *In*: M. S. Kang and H.G. Gauch, Jr. (eds). Genotype by environment interaction. pp. 85-122.

GETINET A., G. RAKOW and R.K. DOWNEY, 1996. Agronomic performance and seed quality of Ethiopian mustard in Saskatchewan. *Can. J. Plant Sci.* 76: 387-392.

GOMEZ-CAMPO, C., and M.E. TORTOSA, 1974. The taxonomic and evolutionary significance of some juvenile characters in *Brassicae*. *Bot. J Linn Soc* 69: 105-124.

GOMEZ-CAMPO, C., and S. PRAKASH, 1999. Origin and domestication of the Brassica. *In*: GOMEZ-CAMPO, C. (ed.). Biology of *Brassica* Coenospecies. Elsevier, Amsterdam, pp 33-58.

GOWER, J.C., 1966. Some distance properties of latent roots and vector methods used in multivariate analysis. *Biometrika* 53: 325-338.

GUPTA, P.K., R.K. VARSHNEY, P.C. SHARMA, and B. RAMESH, 1999. Molecular markers and their applications in wheat breeding. *Plant Breeding* 118: 369-390.

GUPTA, V.P., M.S. SEKHON, and D.R. SATIJA, 1991. Studies on genetic diversity, heterosis and combining ability in Indian mustard (*Brassica juncea* L. Cezern & Coss.). *Indian J. Genet.* 51: 448-453.

HARWOOD, J.L., 1996. Recent advances in the biosynthesis of pant fatty acids. *Biochimica et Biophysica Acta* 1301: 7-56.

HEMINGWAY, J.S., 1995. Mustards: *Brassica* spp. and *Sinapis alba* (Cruciferae). *In*: Smart, J., and Simmonds N.W. (eds.) Evolutions of crop plants, 2nd Edition, Longman Scientific and Technical, Singapore, pp. 82-88.

HILL, L., 1975. Genotype-environment interactions-a challenge for plant breeding. *J. Agric. Sci.* 85: 477-493.

HITCHCOCK, C., 1975. Structure and distribution of plant acyl lipids. *In*: Galliard, T., and E.I. (eds.). Recent advances in the chemistry and biochemistry of plant lipids. Academic Press, London, New York, pp. 1-19.

HOHLS, T., 1995. Analysis of genotype environment interactions. *S. Afr. J. Sci. 91*: 121-124.

HORROBIN, D.F., 1992. Nutritional and medical importance of gamma-linolenic acid. *Prog. Lipid Res.* 31:163-194.

HÜHN, M., 1990. Nonparametric measure of phenotypic stability. Part 1: Theory. *Euphytica*, 47:189-194.

HÜHN, M., 1996. Nonparametric analysis of genotype x environment interaction by ranks. *In*: M.S. Kang, H.G. Gauch, Jr. (Eds.). Genotype by environment interaction. pp. 235-271.

JACKSON, P., M. ROBERTSON, M. COOPER, and G.L. HAMMER, 1996. The role of physiological understanding in plant breeding: From a breeding perspective. *Field Crops Res.* 49: 11-37.

KANG, M.S., 1990. Understanding and utilization of genotype-by-environment interaction in plant breeding. *I.n*: Kang, M.S. (ed.), *Genotype by environment interaction and plant breeding*. Louisiana State University Agricultural Centre, Baton Rouge, Louisiana, USA. pp. 52-68.

KEBEBEW, A., A. MERKER and H. TEFERA, 2002. Quantitative trait variation in Tef [*Eragrostis tef* (Zucc.) Trotter] germplasm from western Ethiopia. *Euphytica* 127: 399-410.

KNOWLES, P.F., T.E. KEARNEY, and D.B. COHEN, 1981. Species of rapeseed and mustard as oil crops in California. *In*: Pryde (ed). New sources of fats and oils. Am. Oil Chem. Soc. Champain, IL, pp 255-268.

KOEBNER, R.M.D., K.M. DEVOS, and M.D. GALE, 1994. Advances in the application of genetic markers in plant breeding. Asian Seed 1994, Chiang Mai, Thailand.

KROON, J.P.M.DE and P. VAN DER LAAN, 1981. Distribution-free test procedures in two-way layouts; a concept of rank-interaction. *Statistica Neerlandica* 35: 189-213.

KUMAR, L.S., 1999. DNA markers in plant improvement: An overview. *Biotechnology Advances* 17: 143-182.

LEGENDRE. L., and P. LEGENDRE, 1983. Numerical Ecology: Developments in environmental modelling 3. Elsevier Scientific Publishing Company.

LI, Y., S. WU, and Y. CAO, 1995. Cluster analysis of an international collection of foxtail millet (*Setaria italica*. (L) P.Beauv.). *Euphytica* 83: 79-85.

LIN, C.S. and M.R. BINNS, 1988. A superiority measure of cultivar performance for cultivar x location data. *Can. J. Plant Sci.* 68: 193-198.

LIN, C.S., M.R., BINNS, and L.P. LEFKOVICH, 1986. Stability analysis: where do we stand? *Crop Sci.* 26: 894-900.

LLOYD, M., J.H. ZAR, and J.R. KARR, 1968. On the calculation of information-theoretical measures of diversity. *Amer. Midland Natur.* 79:257-272.

MALIK, R.S., 1990. Prospects for *Brassica carinata* as an oilseed crop in India. *Experimental Agriculture* 26: 125-129.

MEAD, J.F., 1984. The non-eicosanoid function of the essential fatty acids. *J. Lipid Res.* 25: 1517-1521.

MENSINK, R.P., and M.B. KATAN, 1989. Effect of a diet enriched with monounsaturated or polyunsaturated fatty acids on levels of low-density and high-density lipoprotein cholesterol in healthy women and men. *New England Journal of Medicine* 321: 346-441.

MENSINK, R.P., E.H.M., TEMME, and G. HORNSTRA, 1994. Dietary saturated and trans fatty acids and lipoprotein metabolism. *Annals of medicine* 56:461-464.

MILLAN, T.T., and A.M. CUBERO, 1995. Identification of Rosa by isozymes and RAPD markers. Second International Rosa Symposium ISHS/INRA Antibes, France.

MIZUSHIMA, U., 1980. Genome analysis in *Brassica* and allied genera. *In*: Tsunoda, S., Hinata, K. and Gomez-Campo C., (eds.). Brassica crops and wild allies: Biology and breeding. Japan Scientific Societies Press, Tokyo, pp 89-106.

MNZAVA, N.A., and K. OLSSON, 1990. Studies on tropical vegetables: Part 1: Seed amino, fatty acid and glucosinolate profile of Ethiopian mustard (*Brassica carinata* A. Braun). *Food Chemistry* 35: 229-235.

MULLIS, K., F. FALOONA, S. SCHARE, R. SAIKI, G. HORN and H. ERLICH, 1986. Enzymatic amplification of DNA in vitro: the polymerase enzyme chain reaction. *Cold Spring Harbor Symp. Quant. Biol.*, 51: 263-273.

MUNGOMERY, V.E., R. SHORTER, and E. BYTH, 1974. Genotype x environment interactions and environmental adaptation. I. Pattern analysis- applications to soybean populations. *Aust. J. Agric. Res.* 25: 59-72.

MURPHY, D.J., 1994. Designer oil crops. VCH, Weinheim.

NASSAR, R., and M. HÜHN, 1987. Studies on estimation of phenotypic stability: Test of significance for nonparametric measures of phenotypic stability. *Biometrics* 43: 45-53.

PADLEY, F.B., F.D. GUNSTONE, and J.L. HARWOOD, 1994. Occurrence and characteristics of oils and fats. *In:* F.D. Gunstaone, J.L. Harwood and F.B. Padley (eds). The Lipid Hand Book, London: Chapman and Hall, pp. 47-223.

PERKINS J.M., 1972. Principal component analysis of genotype environment interactions and physical measures of the environment. *Heredity* 29: 51-70.

PERKINS, J.M., and J.L. JINKS, 1968. Environmental and geneotype x environmental components of variability. III. Multiple lines and crosses. *Heredity* 23: 339-356.

PERRY, M.C., and M.S. MCINTOSH, 1991. Geographic patterns of variation in the USDA soybean germplasm: I. Morphological traits. *Crop Sci.* 31:1350-1355.

PIELOU, E.C., 1966. Measurment of diversity in different types of biological collections. *J. Theor. Biol.* 13:131-144.

PINTHUS, M.J., 1973. Estimates of genotypic value: A proposed method. *Euphytica,* 22: 121-123.

PLAISTED, R.L., 1960. A shorter method for evaluating the ability of selections to yield consistently over locations. *Am. Potato J.* 37: 166-172.

PLAISTED, R.L., and L.C., PETERSON, 1959. A technique for evaluating the ability of selections to yield consistently in different locations or seasons. *Am. Potato J.* 36: 381-385.

POLIGNANO, G.B., E. ALBA, P. UGGENTI and G. SCIPPA, 1999. Geographical patterns of variation in Bari faba bean germplasm collection. *Genetic Resources and Crop Evolution* 46: 183-192.

POLIGNANO, G.B., P. UGGENTI, and P. PERRINO, 1989. Pattern analysis and genotype x environment interactions in faba bean (*Vicia faba* L.) populations. *Euphytica 40*: 31-41.

RABBANI, M.A., A. IWABUCHI, Y. MURAKAMI, T. SUZUKI and K. TAKAYANAGI. 1998. Phenotypic variation and the relationships among mustard (*Brassica juncea* L.) germplasm from Pakistan. *Euphytica* 101(3): 357-366.

RAFALSKI, J.A., M. MORGANTE, W. POWELL, J.M. VOGEL, and S.V. TINGEY, 1996. Generating and using DNA markers in plants. *In*: B. Birren, and E. Lai (eds). Analaysis of Non-mammalian Genomes: a Practical Guide. Acadamic Press, Boca Raton. pp. 75-134.

RENAUD, S.C., J.R. RUF, and D. PETITHORY, 1995. The positional distribution of fatty acids in palm oil and lard influences their biological effect in rats. *J. Nutrition* 125-229-237.

REVILLA, P., and W.C. TRACY, 1995. Morphological characterization and classification of open-pollinated sweet corn cultivars. *J. Amer. Soc. Hort. Sci.* 120: 112-118.

RÖBBELEN, G., 1981. Breeding for low content of glucosinolate in rapeseed. *In*: Bunting, E.S. (ed). World crops: Production and utilization of protein in oilseed crops. Martinus Nijhoff, the Netherlands, pp. 91-106.

RÖBBELEN, G., and W. THIES, 1980. Biosynthesis of seed oil and breeding improved oil quality of rapeseed. *In*: S. Tsunoda, K. Hinta, and C. Gomez-Campo (eds.), *Brassica* Crops and wild Allies: Biology and Breeding, Japan Scientific Societies Press, Tokyo. pp. 235-283.

ROMAGOSA, I., and P. N. FOX, 1993. Genotype x environment interaction and adaptation. *In*: M.D. Hayward, N.O. Bosemark & I. ROMAGOSA (Eds.), Plant breeding: principles and prospects, Chapman and Hall, London, pp. 373-390.

SAUER, F.D. and J.K.G. KRAMER, 1983. The problems associated with the feeding of high erucic acid rapeseed oils and some fish oils to experimental animals. *In*: J.K.G. Kramer, F.D. Sauer, and W.J. Pigden (eds), High and Low Erucic Acid Rapeseed Oils, Academic Press, New York. pp. 253-292.

SEEGLER, C.J.P., 1983. Oil plants of Ethiopia: Their taxonomy and agricultural significance. Centre for agricultural publishing and documentation, PUDOC, Wageningen the Netherlands.

SHAHIDI, F., 1996. Natural antioxidants: an overview. In *Natural Antioxidants. Chemistry, Health effects, and applications*, ed., F. Shahidi. Champaign, IL, USA:AOCS Press, pp. 1-11.

SHANNON, C.E., 1948. A mathematical theory of communication. *Bell's System Technical Journal* 27:379-423.

SHANNON, C.E., and W. WEAVER, 1949. The mathematical theory of communication. Univ. Illinois Press. Urbana. Chicago, London.

SHIGESABURO-TSUNODA, 1980. Eco-physiology of wild and cultivated forms in Brassica and allied genera. *In*: Röbelen, G., Downey R.K. and Ashri, A. (eds.). Oil crops of the world. McGraw-Hill New York, p 109-120.

SHORTER, R., D.E. BYTH, and V.E. MUNGOMERY, 1977. Genotype x environment interactions and environmental adaptation. II. Assessment of environmental contribution. *Aust. J. Agric. Res.* 28: 223-235.

SHUKLA, G.K., 1972. Some statistical aspects of partitioning genotype-environmental components of variability. *Heredity* 29: 237-245.

SIMMONDS, N.W., 1979. Principles of crop improvement. Longman, London.

SMITH, J.S.C., and O.S. SMITH, 1989. The description and assessment of distances between inbred lines of maize: The utility of morphological, biochemical and genetic

descriptors and a scheme for the testing of distinctiveness between inbred lines. *Maydica* 34: 151-161.

SMITH, S.E., L. GUARINO, A.L. DOSS and D.M. CONTA, 1995. Morphological and agronomic affinities among Middle Eastern alfalfa's-accession from Oman and Yemen. *Crop Sci.* 35: 1188-1194.

STYMNE, S., and A.K. STOBART, 1987. Triacylglycerol biosynthesis. *In*: P.K. Stumpf (ed). The Biochemistry of plants, vol.9. Orlando, USA: Academic Press, pp. 175-214.

TAI, G.C.C., 1971. Genotype stability analysis and its application to potato regional trials. *Crop Sci.* 11: 184-190.

TAKAHATA, Y., and K. HINATA, 1986. A consideration of the species relationships in subtribe *Brassicinae*: *(Cruciferae)* in view of cluster analysis of morphological characters. *Pl. Sp. Biol.* 1: 79-88.

TATINENI, V.R., G. CANTREIL, and D.D. DAVIS, 1996. Genetic diversity in elite cotton germplasm determined by morphological characteistics and RAPDs. *Crop Sci.* 36: 186-192.

TATUM, V., and C.K. CHOW, 1992. Effects of processing and storage on fatty acids in edible oils. *In:* C.K. Chow (ed). Fatty acids in Foods and Their Health Implications. New York: Marcel Dekker, Inc., pp. 337-351.

TAUTZ, D., 1989. Hypervariability of simple sequences as general sources for polymorphic DNA markers. *Nucleic Acids Res.* 17(16): 6463-6471.

TAYLOR, D.C., S.L. MACKENZIE, A.R. MCCURDY, P.B.E. MCVETTY, E.M. GIBLIN, E.W. PASS, S.J. STONE, R. SCARTH, S.R. RIMMER, and M.D. PICKARD, 1994. Stereospecific analyses of seed tricylglycerol from high-erucic Brassicaceae: detection of erucic acid at sn-2 position in *Brassica oleracae* L. genotypes. *Journal of the American Oil Chemists' Society* 71: 163-167.

THIES, W., 1998. Biochemisch-physiologische Grundlagen der Qualitäts-Pflanzenzüchtung.Institut fÜr Pflanzenbau und Pflanzenzüchtung, Göttingen, Germany.

THORMANN, C.E., M.E. FERREIRA, L.E.A. CAMARGO, J.G. TIVANG, J.G. and T.C. OSBORN, 1994. Comparison of RFLP and RAPD markers to estimating genetic relationships within and among cruciferous species. *Theor. Appl. Genet.* 88: 973-980.

THOTTAPPILLY, G., H.D. MIGNOUNA and O.G. OMITOGUN, 2000. The use of DNA markers for rapid improvement of crops in Africa. *African Crop Science Journal* 8(1): 99-108.

UN, 1935. Genome analysis in *Brassica* with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. *Jap. J. Bot.* 9:389-452.

UPADHYAYA, H.D., P.J. BRAMEL, R. ORTIZ and S. SINGH, 2002. Geographical patterns of diversity for morphological and agronomic traits in groundnut germplasm collection. *Euphytica* 128: 191-294.

VAN EEUWIJK, F.A., 1995. Multiplicative interaction in generalized linear models. *Biometrics* 51: 1017-1032.

VELASCO, L., and J.M. FERNÁNDEZ-MARTÍNEZ, 2002. Breeding oilseed crops for improved oil quality. *In*: A.S. Basra and L.S. Randhawa (eds.). Quality improvement in field crops. Hawworth Press, Inc. New York, USA. pp 309-344.

VODENICHAROVA, M., 1989. Use of proteins as molecular-genetic markers in plants. *Genet Sel.* 22: 269-277.

VOS, P., R. HOGERS, M. BLEEKER, M. REIJANS, T. VAN DE LEE, A. FRIJTERS, J. POT, J. PELEMAN, M. KIPER, and M. ZABEAU, 1995. AFLP: A new technique for DNA fingerprinting. *Nucleic Acids. Res.* 23(21): 4407-4414.

WESTCOTT, B. 1986. Some methods of analysing genotype x environment interaction. *Heredity* 56: 243-253.

WESTCOTT, B., 1987. A method of assessing the yield stability of crop genotypes. *J. Agric. Sci.* 108: 267-274.

WESTMAN, AL., and S. KRESOVICH, 1997. Use of molecular marker techniques for description of plant genetic variation. *In*: Callow, J.A., B.V. Ford-Llyd and H.J. Newbury (eds). Biotechnolgy and plant genetic resources. CAB International, UK. Pp. 9-48.

WESTPHAL, A., and R. MARQUARD, 1980. Yield quality of *Brassica* species in Ethiopia. *Plant Res. Dev.* 13: 114-127.

WILLETT, W.C. and A. ASCHERIO. 1994. Trans fatty acids: are the effects only marginal? *American journal of Public* Health. 84: 722-724.

WILLIAMS, G.K., A.R. KUBELIKE, K.J. LIVAK, J.A. RAFALSKI and S.V. TINGEY, 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res.* 18: 6531-6535.

WILLIAMS, W.T., 1976. Pattern analysis in agricultural sciences. CSIRO, Melbourne and Elsevier, Amesterdam.

WINTER, P., and G. KAHL, 1995. Molecular marker technologies for plant improvement. *World Journal Microbiology and Biotechnology* 11: 438-448.

WOREDE, M., 1993. The role of Ethiopian farmers in conservation and utilization of crop genetic resources. *In*: D.R. Buxton *et al.* (eds.). International crop Science Society of America, Madison, WI: Crop science of America.

WOREDE, M., T. TESEMMA, and R. FEYISSA, 2000. Keeping diversity alive: an Ethiopian perspective. *In*: S.B. Brush (ed.). Genes in the Field: On-farm Conservation of Crop Diversity, S.B. Brush (ed.), Lewis Publishers, Boca Raton. pp. 143-161.

WRICKE, G., 1962. Über eine Methode zur Erfassung der Ökologischen Streuberite in Feldversuchen. *Z. Pflanzenzüchtg.* 47: 92-96.

WRICKE, G., 1965. Die erffassung der Wechselwirkung zwischen genotyp und umwelt bei qquantitativn eigenschaften. *Z. Pflanzenzüchtg.* 53: 266-343.

YAN, W., and L.A. HUNT, 1998. Genotype-by-environment interaction and crop yield. *Plant Breed. Rev.* 16: 35-178.

YAN, W., and L.A. HUNT, 2001. Interpretation of genotype x environment interaction for winter wheat yield in Ontario. *Crop Sci.* 41: 19-25.

YANG, R.C., S. JANA, and J.M. CLARK, 1991. Phenotypic diversity and associations of some potentially drought-responsive characters in durum wheat. *Crop Sci.* 31: 1484-1491.

YATES, F., and W.G. COCHRAN, 1938. The analysis of groups of experiments. *J. Agric. Sci. Camb.* 28: 556-580.

YAU, S.K., 1995. Regression and AMMI analyses of genotype x environment interactions: an empirical comparison. *Agron. J.* 87: 121-126.

YODICE, R., 1990. Nutritional and stability characteristics of high oleic sunflower seed oil. *Fat Science and Technology* 92: 121-126.

YOON, M.S., J.W. AHN, S.J. PARK, H.J. BAEK, N.K. PARK and Y.D. RHO, 2000. Geographical patterns of morphological variation in soybean germplasm. *Korean J. Crop Sci.* 45(4): 267-271.

ZAR, J.H. 1984. Biostatistical analysis. Prentice-Hall, INC., Englewood Cliffs, N.J. New Jersey. Second Edition. pp. 718.

ZOBEL, R.W., M.J. WRIGHT, and H.G. GAUCH, 1988. Statistical analysis of a yield trial. *Agron. J.* 80: 388-393.

CHAPTER III

GEOGRAPHICAL PATTERNS OF MORPHOLOGICAL VARIATION IN ETHIOPIAN MUSTARD GERMPLASM COLLECTIONS

3.1. ABSTRACT

Two hundred and fifty eight accessions of Ethiopian mustard accessions randomly sampled from different geographical areas of the country were evaluated in a field test at Adet, Ethiopia in 2001, to estimate the phenotypic diversity index (H') for six morphological traits (leaf colour, leaf size, stem colour, flower colour, angle of branching, and seed colour). The Shannon Weaver diversity indices (H') were calculated for the traits, geographical regions, and altitudinal classes. Mean diversity index ($\bar{H}'$) across regions for Ethiopia was 0.72±0.06. Polymorphism was high for leaf size (H'=0.95) and stem colour (H'=0.84) and relatively low for angle of branching (H'=0.54). The chi-square (χ^2) analysis indicated that some of the regions had significant phenotypic deviations from expected frequencies for most traits. The mean diversity index for the country across altitudinal class was ($\bar{H}' = 0.71 \pm 0.07$). The highest mean diversity index ($\bar{H}' = 0.75 \pm 0.07$) was recorded for areas ranging in altitude from 2600-2800 m. A decline in mean diversity with increasing elevation was also noted. Based on the characters studied and small collections evaluated, the centres of diversity for B. carinata appears to be Shewa, Bale and Arsi provinces with due consideration to altitudinal parameters. Hence, further exploration and in situ conservation should be undertaken in these regions.

Keywords: *Brassica carinata* , diversity index, Ethiopian mustard, germplasm

3.2. INTRODUCTION

A wide genetic base of germplasm is a prerequisite to the success of plant breeding programmes and to cope with unforeseen breeding challenges in a changing environment. Landraces of crop species have been the priority targets of collection

since Vavilov's expedition to various parts of the world. Although what and how much to conserve has been controversial, there has been a general agreement that landraces should be conserved, either *in situ* or *ex situ* (Bechere *et al.*, 1996).

The determination of the extent of phenotypic and genotypic diversity in a germplasm collection provides an indication of redundancies or exclusion and can aid in managing collections. The information on diversity analysis also, could be used to devise a core collection. Further, easily observable morphological traits are generally used to characterize genetic resources (Perry and Mcintosh, 1991).

One approach for building a germplasm collection is to collect material from diverse geographical origins with a concentration of accessions from proposed centres of diversity. This should capture inherent and unexploited diversity in the individual samples. Representative samples from the complete range of geographical areas of the crop species can help to ensure that co-adapted gene complexes or correlated adaptations are conserved (Frankel and Soule, 1981). The maximum genetic conservation would be achieved by sampling populations from as many distinct environments as possible (Brown, 1978).

In order to determine the distribution of diversity in germplasm, phenotypic and genotypic variation within and between countries and regions of origin have been examined for several crop species including: soybean [*Glycine max* (L.) Merr.] (Perry and Mcintosh, 1991; Yoon *et al.*, 2000;); wheat (*Triticum turgidum* L.) (Jain *et al.*, 1975; Negassa, 1986a; 1986b; Bechere *et al.*, 1996); barley (*Hordeum vulgare L.*) (Tolbert *et al.*, 1979; Bekele, 1983; Negassa, 1985; Asfaw, 1989; Engeles, 1994; Demissie and Björnstad, 1996); sorghum (*Sorghum bicolour* (L.) Moench) (Teshome *et al.*, 1997; Ayana and Bekele, 1999); noug (*Guizotia abyssinica* Cass.) (Genet and Belete, 2000): tef (*Eragrostis tef (Zucc.) Trotter*) (Bekele, 1996; Kefyalew *et al.* 2000; Kebebew *et al.*, 2002); Bari faba bean (*Vicia faba*) (Polignano *et al.* 1999); cacao (*Theobroma cacao* L.) (Bekele and Bekele, 1996) and groundnut (*Arachis hypogae* L.) (Upadhyaya *et al.*, 2002).

There has been increasing interest in recent years in *in-situ* conservation to complement *ex situ* conservation efforts. Choice of sites for *in situ* conservation may be based on high diversity estimates based on markers, and knowledge of adaptive traits linked to certain ecological conditions, e.g., co-evolving host-pathogen systems and adaptation to other stress conditions. Under appropriate ecological and environmental conditions, *in situ* populations of the crop species may not only maintain a high level of variation but also a high frequency of desirable genes (Leur *et al.*, 1989).

Information on the extent and patterns of distribution of genetic variation of a crop species is essential for (1) effective utilization of germplasm in plant breeding programmes (2) devising appropriate sampling procedures for germplasm collection and conservation (3) obtaining core collections for efficient germplasm management and (4) elucidating the taxonomy, evolution and origin of the crop (Allard, 1970; Marshall and Brown, 1975; De Wet *et al.*, 1976; Brown, 1989; Hayward and Breese, 1993; Moreno-Gonzalez and Cubero, 1993).

The Ethiopian or Abyssinian region characterized by a wide range of agro-climatic conditions, often referred to as a major Vavilovian gene centre provides tremendous crop diversity. The need for the conservation of this heritage of genetic stock has been emphasized by several workers (Vavilov, 1951; Harlan, 1969; Zohary, 1970; Mengesha, 1975; Bekele, 1983; Worede, 1988).

Ethiopian mustard *(Brassica carinata*, n=2x=17, BBCC) which has been grown in the Ethiopia since the 4^{th} to 5^{th} Millennia BC (Simmonds, 1979) is one of the oil crops found in great diversity. It is believed to have evolved in the Ethiopian plateau through natural hybridisation of *B. nigra* (n=8, BB) with *B. oleraceae (*n=9, CC) and followed by the chromosome doubling of the hybrid plant (UN, 1935). Wild forms of *B. carinata* have not been reported, but there are diverse ecotypes with morphological and agronomic differences (Abebe *et al.*, 1992).

In its native Ethiopia *B. carinata* is high yielding, resistant to diseases and insect pests, has seed shattering resistance, and tolerance to semiarid conditions (Knowles *et al.*, 1981). In order to use such important genes for germplasm conservation and plant

breeding, quantifying the genetic diversity and studying the eco-geographic pattern of the variability should be of prime importance.

Hence, this work addresses quantifying the morphological diversity in Ethiopian mustard landraces and presents the pattern and estimates of phenotypic diversity based on six morphological traits with respect to geographical origin and altitude class. It tentatively locates areas where significant morphological variation exists for germplasm collection and *in situ* conservation.

3.3. MATERIALS AND METHODS

The study was conducted at Adet Research Centre (longitude 37° 29'E and latitude 11° 16'N and 2240 meters above sea level (masl)) located in north-western Ethiopia during the 2001 main cropping season (keremet). For this study a total of 258 accessions of Ethiopian mustard, collected from different regions of the country (Figure 3.1), which were chosen randomly from 1295 accessions maintained in the Biodiversity Conservation and Research Institute of Ethiopia, were used. They were collected randomly from varying agro-ecological areas of the country and preserved, at the Biodiversity Conservation and Research Institute of Ethiopia since its establishment in 1976. The original samples were collected from farmers' fields, by the use of random sampling technique (Hawkes, 1976).

All accessions were planted on well-drained red soils (Nitosols) in a single row of 4 m length plot, with a spacing of 0.6 m between rows and 10 cm between plants. Fertilizer was applied at the rate of 46/69 kg/ha of N/P_2O_5 at the time of planting. Weeding was carried out 30 days after emergence and again 45 days after the first weeding.

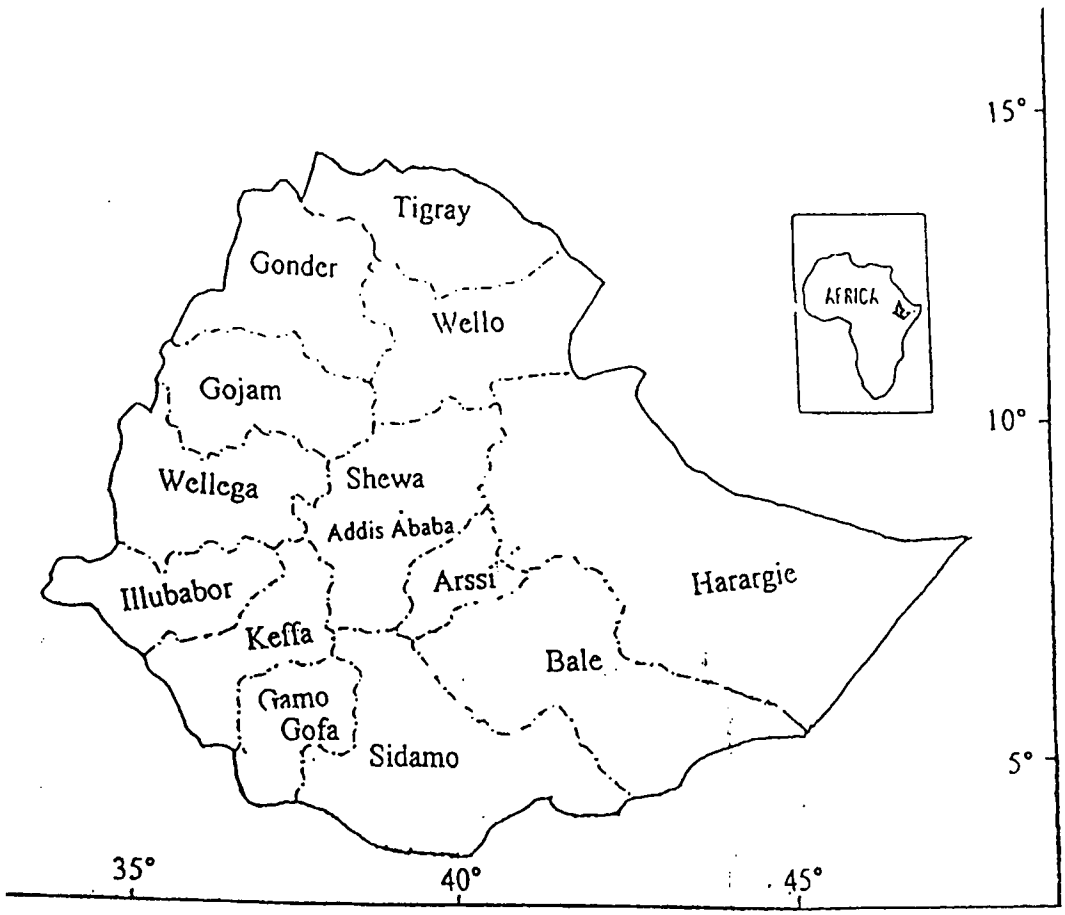


Figure 3.1. Map of Ethiopia depicting regions from where the Ethiopian mustard accessions were collected. (Solid lines represent international boundaries, while broken lines show boundaries between previous provinces).

Plants were scored for six morphological characters (leaf colour, leaf size, stem colour, flower colour, angle of branching, and seed colour) using the Gomenzer description format of Biodiversity Conservation and Research Institute of Ethiopia (Table 3.1). Leaf colour is the most predominant colour of the leaf during 50% flowering. Allowable states are light, medium, dark, green, slightly purple, and intensive purple. Leaf size is a visual measurement taken at the middle part of the plant at 50% flowering. Allowable states are small and narrow, intermediate, and large and broad. Stem colour is the most predominant colour of the stem recorded at 50% flowering. Allowable states are light, medium, dark, green, slightly purple, and intensive purple. Flower colour is the most predominant colour of the flower during

50% flowering. Allowable states are yellow, cream and purple. Angle of branching is the visual measurement taken at 50% flowering at the middle part of the plant. Allowable states are very erect, intermediate and right angle or more. Seed colour is the most predominant colour of the seed recorded after-harvesting. Allowable states are light brown, dark brown, brown, yellow and black. All traits were evaluated in the field for all plants within each row.

Table 3.1. Phenotypic classes of the morphological characters used for the diversity study of *B. carinata* germplasm.

Characters	Abbreviations	No. Classes	Character states
Leaf colour	LC	6	1. Light green 2. Medium 3. Dark green 4. Green 5. Slightly purple 6. Intensive purple
Leaf size	LS	3	1. Small and narrow 2. Intermediate 3. Large and broad
Stem colour	SC	6	1. Light 2. Medium 3. Dark 4. Green 5. Slightly purple 6. Intensive purple
Flower colour	FC	3	1. Yellow 2. Cream 3. Purple
Angle of branching	AB	3	1. Very erect 2. Intermediate 3. Right angle or more
Seed colour	SEC	5	1. Light brown 2. Dark brown 3. Brown 4. Yellow 5. Black

Geographical region, and altitude were used as classifying variables. The country was divided arbitrarily into four regions: north-west and eastern Ethiopia (Gonder, Gojam, and Wello), since the materials from Tigray province were small they were included together with Gonder, western Ethiopia (Wellega, Illubabor, and Keffa), south-weastern Ethiopia (Gamo Gofa, Sidamo, and Bale), and central and eastern Ethiopia (Shewa, Arsi, and Harargie). The altitude range was arbitrarily classified into eight altitude classes, viz., <1800, 1800-2000, 2000-2200, 2200-2400, 2400-2600, 2600-2800, 2800-3000 and >3000 masl.

The χ^2 analysis was performed to test deviations from the expected frequencies of each trait. The diversity index (H') of Shannon and Weaver (1949) was used as a measure of phenotypic diversity for each trait. The index was estimated for each character over all accessions and for characters within regions and altitudinal classes. This index was calculated as follows:

$$H' = -\sum_{i=1}^{n} p_i \log_e p_i ,$$

Where, n is the number of classes of traits and p_i is the proportion of accessions in the ith class of a trait. Each H' value was normalized by dividing it by its maximum value ($\log_e n$), which ensured that all H' values were in the range of 0 to 1. The choice of the base of the logarithm is open to choice, but in most cases the base e should be used.

The average diversity ($\bar{H}'$) over n traits was estimated as:

$$\bar{H}' = \sum H' / n$$

The variance is given by the following expression:

$$\text{var}(H') = \frac{\sum_{i=1}^{n} p_i \log_e{}^2 p_i - (\sum_{i=1}^{n} p_i \log_e p_i)^2}{N} + \frac{n-1}{2N^2}$$

Test for differences among pairs of H' values for groups of entries or regions can be compared by t-tests according to Jain *et al.* (1975).

$$t = (H'_1 - H'_2) / [\text{var}(H'_1) + \text{var}(H'_2)]^{1/2},$$

The degrees of freedom of the test is:

$$df = [\text{var}(H'_1) + \text{var}(H'_2)]^2 / [\text{var}(H'_1)^2 / N_1 + \text{var}(H'_2)^2 / N_2]$$

Where N_1 and N_2 are the number of entries used in calculating $\text{var}(H'_1)$ and $\text{var}(H'_2)$ respectively.

The distribution of H' is described by Bowman *et al.* (1971), and this statistic is widely used to estimate diversity in germplasm collections of economic crops (Tolbert *et al.*, 1979). The interpretation is that the greater the number of variants in each phenotypic class of a given character and the more equal their proportions, the greater is the diversity (Pielou, 1966). The minimum value of the index is zero for a

monomorphic population. The value of the index increases with the increase in polymorphism and reaches the maximum value when all phenotypic classes have equal frequencies (Yang *et al.*, 1991)

3.4. RESULTS AND DISCUSSION

Regional trait distribution

The phenotypic frequencies for individual characters and provinces as percentage of the number of accessions from each geographic origin are presented in Table 3.2. The areas of origins were grouped into regions and regions into provinces and weighted regional frequencies were computed.

Yellow flower types dominated in all regions. Only the material from Keffa displayed a high proportion of cream flower phenotypes. Intermediate types of angle of branching dominated all regions and the very erect type was higher in Wellega. The highest numbers of intermediate type of angle of branching were recorded from Harargie (93%) and Gamo Gofa (92%). The most predominant type of leaf colour was green in all provinces and the entire collection.

The accessions were dominated by green and dark green phenotypes of leaf colour. Medium types were detected in Wellega, Sidamo, and Shewa provinces. Intensive purple types were also detected in Harargie province.

Keffa province had equal proportions of intermediate and large and broad types of leaf size. The intermediate type of leaf size is dominating in almost all regions. All accessions from Gamo Gofa had large and broad types of leaf size.

Stem colour is widely distributed throughout the collection, i.e., 12% (light), 6% (medium), 10% (dark), 42% (green), 26% (slightly purple) and 4% (intensive purple) types of stem colour. Slightly and intensive purple types of stem colour were predominantly found in all provinces.

The phenotypic classes of flower colour types (yellow and cream) occurred in all regions, with the yellow type being the most predominant type in all regions. The

third class of flower colour (purple) occurred only in Shewa province. At regional level the largest representation of green leaf colour was noted in western Ethiopia, followed by southern Ethiopia.

Seed colour was most variable in the Bale and Arssi provinces. Accessions from both these provinces also had the highest proportions of light seed colours. Black seed colour was noted from accessions of Arssi and Bale provinces only. The highest proportion of yellow phenotypes occurred in Wellega, Arssi, and Bale provinces. Brown and dark seed colour phenotypes were found in all regions, with light seed colour phenotype being predominant in Bale province.

South-western Ethiopia showed expected frequencies for most traits, i.e., stem colour, flower colour, and seed colour whereas leaf colour and leaf size displayed significant deviations in chi-square tests. On the other hand, central and eastern Ethiopia displayed expected frequencies for most of the characters except stem colour and seed colour, which showed highly significant deviations in chi-square tests (Table 3.4).

Altitudinal trait distribution

Table 3.3 summarizes the phenotypic frequencies for individual characters and altitudinal classes as percentages of the number of genotypes from each altitude class by pooling populations together.

The frequency of medium phenotype of leaf colour increases with altitude. The green phenotype was the most predominant phenotype in all altitudinal classes, with the highest proportion between 2000 to 2200 masl.

Very erect and intermediate phenotypic classes of angle of branching occurred in all altitude classes, except erect phenotype in >3000 masl. There was no clear association between phenotypic classes and altitude.

Light brown seed coloured occur in all altitude classes, whereas dark and brown phenotypes occurred in all altitudinal classes except >3000 masl. Yellow and cream flower colour phenotypes occurred in all altitudinal classes, however, there was no clear association between flower colour phenotypes with altitude classes.

Stem colour is widely distributed countrywide. Slightly and intensive purple phenotypes were predominant and occurred in all altitudinal classes.

Table 3.2. Frequencies calculated as percentages of phenotypic classes of six morphological characters used for each geographic region. The frequencies of the four geographic regions and Ethiopia as a whole are calculated as weighted mean frequencies.

Geographic region	N-Max	LC*						LS			SC						FC			AB			SEC				
		1	2	3	4	5	6	1	2	3	1	2	3	4	5	6	1	2	3	1	2	3	1	2	3	4	5
NW&E Ethiopia																											
Gonder	19	0	0	42	58	0	0	11	78	11	42	11	0	21	26	0	68	32	0	11	89	0	5	26	69	0	0
Wello	18	0	0	67	33	0	0	89	11	0	6	0	33	50	6	5	89	11	0	17	83	0	0	83	17	0	0
Gojam	33	0	0	24	73	3	0	21	64	15	15	9	0	39	37	0	82	18	0	15	85	0	9	52	33	6	0
Region	70	3	0	37	59	1	0	36	54	10	20	7	9	37	26	1	80	20	0	14	86	0	6	52	39	3	0
W Ethiopia																											
Wellega	52	15	0	8	77	0	0	2	50	48	12	9	9	58	12	0	75	25	0	54	46	0	15	35	44	6	0
Illubabour	9	0	0	22	78	0	0	67	33	0	11	11	11	67	0	0	56	44	0	33	67	0	0	44	56	0	0
Keffa	8	0	0	38	62	0	0	0	50	50	13	0	0	63	24	0	25	75	0	13	87	0	0	75	25	0	0
Region	69	12	0	13	75	0	0	2	52	46	12	9	9	58	12	0	65	35	0	46	54	0	12	32	49	7	0
SW Ethiopia																											
Gamo Gofa	12	0	0	17	75	8	0	0	0	100	0	0	0	67	25	8	58	42	0	8	92	0	8	50	42	0	0
Sidamo	11	18	0	18	64	0	0	9	18	73	9	0	18	27	36	10	64	36	0	27	73	0	27	18	55	0	0
Bale	9	0	0	22	44	34	0	33	33	34	11	0	0	45	22	22	89	11	0	22	78	0	33	22	22	11	12
Region	32	6	0	19	62	13	0	13	16	71	6	0	6	47	28	13	69	31	0	19	81	0	22	31	41	3	3
C&E Ethiopia																											
Shewa	53	9	8	21	45	17	0	34	45	21	11	9	15	19	36	10	70	28	2	30	70	0	11	32	55	2	0
Arssi	19	0	11	16	53	20	0	21	68	11	5	0	11	53	31	0	84	16	0	32	68	0	26	5	37	26	6
Harargie	15	0	0	53	33	7	7	13	73	14	7	0	13	33	40	7	80	20	0	7	93	0	13	47	33	7	0
Region	87	6	7	25	45	16	1	28	55	17	9	6	14	29	36	6	75	24	1	26	74	0	15	44	32	8	1
Ethiopia	258	7	2	24	59	7	1	21	49	30	12	6	10	42	26	4	72	27	1	28	72	0	12	41	40	6	1

*See Table 3.1. for abbreviations; N-Max= maximum number accessions used for each region.

Table 3.3. Frequencies calculated as percentages of phenotypic classes of six morphological characters used for each altitude classes.

Altitude	N-Max	LC						LS			SC						FC			AB			SEC				
Classes		1	2	3	4	5	6	1	2	3	1	2	3	4	5	6	1	2	3	1	2	3	1	2	3	4	5
<1800 m	59	5	2	39	42	12	0	27	39	34	5	7	13	53	20	2	64	36	0	12	88	0	7	54	39	0	0
1800-2000 m	42	5	0	21	62	12	0	14	55	31	7	7	12	28	36	10	69	31	0	17	83	0	14	31	50	5	0
2000-2200 m	43	5	0	19	69	7	0	14	51	35	16	2	19	33	28	2	72	28	0	35	65	0	12	58	25	5	0
2200-2400 m	34	15	3	12	64	6	0	9	62	29	6	6	6	64	15	3	82	18	0	47	53	0	21	41	26	9	3
2400-2600 m	44	0	5	45	43	5	2	27	52	21	20	7	7	34	23	9	77	23	0	36	64	0	9	36	46	7	2
2600-2800 m	26	15	4	19	58	4	0	27	42	31	15	8	8	38	27	4	77	19	4	30	70	0	15	42	31	4	8
2800-3000 m	6	0	17	0	66	17	0	33	67	0	32	17	0	17	17	17	83	17	0	50	50	0	33	17	50	0	0
>3000 m	4	0	25	25	25	25	0	50	25	25	0	0	0	25	75	0	75	25	0	0	100	0	75	0	0	25	0
Ethiopia	258	6	3	27	55	9	0	21	50	29	12	6	11	41	25	5	73	27	0	28	72	0	14	43	37	4	2

See Table 3.1 for abbreviations; N-Max= maximum number accessions used for each altitudinal classes

Table 3.4. Chi-square values of each eco-geographical region for six morphological traits.

Region	LC* χ^2	LS χ^2	SC χ^2	FC χ^2	AB χ^2	SEC χ^2
NW & E Ethiopia	14.06*	31.31**	35.39**	2.55	0.32	15.80*
W Ethiopia	49.78**	9.90*	4.02	8.82**	5.47*	11.44
SW Ethiopia	12.89*	11.59*	8.31	2.43	1.45	9.58
C&E Ethiopia	16.12	5.63	79.96**	2.11	3.66	24.76**

*See Table 3.1. for abbreviations; *P=0.05, **P=0.01

Diversity Index

The Shannon Weaver Diversity Index (H') was calculated to compare phenotypic diversity among characters and regions. According to Brown and Weir (1983) this index is used in genetic resource studies as a convenient measure of both allelic richness and allelic evenness when using genetic data, but because of the log transformation it is not readily interpretable in genetic terms. A low H' indicates extremely unbalanced frequency classes for an individual trait and lack of genetic diversity. Estimates were made for each character and pooled across characters and regions for qualitative traits (Table 3.5).

Table 3.5 gives the estimates of H' individually and pooled over characters and provinces and appropriately weighted by numbers of accessions. The over all province values of Shannon-Weaver Diversity Index (H') ranged from 0.30-0.79, 0.00-1.00, 0.46-0.93, 0.32-0.63, 0.23-0.63, 0.28-95 for leaf colour, leaf size, stem colour, flower colour, angle of branching and seed colour respectively (Table 3.5).

The three highest mean values of H' were from Shewa (H'=0.72), Bale (H'=0.68), and Arssi (H'=0.64) provinces. The lowest values of H' were from Gamu Gofa (H'=0.38) and Wello (H'=0.39).

The estimates of diversity index for individual characters pooled over provinces and regions showed considerable variation among characters (Table 3.5). For example, the Shannon Weaver diversity index varied from 0.54 for angle of branching to 0.95 for leaf size countrywide. The average diversity across eight characters from all the

regions or the over all diversity for the country was 0.72 ± 0.06. Based on the average diversity calculated for each region, across all characters the most diverse region was central and eastern Ethiopia with the H' value of 0.74 ± 0.06 and the least diverse was north-western Ethiopia (0.60 ± 0.08). In general, northern Ethiopia which is considered to be a centre of origin or diversity for many crops had estimates of $H'=0.60 \pm 0.08$ which was lower compared to other regions. This might be due to natural selection in the areas of northern Ethiopia, which are frequently affected by drought and/or it could also be due to sampling effects.

Although mean H' values appear not variable among the collection regions, a closer examination into variances for each trait indicates that the diversities are non-overlapping in some instances. For example, Gojam ($H'=0.58$) and Harargie ($H'=0.58$) exhibited high mean diversity levels. However, the two regions showed substantial differences for diversity in leaf size and leaf colour (Table 3.5). Accessions of Gojam were more variable for leaf size than accessions of Harargie. On the other hand, collections of Harargie were more variable for leaf colour than were collections from Gojam. This suggests that some regions have the potential to offer higher or highest diversity for a particular trait even though their mean variation appears similar to the other regions.

Compared with the regional estimates, H' values pooled over all traits for altitudinal classes showed less variation (Table 3.6). The lower altitude classes showed similar diversity estimates ($H'=0.65$) and declined above 3000 masl. The altitude class with the highest diversity index was between 2400 and 2800 masl, whereas the lowest was for areas above 3000 masl, which is probably associated with the sample size, in addition to the natural and artificial selection pressure operating at high altitudes.

Table 3.5. Estimates of (H') for regions and six morphological characters and mean diversity ($\bar{H}'$) over all characters.

Geographic origin	N-Max	LC*	LS	SC	FC	AB	SEC	$\bar{H}' \pm$ SE
NW Ethiopia								
Gonder	19	0.38	0.62	0.72	0.57	0.32	0.47	0.51±0.06
Wello	18	0.35	0.32	0.67	0.32	0.42	0.28	0.39±0.05
Gojam	33	0.45	0.82	0.69	0.43	0.38	0.68	0.58±0.07
Region	70	0.46	0.85	0.82	0.46	0.37	0.61	0.60±0.08
W Ethiopia								
Wellega	52	0.39	0.71	0.70	0.51	0.63	0.73	0.61±0.06
Illubabur	9	0.30	0.57	0.56	0.63	0.57	0.43	0.51±0.05
Keffa	8	0.37	0.63	0.50	0.51	0.35	0.35	0.45±0.04
Region	69	0.41	0.71	0.70	0.59	0.63	0.71	0.63±0.04
SW Ethiopia								
Gamo Gofa	12	0.40	0.00	0.46	0.62	0.25	0.57	0.38±0.09
Sidamo	11	0.50	0.69	0.83	0.59	0.53	0.61	0.63±0.05
Bale	9	0.59	1.00	0.71	0.32	0.48	0.95	0.68±0.01
Region	32	0.59	0.73	0.73	0.56	0.44	0.79	0.64±0.05
C&E Ethiopia								
Shewa	53	0.79	0.95	0.93	0.48	0.55	0.62	0.72±0.08
Arssi	19	0.66	0.75	0.61	0.40	0.57	0.86	0.64±0.06
Harargie	15	0.60	0.70	0.77	0.45	0.23	0.73	0.58±0.08
Region	87	0.78	0.90	0.87	0.56	0.52	0.78	0.74±0.06
Ethiopia	258	0.64	0.95	0.84	0.58	0.54	0.75	0.72±0.06

*See Table 3.1. for abbreviations; N-Max maximum number of accessions used in each region.

Table 3.6. Estimates of (H') for altitudinal classes and six morphological characters and mean diversity ($\bar{H}'$) over all characters.

Altitude classes	N-Max	LC*	LS	SC	FC	AB	SEC	$\bar{H}' \pm$SE
<1800 m	59	0.68	0.99	0.75	0.59	0.33	0.55	0.65±0.08
1800-2000 m	42	0.57	0.88	0.88	0.56	0.42	0.71	0.67±0.07
2000-2200 m	43	0.51	0.90	0.83	0.54	0.59	0.66	0.67±0.06
2200-2400 m	34	0.61	0.79	0.66	0.43	0.63	0.85	0.66±0.06
2400-2600 m	44	0.61	0.93	0.90	0.49	0.59	0.75	0.71±0.07
2600-2800 m	26	0.66	0.98	0.86	0.59	0.56	0.83	0.75±0.07
2800-3000 m	6	0.49	0.58	0.88	0.42	0.63	0.63	0.61±0.06
>3000 m	4	0.77	0.95	0.31	0.51	0.00	0.35	0.48±0.01
Ethiopia	258	0.65	0.94	0.85	0.53	0.54	0.75	0.71±0.07

*See Table 3.1. for abbreviations; N-Max maximum number of accessions used in each altitude

Based on the characters evaluated across the regions, the highest diversity index was found in Shewa, Bale and Arssi provinces. The highest polymorphism is concentrated in areas between altitudinal brackets of 2600 and 2800 masl. This altitudinal range includes the major Gomenzer growing areas in the country. These phenomena are

indicative of high genetic diversity and abundance of Gomenzer in a particular altitude range.

The results of the character distribution pattern and phenotypic diversity analyses permit some broad generalization about collection and conservation of Gomenzer landraces in Ethiopia. The present study indicates that the amount of variability is not uniformly distributed in all geographic regions and altitude classes. From a genetic conservation point of view, sites in Shewa and Arssi (central and eastern regions), and Bale (south west regions) coupled with the appropriate altitude parameter appear to be suitable for exploration and in *situ* conservation.

Future collection missions of Gomenzer germplasm as a source of diversity should take account of the distribution of polymorphisms. Priorities of germplasm collection should focus on areas with relatively large variation (see Tables 3.5 and 3.6) with due consideration to the cause of genetic erosion, drought and environmental degradation. Both *in situ* and *ex situ* conservation strategies shall be complemented each other for sustainable conservation and utilization of Gomenzer germplasm.

3.5. REFERENCES

ABEBE, D., T. DAWIT, M. GETAHUN, and M. DEBERITU, 1992. Ethiopia's genetic resources of oilseeds. *In*: Proc. 1st National Oilseeds Workshop, Institute of Agricultural Research, Addis Ababa, Ethiopia, pp 12-17.

ALLARD, R.W., 1970. Population structure and sampling methods. *In*: O.H. Frankel and E. Bennett (Eds.), Genetic Resources in plants Their Exploitation and Conservation, IBP Handbook No. 11. F.A. Davis Company, Phildelphia, PA. pp. 97-107.

ASFAW, Z., 1989. Relationships between spike morphology, hordeins and altitude within Ethiopian barley, *Hordeum vulgarae* L. *Hereditas* 110: 203-209.

AYANA, A., and E. BEKELE, 1999. Geographical patterns of morphological variation in Sorghum (*Sorghum bicolour* (L) Moench.) germplasm from Ethiopian and Eritrea. Qualitative characters. *Hereditas* 129: 195-205.

BECHERE, E., G. BELAY, D. MITIKU, and A. MERKER, 1996. Phenotypic diversity of tetraploid wheat landraces from northern and north central regions of Ethiopia. *Hereditas* 124: 165-172.

BEKELE, E., 1983. Some measures of gene diversity analysis on landrace populations of Ethiopian barley. *Hereditas* 98: 127-143.

BEKELE, E., 1996. Morphological analysis of *Eragrosties*: detection for regional patterns of variation. *SINET*: *Eth. J. Sci.* 19: 117-140.

BEKELE, F., and I. BEKELE, 1996. A sampling of the phenetic diversity of cacao in the International Cacao Gene Bank of Trinidad. *Crop Sci.* 36: 57-64.

BOWMAN, K.O., K. HUTCHESON, E.P. ODUM, and L.R. SHENTON, 1971. Comments on the distribution of diversity. *In*: G.P. Patil, E.C. Pielou, W.E. Waters (eds.), Statistical Ecology, Vol. 3. Many Species, Populations, Ecosystems, and System analysis. Pennsylvania University, University Park, pp. 315-366.

BROWN, A.H.D., 1978. Isozymes, plant population genetic structure and genetic conservation. *Theor. Appl. Genet.* 52: 145-157.

BROWN, A.H.D., 1989. Core collections: a practical approach to genetic resource management. *Genome* 31: 818-824.

BROWN, A.H.D. and B.S. WEIR, 1983. Measuring genetic variability in plant populations. *In*: S.D. Tanksley and T.J. Orton (ed.). Isozyme in plant genetics and breeding. Part A. Elsevier Science Publ. Amsterdam. pp. 219-238.

DEMISSIE, A. and A. BJÖRNSTAD, 1996. Phenotypic diversity of Ethiopian barleys in relation to geographical regions, altitudinal range, and agro-ecological zones: as an aid to germplasm collection and conservation strategy. *Hereditas* 124: 17-29.

De WET, J.M.J., J.R. HARLAN and E.G. PRICE, 1976. Variability in Sorghum bicolour. *In*: J.R. Harlan, J.M.J. de Wet and A.B.L. Stemler (Eds.). Origins of African Plant Domestication. pp. 453-463.

ENGELES, J.M.J., 1994. Genetic diversity in Ethiopian barley in relation to altitude. *Genetic Resources and Crop Evolution* 41: 67-73.

FRANKEL, O.H., and M.E. SOULE, 1981. Conservation and evolution. Cambridge Univ. Press, New York.

GENET, T., and K. BELETE, 2000. Phenotypic diversity in the Ethiopian noug germplasm. *African Crop Science Journal* 8(2): 137-143.

HARLAN, J.R., 1969. Ethiopia: a centre of diversity. *Econ. Bot.* 23: 309-314.

HAYWARD, M.D., and E.L. BREESE, 1993. Population structure and variability. *In*: M.D. Hayward, N.O. Bosemark and I. Romagosa (Eds.). Plant Breeding: Principles and prospects, Chapman & Hall, pp. 16-29.

HAWKES, J.G., 1976. Sampling gene pools. Proceedings of Nato conference on conservation of threatened plants. Series 1. Ecology, Plenum, London.

JAIN, S.K., C.O. QUALSET, G.M. BHATT, and K.K. WU, 1975. Geographical patterns of phenotypic diversity in world collections of durum wheat. *Crop Sci.* 15: 700-704.

KEBEBEW, A., A. MERKER and H. TEFERA, 2002. Quantitative trait variation in Tef [*Eragrostis tef* (Zucc.) Trotter] germplasm from western Ethiopia. *Euphytica* 127: 399-410.

KEFYALEW, T., H. TEFERA, K. ASSEFA, and M. AYELE, 2000. Phenotypic diversity for qualitative and phenologic characters in germplasm collections of tef (*Eragrostis tef*). *Genetic Resources and Crop Evolution* 47: 73-80.

KNOWLES, P.F., T.E. KEARNEY, and D.B. COHEN, 1981. Species of rapeseed and mustard as oil crops in California. *In*: Pryde E.H. (ed.). New sources of fats and oils. Am. Oil Chem. Soc. Champain, IL, pp 255-268.

LEUR, van J.A.G., S. CECCARELLI, and S. GRANDO, 1989. Diversity for disease resistance in barley landraces from Syria and Jordan. *Plant Breeding* 103: 324-335.

MARSHALL, D.R., and A.H.D. BROWN, 1975. Optimum sampling strategies in genetic conservation. *In*: O.H. Frankel and J.G. Hawkes (Eds.). Crop Genetic Resources for Today and Tomorrow, Cambridge Univ. Press, Cambridge. pp. 53-80,

MENGESHA, M., 1975. Crop germplasm diversity and resources in Ethiopia. *In*: Frankle, O.H., and J.G., Hawkes (eds.), Crop Genetic Resources for Today and Tomorrow. Cambridge: Univ. press, Cambridge, pp. 449-453.

MORENO-GONZALEZ, J., and J.J. CUBERO, 1993. Selection strategies and choice of breeding methods. *In*: M.D. Hayward, N.O. Bosemark and I. Romagosa (Eds.). Plant Breeding principles and prospects, Chapman & Hall. pp. 16-29,

NEGASSA, M., 1985. Pattern of phenotypic diversity in an Ethiopian barley collection, and the Arusi-Bale Highland as a centre of origin of barley. *Hereditas* 102: 139-150.

NEGASSA, M., 1986a. Estimates of phenotypic diversity and breeding potential of Ethiopian wheat's. *Hereditas* 104: 41-48.

NEGASSA, M., 1986b. Patterns of diversity in Ethiopian wheat's (*Triticum* spp.) and a gene centre for quality breeding. *Plant Breeding* 97: 147-162.

PERRY, M.C., and M.S. MCINTOSH, 1991. Geographical patterns of variation in the USDA soybean germplasm collection: I. Morphological traits. *Crop Sci.* 31: 1350-1355.

PIELOU, E.C., 1966. The measurement of diversity in different types of biological collections. *J. Theor. Biol.* 13: 131-144.

POLIGNANO, G.B., E. ALBA, P. UGGENTI and G. SCIPPA, 1999. Geographical patterns of variation in Bari faba bean germplasm collection. *Genetic Resources and Crop Evolution* 46: 183-192.

SHANNON, C.E., and W. WEAVER, 1949. The mathematical theory of communication. Univ. Illinois Press. Urbana. Chicago, London.

SIMMONDS, N.W., 1979. Evolution of crop plants. Longman Group Ltd; London.

TESHOME, A., and B.R. BAUM, L. HAHRIG, J.K. TORRANCE, T.J. ARNASON, J.D. LAMBERT, 1997. Sorghum (*Sorghum bicolour* (L.) Moench) landrace variation and classification in North Shewa and South Wello, Ethiopia. *Euphytica* 97: 255-263.

TOLBERT, D.M., C.O. QUALSET, S.K. JAIN, and J.C. CRADDOCK, 1979. A diversity analysis of a world collection of barley. *Crop Sci.* 19: 789-794.

U N, 1935. Genome analysis in *Brassica. Jap. J. Bot.* 7: 389-452.

UPADHYAYA, H.D., P.J. BRAMEL, R. ORTIZE and S. SINGH, 2002. Geographical patterns of diversity for morphological and agronomic traits in the groundnut germplasm collection. *Euphytica* 128: 191-204.

VAVILOV, N.I. 1951. Phytogeographic basis of plant breeding. The origin, variation, immunity and breeding of cultivated plants. *Chron. Bot.* 13: 13-54.

WOREDE, M., 1988. Diversity and the genetic resource base. *Ethiopian J. Agric. Sci.* 10(1-2): 39-52.

YANG, R.C., S. JANA, and J.M. CLARK, 1991. Phenotypic diversity and associations of some potentially drought-responsive characters in durum wheat. *Crop Sci.* 31: 1484-1491.

YOON, M.S., J.W. AHN, S.J. PARK, H.J. BAEK, N.K. PARK and Y.D. RHO, 2000. Geographical patterns of morphological variation in soybean germplasm. *Korean J. Crop Sci.* 45(4): 267-271.

ZOHARY, D., 1970. Centres of diversity and centres of origin. *In*: O.H. Frankel and E. Bennett (eds.). Genetic resources in plants, their exploitation and conservation. Black Well Scientific Publication, Oxford. pp. 33-42.

CHAPTER IV

CAPILLARY GAS CHROMATOGRAPHY ANALYSIS OF ETHIOPIAN MUSTARD TO STUDY THE VARIABILITY OF FATTY ACID COMPOSITION

4.1. ABSTRACT

Ethiopian mustard (Brassica carinata) is one of the oil crops grown in Ethiopia. It possesses a number of agronomic advantages over other oilseed crops with similar ecological adaptation. However, the oil is considered low quality, as it contains long chain monounsaturated fatty acids, mainly erucic acid. High erucic acid content is beneficial for the polymer industry, whereas low erucic acid is recommended for food purposes. Oil high in oleic acid is in demand for commercial food-service applications due to long shelf life and cholesterol reducing properties. Both linoleic and linolenic acids are essential fatty acids; however, less than 3% linolenic acid is preferred for oil stability. Therefore, it is important to undertake systematic characterization of the available gene pool for its variable fatty acid profile to be utilized for specific purposes.

Although efforts have been made to improve its quality, much has to be done to use natural variations that might exist within the species for fatty acid composition. The objectives of this study were to determine the oil content and fatty acid composition of Ethiopian mustard to determine the range of genetic diversity for these traits, and to select genotypes that can be used in crosses to improve these traits. In this study 98 B. carinata genotypes were analyzed by Capillary Gas Chromatography (CGC) to study the variability of fatty acid composition. The CGC analysis revealed a wide variation in fatty acid composition. Oil content varied from 25 to 48 %. Twenty-six fatty acids were identified. In all accessions, the predominant fatty acids were erucic [C22: 1c13 (n-9)] *(6.91-35.05%), linoleic* [C18: 2c9, 12(n-6)] *(17.21-28.53%), α-linolenic* [ALA, C18:3c9,12,15(n-3)] *(10.50-22.52%), and oleic [*C18:1c9(n-9)] *(8.91-24.64%), followed by gadoleic [C20:1c11(n-11)] (1.22-10.24%) and palmitic (C16:0) (3.38-16.42%). To a lesser extent stearic (C18:0) (1.22-7.43%), vaccinic [C18:1c7(n-7)]*

(0.86-3.10%), nervonic [C24:1c15(n-9)] (0.42-1.38%) and behenic acids (C22:0) (0.17-0.51%) were also found in all accessions. Gamma linolenic acid [GLA, C18:3c6,9,12(n-6)], lignoceric (C24:0), dimo-γ-linolenic [C20:3c8,11,14(n-6)], margaric (C17:0), palmitoleic [C16:1c9(n-7)], myristoleic (C14:1c9), myristic (C14:0) and butyric acids (C4:0) were also rarely detected. As far as we aware some of these fatty acids have not previously been reported in the seed oil of Ethiopian mustard.

Palmitic acid had significant positive correlations with stearic acid. Erucic acid showed significant and negative linear correlations with palmitic, stearic, vaccinic, oleic, linoleic, α-linolenic acids and positive correlation with eicosenoic acid. These significant correlations could be associated with the biosynthetic pathways of the fatty acids, which are not fully elucidated. They suggest, however, that selection for a particular fatty acid could lead to and increase of those positively correlated. Selection and hybridization techniques could then be applied to modify the oil content and fatty acid composition of Ethiopian mustard, considering the variability observed.

Key words: *Brassica carinata,* capillary gas chromatography, Ethiopian mustard, fatty acid composition, oil content

4.2. INTRODUCTION

Ethiopian mustard or Abyssinian mustard or locally known as Gomenzer (*Brassica carinata A.* Braun) is one of the oilseed *Brassica* species. In its native Ethiopia it is mainly used as leaf vegetable and as an oilseed crop. It is high yielding, resistant to diseases (Gugel *et al.* 1990; Yitbarek, 1992), insect pests (Bayeh and Gebre Medhin, 1992), tolerant to heat and drought (Malik, 1990) and shattering resistance than any of the oilseed crops adapted to comparable ecological amplitude (Alemayehu and Becker, 2001). Despite all these agronomic advantages the germplasm is characterized by high erucic acid content in the oil, which is reported to be in the range of 35-44% (Röbbelen and Thies, 1980, Becker *et al.*, 1999; Westphal and Marquard, 1980). This is above the acceptable range (<2%) from nutritional standpoint, in quality rapeseed (Sauer and Kramer, 1983). The research effort in the

past years have led to the creation of some low erucic acid genotypes (Fernández-Escobar *et al.*, 1988; Alonso *et al.*, 1991; Getinet *et al.*, 1994; Velasco *et al.*, 1995). Oils high in erucic acid content are beneficial for the polymer industry; where as low erucic acid contents are preferred for food purposes. High oleic acid is in demand for commercial food-service applications due to a long shelf life and cholesterol reducing properties. Both linoleic and linolenic acids are essential fatty acids; however, less than 3% linolenic acid is preferred for oil stability. Hence, both the development of commercial varieties free of erucic acid and with very high erucic acid content are breeding objectives in *Brassica* oil crops. Other important objectives are the increase of oleic and linoleic acids, and the reduction of linolenic acid content (Röbbelen, 1991). Therefore it is important to undertake systematic characterization of the available gene pool for its variable fatty acid profile to be utilized for specific purposes. The surveys of fatty acid compositions in the international germplasm of soybean (Rebetzke *et al.*, 1996), safflower (Fernández-Martinez *et al.*, 1993), maize (Cheesbrough *et al.*, 1997) and rapeseed-mustard (Kaushik and Agnihotri, 2000) have revealed wide variation in the proportion of the different fatty acids, offering possibilities of developing superior quality edible oils and specialized industrial oils (Ohlrogge, 1994; Byrum *et al.*, 1997; Rahman *et al.*, 1997).

A survey of the fatty acid composition of Ethiopian mustard would be useful to identify land races of different seed oil compositions for breeding improved varieties with high yield and superior quality edible oil. In this study seed oil content and fatty acid composition was determined for 98 accessions/varieties of *B. carinata* germplasm.

4.3. MATERIALS AND METHODS

Plant material

A total of 98 accessions of *B. carinata*, obtained from the Ethiopian National Breeding Program and from the germplasm collections of the Institute of Biodiversity Conservation and Research, Addis Ababa, Ethiopia and two accessions from Canada were used in this study. The germplasm collections represented the major *B. carinata* growing regions of the country. A detailed description of the materials used in this study is shown in Table 4.1. The plants were grown in the fields of Adet Research

Center in 1999/2000 main cropping season. Each sample consisted of seeds from different plants of each accession, which were analyzed for seed oil content and fatty acid composition.

Table 4.1. List of Ethiopian mustard (*Brassica carinata*) accessions/varieties used in this study.

Code	Accession	Code	Accession	Code	Accession
1	Yellow Dodolla-03	34	Adet	67	PGRC/E 21361-03
2	S-67-02	35	PGRC/E 21358/2	68	PGRC/E 21261-04
3	PGRC/E 21261-03	36	PGRC/E 207928	69	PGRC/E 21001
4	PGRC/E 203221-03	37	PGRC/E 21320/5	70	PGRC/E 21261-01
5	C94-S-67	38	PGRC/E 20112/2	71	PGRC/E 21324
6	C94-Dodolla	39	PGRC/E 20165	72	PGRC/E 21163/1
7	Yellow Dodolla-01	40	PGRC/E 21051	73	PGRC/E 20163
8	PGRC/E 203221-01	41	PGRC/E 21261-05	74	PGRC/E 21324/1
9	PGRC/E 21261-02	42	PGRC/E 210406	75	PGRC/E 21236/1
10	PGRC/E 20130	43	PGRC/E 21169	76	PGRC/E 20104
11	Merawi	44	PGRC/E 21184	77	PGRC/E 21170
12	PGRC/E 20059	45	PGRC/E 20113	78	PGRC/E 208410
13	PGRC/E 21207	46	PGRC/E 20021	79	PGRC/E 207931
14	PGRC/E 20080	47	PGRC/E 21252	80	PGRC/E 208004
15	PGRC/E 207929	48	PGRC/E 20126	81	PGRC/E 21304
16	PGRC/E 207975	49	PGRC/E 21163	82	PGRC/E 21010
17	PGRC/E 20168/1	50	PGRC/E 20013	83	PGRC/E 21031
18	PGRC/E 20095/1	51	PGRC/E 21057/1	84	PGRC/E 20153
19	PGRC/E 21356/1	52	PGRC/E 20076/3	85	PGRC/E 20126/1
20	PGRC/E 21237	53	PGRC/E 20168/2	86	PGRC/E 20120
21	PGRC/E 21156	54	PGRC/E 200394	87	PGRC/E 208551
22	PGRC/E 20112/2	55	PGRC/E 207481	88	PGRC/E 20103
23	PGRC/E 200413	56	PGRC/E 21172	89	PGRC/E 21373
24	(4DxZem-1)x(Zem-1-AD/88)	57	PGRC/E 20165/1	90	PGRC/E 20164
25	(4DxZem-1)x(Zem-1-F5/10)	58	PGRC/E 208401	91	PGRC/E 20165/2
26	PGRC/E 20156	59	PGRC/E 21223/2	92	PGRC/E 20090/1
27	PGRC/E 21162/1	60	PGRC/E 21224/3	93	PGRC/E 20162/1
28	PGRC/E 20163/1	61	PGRC/E 20168/3	94	PGRC/E 21169/1
29	PGRC/E 20076/2	62	PGRC/E 21328/1	95	PGRC/E 20175/1
30	PGRC/E 20147/1	63	PGRC/E 21162	96	PGRC/E 21058/2
31	PGRC/E 208404	64	PGRC/E 21235/1	97	PGRC/E 21261
32	Yellow Dodolla	65	PGRC/E 21263	98	PGRC/E 20163/5
33	S-67-01	66	PGRC/E 21057		

Analytical methods

Lipid extraction

Total lipid was extracted with chloroform-methanol (2:1 v/v) as described by Folch *et al.* (1957). Butylated hydroxy toluene (20 g) was placed in a 2000 ml volumetric flask. Chloroform (1333.33 ml) and methanol (666.67 ml) were added. In the lower phase, chloroform (1700.97 ml) was placed in 2000 ml volumetric flak, methanol (277.23 ml) and water (19.8 ml) were added with 86:14:1 v/v proportions respectively. About 200 g of seed from each sample were ground with mortar and pestle. A ± 0.5 g ground

seed was used for lipid extraction in a 250 ml round bottom flask. A 30 ml of Chloroform: methanol (2:1 v/v) was added and left overnight in a refrigerator at 4°C.

The sample was transferred into a round bottom flask through a pre-weighed 18.5 cm Whatman No. 1 filter paper on top of the separating funnel. The round bottom flask was washed three times with 12.33 ml of chloroform: methanol (2:1 v/v) and transferred each time to filter paper. After allowing everything to run through, the filter paper was removed. Then 16.5 ml distilled H_2O was added to each separating funnel, which was recapped, shaken thoroughly, and allowed one hour for separation. After this the lower phase was drained into a 500 ml round bottom flask and 50 ml of lower phase was added to each separating funnel and allowed 15 minutes for separation. After 15 minutes the lower phase was drained into the same 500 ml round bottom flask and the upper phase discarded. Contents from the 500 ml of round bottom flask were evaporated under vacuum in a rotary evaporator at 60°C for 20 minutes at 15 psi. The remaining water was removed by the addition of 30 ml methanol and then by evaporating each round bottom flask at 60 °C at 15 psi. Contents of each 500 ml round bottom flask were washed six times with 5 ml portions of diethyl ether into a pre-weighed polytop. Diethyl ether was removed from each polytop by evaporation under a stream of nitrogen on a heating block at 60°C for 20 minutes. Filter papers and capped polytops were put in a vacuum oven, dried at 50°C overnight and weighed the next morning. The oil-solvent mixture was dried over anhydrous sodium sulphate, the solvent evaporated under nitrogen, and the oil percentage determined by weighing. A ±10 mg of the lipid was transferred to another polytop for methylation.

Methylation

A modification of the procedure described by Slover and Lanza (1979) was used. An amount of ±10.0 mg of lipid was put or washed with 6x 1 ml hexane into the test tube with a Teflon-lined cap. Hexane was removed by N_2 evaporation without application of heat. Methanolic 0.5 N NaOH (1 ml) was added, and the tube capped and heated in a boiling water bath for 15 min. After the tube had cooled, 2 ml of BF_3/CH_3OH (14%) was added; then the tube was recapped and heated in the boiling water bath for an additional 15 min. The tube was cooled, and 1 ml of hexane and 2 ml of saturated aqueous NaCl were added. The tube was shaken vigorously for 1 min, and allowed to

stand for 10 min until the phases separated. The upper 70% of hexane layer was transferred with a Pasteur pipette to a 45x11 mm vial containing a 1 mm layer of anhydrous Na_2SO_4. The vial was capped, shaken, and allowed to stand for at least 20 min. to remove traces of water, then 100 µl hexane from each vial was transferred to a clean labeled auto-sampler vial and 900 µl hexane was added to each vial and stored below freezing point for GC analyses.

Determination of fatty acid composition by gas chromatography

After methylation individual fatty acid composition was determined by gas chromatography for all accessions. Fatty acids were quantified using a Varian GX 3400 flame ionization gas chromatograph, with a fused silica capillary column, Chrompack CPSIL 88 (100 m length, 0.25 µm ID, 0.2 µm film thickness). Column temperature was 40-230 °C (hold 2 minutes 4°C/min; hold 10 minutes). Fatty acid methyl esters in hexane (1µl) were injected into the column using a Varian 8200 CX Auto-sampler with a split ratio of 100:1. The injection port and detector were both maintained at 250 °C. Hydrogen was used as a carrier gas at 45 psi and nitrogen was used as the makeup gas. Chromatograms were recorded with Varian Star Chromatography Software version 4.0. Identification of sample fatty acids was made by comparing the relative retention times of fatty acid methyl ester peaks from samples with those of standards obtained from SIGMA.

Fatty acid ratios

It is difficult to evaluate the potential of different phenotypes for plant breeding by comparing individual fatty acid values, because they are interconnected by virtue of bio-synthetic pathways and selection will affect the whole system (Velasco *et al.*, 1997). For this reason several ratios were additionally used. They are called elongation ratio (ER), desaturation ratio (DR), oleic desaturation ratio (ODR), and linoleic desaturation ratio (LDR). The ER estimates the relative weight of the elongation pathway from oleic acid (C18: 1) to eicosenoic (C20: 1) and erucic acid (C22: 1), while the DR estimates the relative weight of the desaturation pathway from oleic acid (18:1) to linoleic (C18: 2) and linolenic acid (C18: 3) within the over all fatty acid biosynthesis system (Velasco *et al.*, 1998; Pleines and Friedt, 1988). The ODR and LDR are calculated following formulas of Pleines and Friedt (1988) and they estimate, within the desaturation pathway, the efficiency of the desaturation from

oleic to linoleic (ODR) and from linoleic to linolenic acid (LDR). These ratios were calculated as follows:

$$ER = \frac{\%C20:1 + \%C22:1}{\%C20:1 + \%C22:1 + \%C18:1 + \%C18:2 + \%C18:3}$$

$$DR = \frac{\%C18:2 + \%C18:3}{\%C20:1 + \%C22:1 + \%C18:1 + \%C18:2 + \%C18:3}$$

$$ODR = \frac{\%C18:2 + \%C18:3}{\%C18:1 + \%C18:2 + \%C18:3}$$

$$LDR = \frac{\%C18:3}{\%C18:2 + \%C18:3}$$

Therefore, the results reported in the present study represent the average composition of the germplasm accessions; the within accessions variability was not investigated.

Statistical analysis

The statistical package NCSS 2000 (Hintze, 2000) was used to do simple descriptive statistics and Pearson correlation coefficients.

4.4. RESULTS AND DISCUSSION

Description of variability for fatty acid composition

Fatty acid profile and descriptive statistics for oil content and fatty acid composition in *B. carianta* accessions are shown in Tables 4.2 and 4.3. Trival and systematic names, the number of carbon atoms, and the position of carbon double bonds in the fatty acid composition are also listed in Table 4.3. Considerable variation was observed among the quality parameters recorded for the 98 accessions.

The fatty acid composition, fatty acid methyl esters as % (w/w) of total fatty acids, of the seed oil of the accessions identified 26 fatty acids (Table 4.2). The predominant fatty acids found were erucic acid [C22: 1c13 (n=9)], linoleic acid [C18: 2c9, 12(n-6)], alpha-linolenic acid or α-linolenic acid [ALA, C18:3c9,12,15(n-3)], oleic acid (C18:1c9(n-9), eicosenoic or gadoleic acid [C20:1c11(n-11)] and palmitic acid (C16:0). To a lesser extent stearic acid (C18:0), vaccinic acid [C18:1c7(n-7)], nervonic acid [C24:1c15(n-9)] and eicosanoic (arachidic) acid (C20:0) were also found in all accessions.

Lignoceric acid (C24:0), dimo-γ-linolenic acid [C20:3c8,11,14(n-6)], gamma-linolenic acid or γ-linolenic acid [GLA, C18:3c6,9,12 (n-6)], margaric acid (C17:0), palmitoleic acid [C16:1c7(n-7)], myristoleic acid (C14:1c9), myristic acid (C14:0) and butyric acid (C4:0) were rarely detected in some accessions. As far as we aware undecanoic, tridecannoic, pentadecenoic, margaric, nonadecanoic, lignoceric, myristoleic, nervonic, fatty acids have not been previously reported.

Table 4.2. Fatty acid profiles of Ethiopian mustard accessions/varieties determined by capillary gas chromatography.

No	Accessions/varieties	%Fat	C4:0	C11:0	C13:0	C14:0	C14:1e9	C15:0	C16:0	C16:1e9	C17:0	C18:0	C18:1c7	C18:1e9	C18:2e9,12	C18:3c6,9,12	C18:3e9,12,15	C19:0	C20:0	C20:1e11	C20:2e11,14	C20:3e8,11,14	C20:5e5,8,11,14,17	C22:0	C22:1e13	C22:2e13,16	C24:0	C24:1e15
1	Yellow Doddolla-03	41.85	0.00	0.00	0.00	0.12	0.00	0.00	4.11	0.17	0.00	1.41	1.21	9.44	23.38	0.00	17.33	0.00	0.79	7.68	1.31	0.30	0.25	0.42	29.25	1.08	0.00	1.01
2	S-67-02	40.14	0.00	0.00	0.00	0.12	0.00	0.00	4.17	0.17	0.00	1.55	1.48	9.78	24.58	0.00	16.33	0.00	0.82	7.10	1.42	0.36	0.24	0.41	27.99	1.05	0.00	1.07
3	PGRC/E 21261-03	37.36	0.00	0.00	0.00	0.00	0.00	0.00	4.09	0.15	0.00	1.68	1.39	9.49	21.76	0.00	17.75	0.00	0.83	6.30	1.25	0.42	0.31	0.51	30.84	1.13	0.00	1.03
4	PGRC/E 203221-03	41.15	0.00	0.00	0.00	0.15	0.00	0.03	5.79	0.24	0.07	1.90	2.66	16.36	26.55	0.07	19.36	0.03	0.61	8.74	1.09	0.20	0.12	0.18	13.32	0.41	0.00	0.65
5	C94-S-67	24.92	0.00	0.03	0.02	0.14	0.00	0.03	4.64	0.22	0.08	1.96	2.85	19.74	28.51	0.10	22.20	0.02	0.62	5.91	0.82	0.14	0.12	0.22	8.73	0.30	0.00	0.57
6	C94-Dodolla	34.98	0.02	0.03	0.02	0.17	0.00	0.04	5.48	0.29	0.08	1.77	3.10	21.59	28.53	0.10	22.52	0.01	0.52	4.90	0.74	0.11	0.10	0.18	6.91	0.22	0.00	0.48
7	Yellow Dodolla-01	37.52	0.00	0.00	0.00	0.00	0.00	0.00	3.93	0.15	0.00	1.52	1.40	9.93	20.31	0.00	17.80	0.00	0.84	8.49	1.21	0.38	0.27	0.42	30.82	1.06	0.00	1.20
8	PGRC/E 203221-01	42.04	0.00	0.00	0.00	0.07	0.00	0.00	3.94	0.09	0.00	1.47	1.71	12.30	22.59	0.00	19.12	0.00	0.69	8.20	1.27	0.34	0.26	0.30	25.33	0.78	0.00	0.91
9	PGRC/E 21261-02	47.97	0.00	0.00	0.00	0.09	0.00	0.00	4.22	0.14	0.00	1.54	1.51	11.42	21.47	0.00	19.45	0.00	0.68	7.67	1.11	0.36	0.17	0.33	27.06	0.83	0.00	1.02
10	PGRC/E 20130	40.94	0.00	0.00	0.00	0.00	0.00	0.00	4.05	0.22	0.00	1.51	1.59	10.29	22.38	0.00	18.13	0.00	0.77	7.87	1.36	0.44	0.26	0.38	27.88	1.03	0.00	1.07
11	Merawi	41.92	0.00	0.00	0.00	0.13	0.00	0.00	4.29	0.18	0.00	1.61	1.66	12.11	21.43	0.08	20.10	0.00	0.74	7.77	1.17	0.36	0.23	0.33	24.70	0.77	0.00	0.93
12	PGRC/E20059	41.88	0.00	0.00	0.00	0.00	0.00	0.00	3.54	0.00	0.00	1.22	1.58	9.53	19.10	0.00	19.72	0.00	0.76	8.12	1.45	0.53	0.26	0.43	30.55	1.02	0.00	1.25
13	PGRC/E 21207	39.36	0.00	0.00	0.00	0.00	0.00	0.00	4.19	0.17	0.00	1.43	1.53	9.72	21.84	0.00	19.04	0.33	0.70	7.28	1.45	0.47	0.24	0.37	28.36	1.09	0.00	1.07
14	PGRC/E 20080	39.41	0.00	0.00	0.00	0.00	0.00	0.00	4.14	0.19	0.00	1.57	1.58	10.22	21.72	0.00	18.75	0.00	0.79	7.60	1.16	0.41	0.24	0.36	28.03	1.02	0.00	1.06
15	PGRC/E 207929	40.12	0.00	0.00	0.00	0.12	0.00	0.00	4.22	0.14	0.05	1.64	0.96	14.55	19.84	0.07	19.32	0.00	0.83	8.60	0.88	0.26	0.18	0.35	25.17	0.61	0.00	0.79
16	PGRC/E 207975	40.50	0.00	0.00	0.00	0.00	0.00	0.00	3.68	0.00	0.00	1.71	1.32	12.72	20.29	0.00	19.22	0.00	0.85	8.76	1.12	0.32	0.23	0.40	27.25	0.70	0.00	0.89
17	PGRC/E 20168/1	35.18	0.00	0.00	0.00	0.13	0.00	0.00	4.32	0.15	0.00	1.49	1.89	10.65	22.12	0.00	17.97	0.00	0.74	8.29	1.27	0.43	0.25	0.31	27.26	0.98	0.00	1.12
18	PGRC/E 20095/1	38.43	0.00	0.00	0.00	0.14	0.00	0.00	4.01	0.17	0.04	1.50	1.44	9.25	20.58	0.07	17.80	0.00	0.86	8.01	1.42	0.33	0.26	0.42	29.90	1.10	0.00	1.10
19	PGRC/E 21356/1	42.78	0.00	0.05	0.03	0.17	0.00	0.00	4.10	0.15	0.05	1.53	1.50	12.77	20.95	0.07	18.50	0.04	0.71	8.05	1.12	0.06	0.18	0.31	25.92	0.76	0.00	0.85
20	PGRC/E 21237	39.48	0.00	0.00	0.00	0.15	0.00	0.00	4.77	0.17	0.00	1.63	1.61	10.05	22.18	0.08	19.36	0.00	0.76	7.50	1.18	0.39	0.24	0.34	26.66	0.88	0.00	1.07
21	PGRC/E 21156	38.03	0.00	0.00	0.00	0.09	0.00	0.00	3.87	0.14	0.00	1.50	1.47	10.37	22.25	0.00	17.80	0.00	0.80	7.95	1.25	0.41	0.26	0.42	28.68	1.01	0.00	1.09
22	PGRC/E 20112/2	40.09	0.00	0.00	0.00	0.00	0.00	0.00	3.74	0.20	0.00	1.36	1.52	10.06	21.71	0.07	17.89	0.00	0.76	7.97	1.26	0.38	0.33	0.38	29.49	1.06	0.00	1.08
23	PGRC/E 200413	41.87	0.00	0.00	0.00	0.18	0.00	0.00	3.68	0.10	0.00	1.71	0.86	12.80	20.58	0.08	18.49	0.00	0.83	8.72	1.15	0.26	0.24	0.40	26.79	0.71	0.00	0.87
24	(4DXZem-1)X(Zem-1-AD/ 88)	39.42	0.00	0.00	0.00	0.09	0.00	0.00	4.28	0.16	0.00	1.52	1.61	11.12	22.83	0.00	18.22	0.00	0.75	7.48	1.27	0.34	0.21	0.37	26.95	0.93	0.00	0.97
25	(4DXZem-1)X(Zem-1-F5/10)	39.93	0.00	0.00	0.00	0.13	0.00	0.00	4.53	0.15	0.00	1.62	1.68	11.74	23.58	0.00	18.70	0.00	0.71	7.60	1.27	0.32	0.22	0.36	24.80	0.84	0.00	0.88
26	PGRC/E 20156	37.89	0.00	0.00	0.00	0.16	0.00	0.00	4.28	0.16	0.00	1.53	1.63	11.71	22.77	0.00	19.11	0.00	0.68	7.89	1.15	0.37	0.19	0.32	25.52	0.75	0.00	0.93
27	PGRC/E 21162/1	39.21	0.00	0.00	0.00	0.00	0.00	0.00	3.98	0.17	0.00	1.40	1.69	10.15	21.52	0.00	17.68	0.00	0.74	7.92	1.24	0.43	0.28	0.38	29.32	1.04	0.00	1.23
28	PGRC/E 20163/1	42.23	0.00	0.00	0.00	0.12	0.00	0.00	3.94	0.13	0.00	1.57	1.42	11.50	20.73	0.00	19.99	0.00	0.81	8.38	1.28	0.33	0.21	0.33	26.64	0.80	0.00	0.96
29	PGRC/E 20076/2	38.09	0.00	0.00	0.00	0.09	0.00	0.00	3.92	0.15	0.00	1.37	1.37	9.88	21.76	0.00	18.86	0.00	0.78	7.52	1.25	0.38	0.25	0.39	29.03	1.06	0.00	1.04
30	PGRC/E 20147/1	30.87	0.00	0.00	0.00	0.09	0.00	0.00	4.13	0.14	0.00	1.36	1.46	11.07	22.20	0.00	18.44	0.00	0.70	7.40	1.22	0.29	0.23	0.33	28.22	0.95	0.00	1.06
31	PGRC/E 208404	42.00	0.00	0.00	0.00	0.15	0.00	0.00	4.26	0.00	0.00	1.77	1.60	12.88	23.35	0.00	17.89	0.00	0.76	8.92	1.16	0.33	0.13	0.33	23.45	0.63	0.00	0.84
32	Yellow Dodolla	43.23	0.00	0.00	0.00	0.00	0.00	0.00	4.18	0.00	0.00	1.63	1.47	10.95	22.85	0.00	18.27	0.00	0.76	8.15	1.23	0.35	0.20	0.34	27.16	0.94	0.00	1.03
33	S-67-01	38.67	0.00	0.00	0.00	0.00	0.00	0.00	4.01	0.16	0.00	1.38	1.61	10.90	23.24	0.00	18.60	0.00	0.68	7.37	1.20	0.43	0.25	0.36	27.30	1.01	0.00	1.09
34	Adet	40.38	0.00	0.00	0.00	0.19	0.00	0.00	4.25	0.16	0.00	1.76	1.72	11.32	22.34	0.00	17.67	0.00	0.71	8.49	1.22	0.39	0.18	0.34	26.05	0.77	0.00	0.97
35	PGRC/E 21358/2	43.52	0.00	0.00	0.00	0.09	0.00	0.00	4.05	0.15	0.00	1.49	1.44	13.16	20.36	0.00	18.03	0.00	0.77	8.76	0.99	0.29	0.19	0.36	27.68	0.75	0.00	0.85
36	PGRC/E 207928	41.38	0.00	0.00	0.00	0.11	0.00	0.00	3.74	0.13	0.00	1.65	1.38	13.31	18.13	0.06	18.64	0.00	0.84	8.98	1.02	0.29	0.27	0.34	28.24	0.65	0.00	0.95
37	PGRC/E 21320/5	38.98	0.00	0.00	0.00	0.11	0.00	0.00	3.95	0.11	0.00	1.56	1.46	12.32	21.32	0.06	18.43	0.00	0.74	8.05	1.15	0.33	0.20	0.41	26.84	0.82	0.00	0.95
38	PGRC/E 20112/2	40.98	0.00	0.00	0.00	0.00	0.00	0.00	4.29	0.00	0.00	1.57	1.64	10.23	21.27	0.00	19.96	0.00	0.73	7.71	1.26	0.43	0.23	0.37	27.65	1.00	0.00	1.15
39	PGRC/E 20165	43.88	0.00	0.00	0.07	0.29	0.00	0.00	4.68	0.16	0.00	1.77	1.59	9.80	21.83	0.07	19.29	0.00	0.71	6.94	1.28	0.39	0.24	0.33	25.49	0.90	0.00	1.03
40	PGRC/E 21051	39.23	0.00	0.03	0.02	0.11	0.02	0.02	4.44	0.15	0.04	1.57	1.61	12.13	23.90	0.07	17.98	0.01	0.79	8.69	1.31	0.11	0.19	0.31	23.00	0.78	0.00	0.88
41	PGRC/E 21261-05	40.00	0.00	0.00	0.00	0.00	0.00	0.00	4.09	0.00	0.00	1.54	1.62	11.33	24.39	0.00	17.30	0.00	0.82	8.51	1.37	0.33	0.17	0.37	26.24	0.87	0.00	0.89
42	PGRC/E 210406	40.46	0.00	0.00	0.00	0.00	0.00	0.00	4.17	0.00	0.00	1.48	1.49	10.97	21.84	0.00	19.18	0.00	0.79	8.33	1.27	0.36	0.19	0.38	26.93	0.81	0.00	1.02
43	PGRC/E 21169	41.14	0.00	0.00	0.00	0.15	0.00	0.00	5.07	0.25	0.04	1.77	1.83	13.59	23.97	0.07	17.86	0.00	0.69	8.01	1.10	0.26	0.19	0.29	22.44	0.65	0.00	0.79
44	PGRC/E 21184	41.58	0.00	0.00	0.00	0.08	0.00	0.00	4.37	0.16	0.00	1.62	1.79	11.10	24.07	0.00	18.29	0.00	0.78	7.87	1.33	0.38	0.27	0.37	24.82	0.88	0.00	0.96
45	PGRC/E 20113	35.58	0.00	0.00	0.00	0.10	0.00	0.00	3.63	0.16	0.00	1.41	1.55	10.04	20.72	0.00	17.90	0.00	0.73	8.25	1.29	0.43	0.26	0.35	30.03	1.05	0.00	1.30
46	PGRC/E 20021	38.78	0.00	0.00	0.00	0.17	0.00	0.03	4.71	0.18	0.06	1.60	1.79	11.75	23.98	0.07	18.86	0.00	0.67	7.61	1.26	0.25	0.21	0.32	23.10	0.78	0.00	0.90
47	PGRC/E 21252	38.12	0.00	0.00	0.00	0.19	0.00	0.05	6.21	0.26	0.04	1.44	1.94	11.65	25.34	0.08	19.66	0.00	0.57	6.72	1.14	0.28	0.19	0.26	20.00	0.69	0.00	0.86
48	PGRC/E 20126	38.22	0.00	0.00	0.00	0.00	0.00	0.00	3.81	0.18	0.00	1.27	1.40	10.40	21.96	0.09	18.25	0.00	0.71	8.44	1.32	0.40	0.22	0.38	28.56	0.95	0.00	1.13
49	PGRC/E 21163	38.55	0.00	0.00	0.00	0.00	0.00	0.00	3.97	0.19	0.00	1.43	1.53	11.00	23.13	0.00	17.44	0.00	0.75	8.04	1.28	0.36	0.22	0.35	27.14	0.94	0.00	1.04
50	PGRC/E 20013	37.70	0.00	0.00	0.00	0.14	0.00	0.03	5.31	0.21	0.07	2.15	1.66	11.70	22.00	0.08	18.67	1.76	0.63	6.29	1.00	0.24	0.21	0.32	23.44	0.82	0.02	0.90
51	PGRC/E 21057/1	41.78	0.00	0.00	0.00	0.13	0.00	0.00	4.08	0.17	0.00	1.57	1.61	12.94	19.48	0.00	18.95	0.00	0.71	8.48	1.13	0.33	0.21	0.30	27.39	0.71	0.00	0.93
52	PGRC/E 20076/3	43.87	0.00	0.00	0.00	0.13	0.00	0.00	4.12	0.12	0.00	1.53	1.33	8.91	20.94	0.00	19.55	0.00	0.77	7.44	1.40	0.50	0.19	0.38	29.26	1.10	0.00	1.06
53	PGRC/E 20168/2	38.45	0.00	0.00	0.00	0.12	0.00	0.00	3.99	0.21	0.00	1.53	1.75	10.41	22.43	0.12	17.55	0.00	0.76	7.88	1.32	0.42	0.28	0.36	27.46	0.99	0.00	1.16
54	PGRC/E 200394	37.16	0.00	0.00	0.00	0.08	0.00	0.00	3.79	0.13	0.00	1.52	1.50	11.70	21.17	0.07	17.48	0.00	0.79	8.70	1.16	0.33	0.26	0.37	28.20	0.88	0.00	1.01
55	PGRC/E 207481	39.90	0.00	0.00	0.00	0.00	0.00	0.00	3.63	0.11	0.00	1.64	1.28	10.48	20.56	0.00	16.87	0.00	0.92	1.22	1.22	0.35	0.25	0.42	30.51	1.05	0.00	1.10
56	PGRC/E 21172	42.64	0.00	0.00	0.00	0.00	0.00	0.00	3.75	0.00	0.00	1.51	1.43	12.05	21.54	0.00	17.78	0.00	0.86	8.93	1.22	0.36	0.14	0.42	28.16	0.80	0.00	0.95
57	PGRC/E 20165/1	41.55	0.00	0.00	0.00	0.14	0.00	0.00	4.16	0.14	0.00	1.60	1.41	10.47	20.18	0.00	18.31	0.00	0.83	7.85	1.29	0.38	0.24	0.40	29.59	0.99	0.00	1.02

Table 4.2 (... continued).

No	Accessions/varieties	%Fat	C4:0	C11:0	C13:0	C14:0	C14:1c9	C15:0	C16:0	C16:1c9	C17:0	C18:0	C18:1c7	C18:1c9	C18:2c9,12	C18:3c6,9,12	C18:3c9,12,15	C19:0	C20:0	C20:1c11	C20:2c11,14	C20:3c8,11,14	C20:5c5,8,11,14,17	C22:0	C22:1c13	C22:2c13,16	C24:0	C24:1c15
58	PGRC/E 208401	33.14	0.00	0.00	0.00	0.10	0.00	0.00	3.53	0.08	0.00	1.63	1.33	12.23	19.43	0.00	16.68	0.00	0.93	10.24	1.13	0.28	0.21	0.40	29.38	0.79	0.00	0.90
59	PGRC/E 21223/2	40.18	0.00	0.03	0.00	0.15	0.00	0.04	4.75	0.18	0.06	1.44	1.63	11.94	23.08	0.07	18.54	0.03	0.69	7.39	1.16	0.21	0.21	0.33	24.38	0.76	0.00	0.95
60	PGRC/E 21224/3	36.17	0.00	0.00	0.00	0.08	0.00	0.00	3.59	0.12	0.00	1.37	1.48	9.01	20.66	0.00	16.77	0.00	0.80	7.97	1.28	0.48	0.36	0.44	31.92	1.21	0.00	1.38
61	PGRC/E 20168/3	39.00	0.00	0.00	0.00	0.12	0.00	0.00	4.57	0.12	0.00	1.55	1.27	10.97	21.91	0.08	19.04	0.00	0.76	7.49	1.20	0.29	0.21	0.39	27.32	0.88	0.00	0.85
62	PGRC/E 21328/1	39.09	0.00	0.00	0.00	0.12	0.00	0.00	4.16	0.18	0.00	1.73	1.79	11.75	23.01	0.00	17.89	0.00	0.85	9.26	1.34	0.36	0.22	0.34	24.73	0.76	0.00	0.83
63	PGRC/E 21162	26.78	0.00	0.00	0.00	0.08	0.00	0.00	4.19	0.15	0.00	1.50	1.55	12.22	21.11	0.00	17.96	0.00	0.81	8.58	1.22	0.34	0.23	0.36	27.09	0.79	0.00	0.98
64	PGRC/E 21235/1	43.47	0.00	0.00	0.00	0.12	0.00	0.00	3.81	0.19	0.00	1.50	1.41	10.39	19.82	0.00	18.30	0.00	0.87	8.58	1.22	0.36	0.23	0.43	30.11	0.92	0.00	0.99
65	PGRC/E 21263	38.38	0.00	0.00	0.00	0.11	0.00	0.00	4.24	0.16	0.00	1.53	1.65	9.77	22.08	0.00	18.66	0.00	0.84	7.79	1.30	0.43	0.23	0.36	28.09	1.04	0.00	1.07
66	PGRC/E 21057	38.88	0.00	0.00	0.00	0.08	0.00	0.00	3.96	0.22	0.00	1.39	1.73	9.20	21.68	0.00	18.28	0.00	0.72	7.43	1.35	0.51	0.26	0.41	29.56	1.10	0.00	1.26
67	PGRC/E 21361-03	34.98	0.00	0.00	0.00	0.10	0.00	0.00	4.80	0.16	0.00	1.64	1.59	9.41	24.91	0.00	17.60	0.00	0.78	6.84	1.33	0.41	0.29	0.44	26.40	1.02	0.00	0.92
68	PGRC/E 21261-04	36.50	0.00	0.00	0.00	0.00	0.00	0.00	4.61	0.18	0.00	1.82	1.99	10.38	22.63	0.00	20.65	0.00	0.85	6.70	1.15	0.51	0.27	0.41	25.37	0.91	0.00	0.78
69	PGRC/E 21001	39.45	0.00	0.00	0.00	0.00	0.00	0.00	3.95	0.16	0.00	1.41	1.68	10.94	23.01	0.00	19.32	0.00	0.69	7.52	1.41	0.40	0.13	0.29	26.75	0.89	0.00	0.97
70	PGRC/E 21261-01	43.32	0.00	0.00	0.00	0.00	0.00	0.00	4.01	0.15	0.00	1.52	1.45	9.67	23.01	0.00	17.88	0.00	0.87	7.58	1.44	0.38	0.22	0.43	28.76	1.02	0.00	0.94
71	PGRC/E 21324	40.64	0.00	0.00	0.00	0.00	0.00	0.00	4.16	0.11	0.00	1.63	1.76	12.81	23.25	0.00	20.36	0.00	0.75	8.82	1.24	0.37	0.25	0.31	22.29	0.73	0.00	0.85
72	PGRC/E 21163/1	41.80	0.00	0.00	0.00	0.00	0.00	0.00	3.96	0.11	0.00	1.33	1.56	10.94	22.35	0.00	18.19	0.00	0.67	8.04	1.24	0.35	0.26	0.33	27.86	0.90	0.00	1.11
73	PGRC/E 20163	39.20	0.00	0.00	0.02	0.16	0.00	0.03	4.79	0.20	0.04	1.43	1.65	9.55	22.84	0.07	19.89	0.00	0.70	7.06	1.33	0.32	0.23	0.34	25.36	0.98	0.00	0.97
74	PGRC/E 21324/1	42.10	0.00	0.00	0.00	0.10	0.00	0.00	4.17	0.15	0.00	1.62	1.64	13.60	22.41	0.00	18.73	0.00	0.77	9.99	1.26	0.31	0.20	0.32	22.51	0.65	0.00	0.81
75	PGRC/E 21236/1	44.67	0.00	0.00	0.00	0.00	0.00	0.00	3.91	0.00	0.00	1.49	1.35	11.47	20.25	0.00	20.34	0.00	0.72	8.18	1.14	0.35	0.17	0.35	27.98	0.81	0.00	0.96
76	PGRC/E 20104	35.07	0.00	0.00	0.00	0.76	0.00	0.05	16.42	1.04	0.25	7.43	1.28	24.64	17.21	0.04	10.50	0.00	0.47	4.36	0.77	0.18	0.13	0.17	11.59	0.37	0.00	0.42
77	PGRC/E 21170	39.82	0.00	0.00	0.00	0.00	0.00	0.00	4.18	0.13	0.00	1.38	1.57	10.97	22.38	0.00	19.05	0.00	0.70	7.83	1.33	0.37	0.21	0.37	27.27	0.91	0.00	1.01
78	PGRC/E 208410	41.88	0.00	0.00	0.00	0.00	0.00	0.00	3.68	0.00	0.00	1.49	1.41	11.54	20.71	0.00	17.72	0.00	0.81	9.34	1.30	0.33	0.20	0.39	28.73	0.83	0.00	0.95
79	PGRC/E 207931	41.32	0.00	0.05	0.03	0.14	0.00	0.03	4.31	0.17	0.04	1.44	1.44	12.29	21.35	0.08	18.50	0.00	0.73	7.72	1.14	0.10	0.22	0.36	26.02	0.76	0.00	0.88
80	PGRC/E 208004	41.30	0.00	0.00	0.00	0.00	0.00	0.00	4.25	0.13	0.00	1.57	1.69	13.91	24.38	0.00	16.01	0.00	0.78	10.15	1.22	0.27	0.17	0.34	22.59	0.67	0.00	0.81
81	PGRC/E 21304	42.54	0.00	0.00	0.00	0.09	0.00	0.00	4.21	0.13	0.00	1.46	1.43	11.43	21.96	0.00	19.30	0.00	0.76	8.17	1.29	0.32	0.20	0.31	26.38	0.86	0.00	0.89
82	PGRC/E 21010	43.50	0.00	0.00	0.00	0.00	0.00	0.00	3.38	0.00	0.00	1.32	1.30	10.31	17.38	0.00	15.31	0.00	0.87	9.13	1.28	0.45	0.26	0.49	35.01	1.04	0.00	1.31
83	PGRC/E 21031	41.21	0.00	0.00	0.00	0.26	0.00	0.00	4.04	0.20	0.00	1.86	1.73	10.20	21.84	0.00	17.32	0.00	0.75	7.78	1.21	0.42	0.30	0.42	27.61	0.96	0.00	1.14
84	PGRC/E 20153	37.36	0.00	0.00	0.00	0.12	0.00	0.00	4.31	0.17	0.00	1.50	1.49	10.53	22.51	0.06	17.48	0.00	0.74	7.98	1.25	0.35	0.27	0.40	27.89	0.90	0.00	1.04
85	PGRC/E 20126/1	36.02	0.00	0.04	0.00	0.14	0.00	0.02	4.58	0.20	0.06	1.41	1.73	11.43	24.08	0.07	18.74	0.02	0.66	7.48	1.23	0.20	0.21	0.30	23.42	0.84	0.00	0.90
86	PGRC/E 20120	34.84	0.00	0.00	0.00	0.11	0.00	0.00	4.08	0.18	0.00	1.44	1.56	9.42	21.76	0.07	17.72	0.00	0.81	7.92	1.29	0.46	0.26	0.37	29.29	1.15	0.00	1.16
87	PGRC/E 208551	35.33	0.00	0.00	0.00	0.15	0.00	0.00	4.05	0.17	0.00	1.46	1.59	9.96	21.68	0.00	17.35	0.00	0.75	7.52	1.19	0.41	0.32	0.40	29.83	1.04	0.00	1.23
88	PGRC/E 20103	37.08	0.00	0.00	0.00	0.18	0.00	0.00	4.64	0.20	0.00	1.73	1.68	11.91	22.85	0.05	17.78	0.00	0.75	8.44	1.18	0.33	0.23	0.35	24.21	0.78	0.00	0.93
89	PGRC/E 21373	37.04	0.00	0.00	0.00	0.00	0.00	0.00	4.01	0.00	0.00	1.42	1.66	11.28	23.16	0.00	18.53	0.00	0.85	8.41	1.30	0.40	0.21	0.35	26.38	0.88	0.00	0.96
90	PGRC/E 20164	35.09	0.00	0.03	0.02	0.17	0.00	0.03	4.74	0.22	0.04	1.45	1.74	10.23	23.97	0.07	18.60	0.03	0.70	7.01	1.28	0.21	0.26	0.35	24.75	0.93	0.00	1.04
91	PGRC/E 20165/2	38.36	0.00	0.00	0.00	0.00	0.00	0.00	4.29	0.13	0.00	1.50	1.61	10.50	22.28	0.00	18.49	0.00	0.74	8.11	1.27	0.41	0.25	0.37	27.38	1.03	0.00	1.05
92	PGRC/E 20090/1	39.27	0.00	0.00	0.00	0.11	0.00	0.00	4.13	0.17	0.00	1.58	1.56	9.59	22.66	0.00	17.79	0.00	0.75	7.80	1.27	0.45	0.26	0.33	28.41	1.01	0.00	1.16
93	PGRC/E 20162/1	38.80	0.00	0.00	0.00	0.26	0.00	0.00	4.19	0.20	0.00	1.70	1.73	9.48	22.46	0.00	17.96	0.00	0.77	7.11	1.38	0.46	0.26	0.35	27.73	1.06	0.00	1.21
94	PGRC/E 21169/1	41.92	0.00	0.00	0.00	0.00	0.00	0.00	3.54	0.00	0.00	1.47	1.36	12.39	18.76	0.00	18.69	0.00	0.87	9.18	1.06	0.35	0.20	0.37	29.82	0.76	0.00	0.91
95	PGRC/E 20175/1	37.69	0.00	0.00	0.00	0.00	0.00	0.00	4.57	0.12	0.00	1.59	1.85	10.43	24.22	0.00	18.80	0.00	0.67	7.42	1.20	0.48	0.26	0.33	25.31	0.95	0.00	1.01
96	PGRC/E 21058/2	40.52	0.00	0.00	0.00	0.14	0.00	0.00	3.85	0.16	0.00	1.52	1.30	9.53	19.80	0.00	17.04	0.00	0.88	8.25	1.31	0.35	0.29	0.47	31.86	1.06	0.00	1.16
97	PGRC/E 21261	40.36	0.00	0.00	0.00	0.00	0.00	0.00	4.22	0.00	0.00	1.53	1.48	11.23	23.39	0.00	17.15	0.00	0.77	8.13	1.33	0.35	0.29	0.36	27.60	0.89	0.00	1.03
98	PGRC/E 20163/5	42.10	0.00	0.00	0.00	0.00	0.00	0.00	3.78	0.11	0.00	1.56	1.30	10.35	19.85	0.00	19.00	0.00	0.84	8.50	1.27	0.34	0.25	0.40	29.62	0.88	0.00	1.01

The fatty acid composition of seed samples revealed a large variation for most of the fatty acids examined. The variation, however, was found to be greater for erucic acid (6.91 – 35.01 %), linoleic acid (17.21 – 28.53 %), oleic acid (8.91 – 24.64%), α-linolenic acid (10.5 – 22.52 %), palmitic acid (3.38 –16.42 %) and gadoleic or eicosenoic acid (1.22 –10.24%) in all accessions. To a lesser extent vaccinic acid (0.86-3.1), nervonic acid (0.42-1.38%), and behenic acid (0.17-0.51%) were also found in all accessions.

The three-monounsaturated fatty acids: erucic, oleic, and eicosenoic acids comprised a significant proportions of the fatty acids found in the accessions (Tables 4.2 and 4.3). The two primary polyunsaturated fatty acids found in Ethiopian mustard are linoleic and α-linolenic acid with a mean of 22 % and 18 % fatty acid concentration respectively. The predominance of erucic acid was in agreement of the findings of other researchers (Alemayehu and Becker, 2001; Tahoun *et al.*, 1999; Auld *et al.* 1989; Mnzava and Olsson, 1990). The identified fatty acids in this study were composed of approximately 5.96-25.55% saturated, 72.66-93.48% unsaturated, 37.33-57.05% monounsaturated, 29.18-52.27% polyunsaturated and 27.70-51.05% essential fatty acids.

It would appear that in the absence of artificial selection, most accessions of *B. carinata* have an average fatty acid composition of approximately 4 % palmitic acid, 2 % stearic acid, 2 % vaccinic acid, 11 % oleic acid, 22 % linoleic acid, 18 % α-linolenic acid, 8 % eicosenoic acid, and 27 % for erucic acid. Auld *et al.* (1989) in their analysis of fatty acid composition of Ethiopian mustard reported pamitic, stearic, oleic, linolenic, eicosenoic, and erucic acids with a mean of 3.8, 0.9, 10.7, 17.1, 15.0, 7.6, and 42.6 % respectively. Alemayehu and Becker (2001) studied inbred lines as well as F_2 populations of *B. carinata*. They found a wide variation in fatty acid contents. Oleic acid varied from 5 to 34% and erucic acid content from 6 to 51%.

Table 4.3. Descriptive statistics, molecular formula, trival names and fatty acid composition (as % of total fatty acids) of 98 Ethiopian mustard accessions/varieties determined by capillary gas chromatography.

Trival name	IUPAC/Systematic name[a]	Abbreviations[b]	Molecular formula	Mean	Min	Max	Std.Dev.	Range	CV%
SFA									
Butyric	Butanoic	4:0	$CH_3(CH_2)_2COOH$	-	0.02	0.02	0.00	0.00	0.00
	Undecanoic	11:0	$CH_3(CH_2)_9COOH$	0.04	0.03	0.05	0.01	0.02	0.25
	Tridecanoic	13:0	$CH_3(CH_2)_{11}COOH$	0.03	0.02	0.07	0.01	0.05	0.60
Myristic	Teteradecanoic	14:0	$CH_3(CH_2)_{12}COOH$	0.09	0.76	0.76	0.00	0.76	1.07
	Pentadecenoic	15:0	$CH_3(CH_2)_{13}COOH$	0.03	0.02	0.05	0.01	0.03	0.29
Palmitic	Hexadecanoic	16:0	$CH_3(CH_2)_{14}COOH$	4.32	3.38	16.42	1.32	13.4	0.30
Margaric	Heptadecanoic	17:0	$CH_3(CH_2)_{15}COOH$	0.01	0.00	0.25	0.03	0.25	2.83
Stearic	Octadecanoic	18:0	$CH_3(CH_2)_{16}COOH$	1.61	1.22	7.43	0.61	6.21	0.38
	Nonadecanoic	19:0	$CH_3(CH_2)_{17}COOH$	0.22	0.01	1.76	0.55	1.75	2.40
Eicosanoic	Arachidic	20:0	$CH_3(CH_2)_{18}COOH$	0.76	0.47	0.93	0.08	0.46	0.10
Behenic	Docosanoic	22:0	$CH_3(CH_2)_{20}COOH$	0.36	0.17	0.51	0.06	0.3	0.16
Lignoceric	Teteracosanoic	24:0	$CH_3(CH_2)_{22}COOH$	-	0.02	0.02	0.00	0.00	0.00
MUFA									
Myristoleic acid	9-Teteradecenoic	14:1c9	$CH_3(CH_2)_3CH=CH(CH_2)_7COOH$	-	0.02	0.02	0.00	0.00	0.00
Palmitoleic acid	9-Hexadecenoic	16:1c9 (n-7)	$CH_3(CH_2)_5CH=CH(CH_2)_7COOH$	0.15	0.00	1.04	0.11	1.04	0.76
Vaccinic	11-Octadecenoic	18:1c7 (n-7)	$CH_3(CH_2)_5CH=CH(CH_2)_9COOH$	1.58	0.86	3.10	0.29	2.24	0.19
Oleic	9-Octadecenoic	18:1c9 (n-9)	$CH_3(CH_2)_7CH=CH(CH_2)_7COOH$	11.43	8.91	24.64	2.33	15.73	0.24
Gadoleic	9-Eicosenoic	20:1c11 (n-11)	$CH_3(CH_2)_9CH=CH(CH_2)_7COOH$	7.87	1.22	10.24	1.12	9.02	0.14
Erucic acid	13-Docosenoic	22:1c13 (n-9)	$CH_3(CH_2)_7CH=CH(CH_2)_{11}COOH$	26.51	6.91	35.01	4.19	28.01	0.16
Nervonic acid	15-Teteracosenoic	24:1c15 (n-9)	$CH_3(CH_2)_7CH=CH(CH_2)_{13}COOH$	0.98	0.42	1.38	0.16	0.96	0.16
PUFA									
Linoleic	9,12-Octadecadienoic	18:2c9,12 (n-6)	$CH_3(CH_2)_4(CH=CHCH_2)_2(CH_2)_6COOH$	22.09	17.21	28.53	1.87	11.32	0.08
α-linolenic	9,12,15-Octadecatrienoic	18:3c9,12,15 (n-3)	$CH_3(CH_2)(CH=CHCH_2)_4(CH_2)_6COOH$	18.36	10.50	22.52	1.36	12.02	0.07
γ-Linolenic	6,9,12-Octadecatrienoic	18:3c6,9,12 (n-6)	$CH_3(CH_2)_4(CH=CHCH_2)_3(CH_2)_3COOH$	0.02	0.00	0.12	0.04	0.12	1.52
	11,14-Eicosadienoic	20:2c11,14	$CH_3(CH_2)_4(CH=CHCH_2)_2(CH_2)_8COOH$	1.23	0.68	1.45	0.13	0.77	0.11
Dihomo-γ-linolenic	8,11,14-eicosatrienoic	20:3c8,11,14 (n-6)	$CH_3(CH_2)_4(CH=CHCH_2)_3(CH_2)_5COOH$	0.35	0.06	0.53	0.09	0.47	0.26
Timnodonic (EPA)	5,8,11,14,17-Eicosapentaenoic	20:5c5,8,11,14,17 (n-3)	$CH_3CH_2(CH=CHCH_2)_5(CH_2)_2COOH$	0.23	0.10	0.36	0.05	0.26	0.20
	13,16-Docosadienoic	22:2c13,16	$CH_3(CH_2)_4(CH=CHCH_2)_2(CH_2)_{10}COOH$	0.88	0.22	1.21	0.18	0.99	0.20
ER				0.39	0.13	0.50	0.49	0.36	0.13
DR				0.46	0.39	0.58	0.03	0.21	0.06
ODR				0.76	0.52	0.80	0.03	0.28	0.04
LDR				0.45	0.38	0.51	0.02	0.13	0.05
O/L				0.52	0.38	1.43	0.12	1.40	0.23
U/S				8.99	2.19	10.69	0.91	8.5	0.10
n-6/n-3				1.21	0.97	1.64	0.11	0.67	0.09
n-3				18.59	10.62	22.62	1.35	12.00	0.07
n-6				22.46	17.42	28.75	1.86	11.33	0.08
n-7				1.73	0.96	3.38	0.33	2.42	0.19
n-9				38.92	28.96	46.62	2.65	17.66	0.07
SFA				7.18	5.96	25.55	1.97	19.59	0.07
USFA				91.69	72.66	93.48	2.32	20.82	0.03
MUFA				48.54	37.33	57.05	2.89	19.72	0.06
PUFA				43.15	29.18	52.27	2.65	23.09	0.06
EFA				40.45	27.70	51.05	2.70	23.35	0.07
Total				98.87	91.61	99.89	0.93	8.28	0.01

ER=Elongation ratio; DR=Desaturation ratio; ODR=Oleic desaturation ratio; LDR=Linoleic desaturation ratio; O/L= Oleic to linoeic ratio; U/S=unfaturation ratio; EFA=Essential fatty acid;

MUFA=Monounsaturate fatty acid; PUSFA=Polyunsaturated fatty acid; USFA=Unsaturate fatty acid; SFA=Saturated fatty acid;

[a]The full designation for double bonds would be all-*cis*, for example, all-*cis*-9,12,15-Octadecatrienoic acid. Unless otherwise indicated as *t* for trans double bond, all bonds are of the *cis* geometric configuration.

[b]The name of the fatty acid consists of a number which indicates the number of carbon atoms in the acid and the number of unsaturated centers (i.e., 18:3 carbon atoms with 3 double bonds).

The c denotes a cis double bond and the position is indicated by the number i.e., 9c.

The accessions with a low level of erucic acid (PGRC/E 20104, C94-S-67, C94-Dodolla, and PGRC/E 203221-03) had high levels of oleic acid (Table 4.2). The higher the erucic acid content, the lower the amount of oleic acid. This can be explained by the fact that the former is synthesized through the elongation of two molecules of active acetate without shift in the double-bond position of the starting synthetic material (oleic acid). The relationship between the syntheses of the two fatty acids is very strong. This was also confirmed by Tahoun *et al.* (1999) on *B. carinata* and *B. napus* accessions.

The accessions with a low level of erucic acid showed almost a directly proportional increase in the concentration of oleic acid. The high level of oleic acid occurs when the fatty acid elongation pathway responsible for the formation of eicosenoic acid and erucic acid is genetically blocked.

Accession PGRC/E 20104 showed the lowest α-linolenic acid concentration (10.50%) among all accessions studied. However, it also had a low level of erucic acid (11.59) unlike accessions C94-S-67 and C94-Dodolla, which had 22.52 and 22.20% α-linolenic acid respectively (Table 4.2).

Brassica carinata oil contains significant amounts of nutritionally important fatty acids, namely α-linolenic acid [ALA, 18:3(n-3), about 18.36% %]. ALA is an essential fatty acid that is metabolized to eicosapentaenoic acid (EPA), a precursor of eicosanoids with anti-inflammatory and antithrombotic activity (Ruiz *et al.*, 2002).

In general, there are two well-defined patterns. The first one showed a larger accumulation of polyunsaturated fatty acids, C18:2 and C18:3. The second one was characterized by the main formation of long-chain monounsaturated fatty acids, C20:1 and C22:1.

Fatty acid ratios

The values of the fatty acid ratio's ER and DR in the accessions were calculated and the average values for the various accessions are shown in Table 4.3. These ratios were very useful to get a characterization of the relative efficiency of the elongation and desaturation pathways, and their use will help to design strategies for the breeding

management of this species (Velasco *et al.*, 1998). *Brassica carinata* accessions were characterized by one of the highest average values of DR and therefore by relatively high desaturation efficiency. Furthermore, the average value of ODR for all accessions was 0.76 and for LDR 0.46. These values explain the large increase of C18:3 content (>18%) produced in this species when C22:1 was eliminated through plant breeding (Alonso *et al.*, 1991; Getinet *et al.*, 1994).

The accessions C94-S-67, C94-Dodolla, PGRC/E 20104 and PGRC/E 203221-03 were clearly characterized by very high values of DR and low values of ER, which reveals a higher efficiency of the desaturation pathway than other accessions (Figure 4.1). Figure 4.2 shows the scatter plot of ODR versus LDR, i.e., the efficiency of the desaturation systems from C18:1 to C18:2 and from C18:2 to C18:3, respectively. Genotype PGRC/E 20104 showed very low ODR and LDR. PGRC/E 207928, PGRC/E 20059 and PGRC/E 21236/1 showed high LDR and ODR (Figure 4.2). The oleic to linoleic ratio (O/L) is an indicator of oil stability and shelf life. Higher O/L ratios would suggest better stability and longer shelf life. In addition, to increasing stability of oil, increasing the O/L ratio appears to have health benefits as well (O'Bryne *et al.*, 1997; Renaud *et al.*, 1995). Accordingly accession PGRC/E 20104 and C94-Dodolla had the best O/L ratios of 1.43 and 0.76 respectively.

Polyunsaturated FAs belong to two main series (n-3) and (n-6). PUFAs of the (n-3) series ranges from 10.62 to 22.62% of the observed fatty acids. PUFAs of (n-6) series ranges from 17.42% to 28.75%. The n-6/n-3 ratio of the fatty acids was found to be 0.61 to 1.03. This result showed that *B. carinata* is rich in the n-3 and n-6 series of PUFA's, which are considered essential fatty acids for humans and animals; for example linoleic acid, linolenic acid and arachidonic acids have an important function in growth and protection of the skin (Li *et al.*, 2002).

It is felt that vegetable oils used for dietary purposes should have saturated, monounsaturated and polyunsaturated fatty acids in about equal proportions or at least in the ratio of 2:1 (unsaturated: saturated) (Murphy, 1993). Only one accession, PGRC/E 20104 was found to have fatty acid composition, saturated: monounsaturated: polyunsaturated nearing 0.6:1:0.7 or unsaturated: saturated nearing 2.8:1.

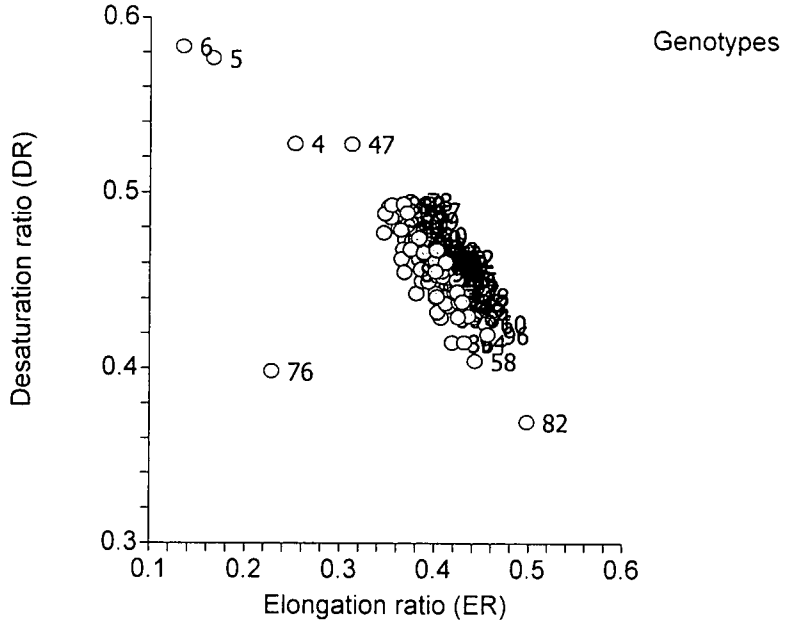


Figure 4.1. Scatter plot of elongation ratio (ER) vs. desaturation ratio (DR) in 98 Ethiopian mustard genotypes. The ratios were calculated from the CGC data of individual fatty acids.

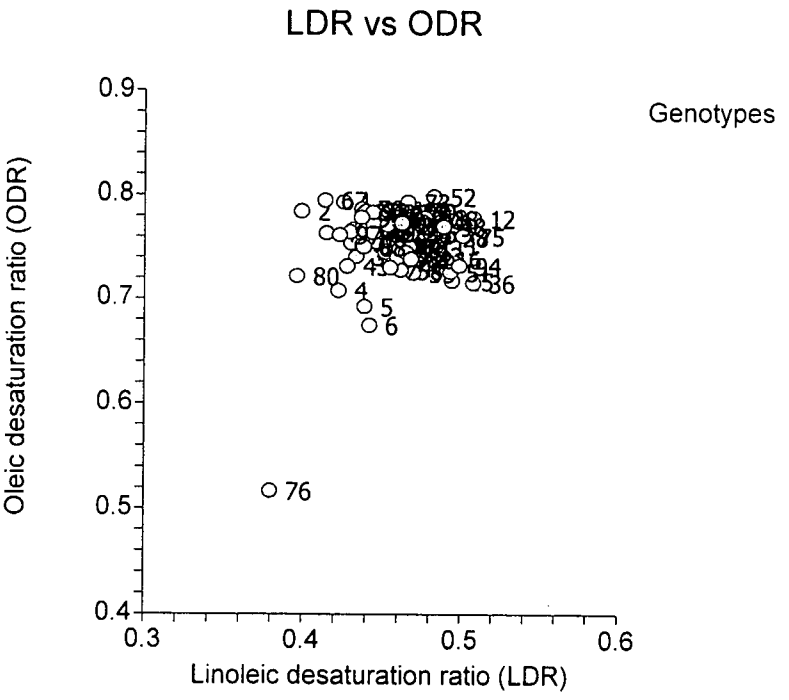


Figure 4.2. Scatter plots of linoleic desaturation ratio (LDR) vs. oleic desaturation ratio (ODR) in 98 Ethiopian mustard genotypes. The ratios were calculated from CGC data of individual fatty acids.

LDR was negatively correlated with palmitic, stearic, vaccinic, oleic acid, and DR. It was also positively correlated with α-linolenic acid, erucic acid, and ER. Since ODR and LDR are not related significantly, a reduction of ODR has no influence on LDR activity and vise versa.

ODR was positively correlated with linolenic acid. Since linolenic acid content is mainly influenced by OD-activity, a reduction of OD decreases the amount of substrate for further desaturation steps, which finally leads to reduced C18:3 contents.

Relationship between oil content and fatty acid composition

The results of correlation analyses between oil content and fatty acid composition and within the fatty acid composition *per se* in oils of different accessions examined are given in Table 4.4. Oil content had negative correlations with palmitic, stearic, vaccinic, oleic and linoleic acids and positive correlations with α-linolenic, eicosenoic, and erucic acids. The palmitic acid content was significantly positively correlated ($P<0.0001$) with stearic acid content. The long chain erucic acid showed highly significant negative correlation ($P<0.0001$) with palmitic acid, stearic acid, vaccinic acid, oleic acid, linoleic acid, and α-linolenic acids. It also showed positive correlations with gadoleic or eicosenoic acid. The decrease in the concentration of erucic acid results in an increase of C16 and C18 fatty acids.

A strong positive correlation ($P<0.0001$) was found between linoleic and α-linolenic acids in the Ethiopian mustard. Kumar and Tsunoda (1978) had also reported positive correlations between linoleic and α-linolenic acids within *Brassicea.* There was also highly significant positive correlation ($P<0.0001$) between vaccinic and linoleic acids.

Since linoleic acid (18:2) and linolenic acid (18:3) contents are positively correlated, selection of genotypes with low linolenic and high linoleic contents may be difficult, but nevertheless feasible (Pleines and Friedt, 1988).

There were strong significant positive correlations between stearic and palmitic, and oleic acids, while a negative correlation was found between stearic, and vaccinic, linoleic, α-linolenic, eicosenoic and erucic acids.

A strong positive correlation between oil content and erucic acid and a weak negative correlation between oil content and oleic, vaccinic, and linoleic acid contents of the seed oil can be noted from Table 4.5.

Positive correlations indicate that selection directed towards increased linoleic acid also would increase palmitic and stearic acids. Negative correlations, however, indicate that selection to increase palmitic or stearic acid would cause a decrease in the proportion of erucic acid. These correlations may be explained as a function of the biosynthetic pathways of the *B. carinata* fatty acids, which are not fully elucidated yet. Highly negative correlations between erucic acid and palmitic, stearic, oleic, and linoleic indicate also that they might participate in the erucic acid biosynthetic pathways. Similar associations were made to explain the interrelationships among the fatty acids of linseed oil (Green and Marshall, 1981).

Table 4.4. Matrix of simple phenotypic correlation coefficients for oil content and fatty acid composition of Ethiopian mustard genotypes.

Variable	PAL															
Palmitic	**1.0**	STE														
Stearic	0.95**	**1.0**	VAC													
Vaccinic	0.12ns	-0.01ns	**1.0**	OLE												
Oleic	0.64**	0.65**	0.41**	**1.0**	LIN											
Linoleic	-0.02ns	-0.19ns	0.75**	0.18ns	**1.0**	ALA										
α-Linolenic	-0.44**	-0.53**	0.44**	-0.03ns	0.38**	**1.0**	EIC									
Eicosenoic	-0.40**	-0.35**	-0.22ns	-0.16ns	-0.20ns	-0.01ns	**1.0**	ERU								
Erucic	-0.57**	-0.47**	-0.72**	-0.84**	-0.61**	-0.23*	0.28**	**1.0**	ER							
ER	-0.55**	-0.44**	-0.74**	-0.81**	-0.64**	-0.29**	0.42**	0.98**	**1.0**	DR						
DR	0.06ns	-0.11ns	0.78**	0.26**	0.89**	0.63**	-0.39**	-0.71**	-0.77**	**1.0**	ODR					
ODR	-0.73**	-0.78**	-0.17ns	-0.93**	0.13ns	0.32**	0.10ns	0.64**	0.58**	0.07ns	**1.0**	LDR				
LDR	-0.43**	-0.35**	-0.30**	-0.23**	-0.58**	0.53**	0.22*	0.39**	0.37**	-0.28**	0.20ns	**1.0**	EFA			
EFA	-0.23*	-0.40**	0.74**	0.11ns	0.88**	0.77**	-0.14ns	-0.54**	-0.59**	0.94**	0.25**	-0.14ns	**1.0**	SFA		
SFA	0.99**	0.98**	0.05ns	0.63**	-0.10ns	-0.49**	-0.40**	-0.53**	-0.50**	-0.01ns	-0.74**	-0.40**	-0.31**	**1.0**	MUFA	
MUFA	-0.44**	-0.28**	-0.70**	-0.45**	-0.75**	-0.35**	0.62**	0.82**	0.88**	-0.90**	0.20ns	0.42**	-0.70**	-0.39**	**1.0**	PUFA
PUFA	-0.30**	-0.47**	0.70**	-0.02ns	0.87**	0.76**	-0.13ns	-0.44**	-0.50**	0.91**	0.37**	-0.13ns	0.99**	-0.38**	-0.65**	**1.0**
Oil content	-0.19ns	-0.14ns	-0.36**	-0.22*	-0.30**	0.03ns	0.25**	0.29**	0.32**	-0.26**	0.14ns	0.31**	-0.19ns	-0.18ns	0.30**	-0.18ns

PAL=Palmitic, STE=Stearic, VAC=Vaccinic, OLE=Oleic, LIN=Linoleic, ALA=α-linoleinic , EIC=Eicosenoic, ERU=Erucic

* P=0.05; **P=0.01

Table 4.5. Fatty acid composition of the seed oil of 98 Ethiopian mustard accessions differing in oil content

% oil	Accessions	% of FA's								
		Palmitic	**Stearic**	**Vaccinic**	**Oleic**	**Linoleic**	**Linolenic**	**Eicosenoic**	**Icosadienoic**	**Erucic**
≤30	2	4.42±0.13	1.73±0.13	2.2±0.38	15.98±2.17	24.81±2.14	20.08±1.22	7.25±0.77	1.02±0.11	17.91±5.30
30.1-35	5	4.40±0.22	1.57±0.05	1.81±0.22	12.74±1.52	23.37±1.04	18.59±0.68	1.13±0.078	7.46±0.58	24.04±2.88
35.1-40	46	4.55±0.32	1.67±0.15	1.59±0.03	11.00±0.38	18.14±0.24	18.14±0.24	7.89±0.21	1.24±0.019	26.85±0.57
40.1-45	44	4.10±0.04	1.56±0.02	1.51±0.03	11.59±0.17	19.45±0.00	18.44±0.12	8.31±0.07	1.22±0.012	26.90±0.35
≥45.1	1	4.22± 0.00	1.54±0.00	1.51 ±0.00	11.42±0.0 0	19.45±0.00	19.45±0.00	7.67±0.00	1.11±0.00	27.06±0.00

[a] Each fatty acid contents is given as mean ± standard error of the mean.

The Ethiopian mustard germplasm were found to contain considerable variability for oil content and fatty acid compositions. The accessions varied widely in their oil content and fatty acid composition. The major fatty acids in the seed oil on average basis were erucic>linoleic>α-linolenic>oleic>gadoleic>palmitic. Gamma linolenic, lignoceric, dimo-γ-linolenic, margaric, palmitoleic, myristic were also rarely detected and only in small number of accessions. The high oil yielding accessions were relatively richer in erucic acid content. There was significant and negative correlation between erucic acid and palmitic, stearic, vaccinic, linoleic, α-linolenic acids and positive correlations with eicosenoic acid.

4.5. REFERENCES

ALEMAYEHU, N., and H.C. BECKER, 2001. Variation and inheritance of erucic acid content in *Brassica carinata* germplasm collections from Ethiopia. *Plant Breeding* 120:331-335.

ALONSO, L., C. FERNÁNDEZ-SERRANO, O. and J. FERNÁNDEZ-ESCOBAR, 1991. The onset of a new oilseed crop: *Brassica carinata* with low erucic acid content. *Proc. 8th Int. Rape seed Congr.*, Saskatoon, SK, Vol. I, pp. 170-176.

AULD, D.I., K.A. MAHLER and D.J. LETOURNEAU, 1989. Evaluation of four *Brassica* germplasm collections for fatty acid composition. *JAOCS* 66(10): 1475-1479.

BAYEH. M., and T. GEBRE MEDHIN, 1992. Insect pests of noug, linseed and oilseed *Brassica* in Ethiopia. *In*: Oilseed research and development in Ethiopia, *Proceedings of the First National Oilseeds Workshop*, 3-5 December 1991, Addis Ababa, Ethiopia. pp. 174-178

BECKER, H.C., H. LÖPTIEN, and G. RÖBBELEN, 1999. Breeding: an overview. *In*: C. Gomez-Campo (ed.). Biology of *Brassica* Coeno-species. Elsevier, Science BV, Amsterdam. pp. 413-460.

BYRUM, J.R., A.J. KINNEY, K.L. STECCA, D.J. GRACE and B.W. DIERS, 1997. Alteration of the omega-3 fatty acid desaturase gene is associated with reduced linolenic acid in the A5 soybean genotype. *Theor. Appl. Genet.* 94: 356-359.

CHEESBROUGH, T.M., K. SNOW, M.C. SCHNEERMAN and D.F.WEBER, 1997. Fatty acid composition of maize races from central and south America. *Maydica* 42: 155-161.

FERNÁNDEZ-ESCOBAR, J., J. DOMINGUEZ, A. MARTI, and J. FERNÁNDEZ-MARTINEZ, 1988. Genetics of erucic acid content in interspecific hybrids of Ethiopian mustard (*B. carinata* A. Braun) and Rapeseed (*B. napus* L.). *Plant Breeding* 100: 310-315.

FERNÁNDEZ-MARTINEZ, J., M. DEL RIO, and A. DE HARO, 1993. Survey of safflower (*Carthamus tinctorius* L.) germplasm for variants in fatty acid composition and other seed characters. *Euphytica* 69: 115-122.

FOLCH, J., M. LEES, and G.M. SLOONE-STANLEY, 1957. A simple method for the isolation and purification of total lipids from animal tissue. *J. Biol. Chem.* 226: 497-509.

GETINET, A., G. RAKOW, J.P. RANEY, and R.K. DOWNEY, 1994. Development of zero erucic acid Ethiopian mustard through an interspecific cross with zero erucic acid Oriental mustard. *Can J. Plant Sci.* 74: 793-795.

GREEN, A.G., and D.R. MARSHALL, 1981. Variation for oil quantity and quality in linseed (*Linum usitatissimum* L.). *Austr. J. Agric. Res.* 32: 599-607.

GUGEL, R.K., G. SÉGUIN-SWARTZ, and A., PETRIE, 1990. Pathogenicity of three isolates of *Liptosphaeria maculans* on *Brassica* species and other crucifers. *Can. J. Plant Pathol.* 12: 387-392.

HINTZE, J.L., 2000. NCSS 2000 Statistical System for Windows. Number Cruncher Statistical Systems. Kaysville, Utah.

KAUSHIK, N. and A. AGNIHOTRI. 2000. GLC analysis of rapeseed-mustard to study the variability of fatty acid composition. *Biochemical Society Transactions*, 28(6): 581-583.

KUMAR, P.R., and S. TSUNODA, 1978. Fatty acid spectrum of Mediterranean wild *Cruciferae. J. Am. Oil Chem. Soc.* 55: 320-323.

LI, X., X. FAN, L. HUN, and Q. LOU, 2002. Fatty acids of some algae from the Bohai sea. *Phytochemistry* 59: 157-161.

MALIK, R.S., 1990. Prospects of *Brassica carinata* as an oilseed crop in India. *Exp. Agric.* 26: 125-129.

MNZAVA, N.A., and K. OLSSON, 1990. Studies on tropical vegetables: Part 1: Seed amino, fatty acid and glucosinolate profile of Ethiopian mustards (*Brassica carinata* A. Braun). *Food Chemistry* 35: 229-235.

MURPHY, D.J., 1993. Designer Oil Crops: Breeding, Processing, and Biotechnology, Weihein, Germany and New York, USA.

O'BRYNE, D.J., D.A. KNAUFT, and R.B. SHIREMAN, 1997. Low fat–monounsaturated rich diets containing high-oleic peanuts improve serum lipoprotein profiles. *Lipids* 32: 687-695.

OHLROGGE, J.B., 1994. Design of new plant products: Engineering of fatty acid metabolism. *Plant Physiol.* 104: 821-826.

PLEINES, S., and W. FRIEDT, 1988. Breeding for improved C18 fatty acid composition in rapeseed (*Brassica napus* L.). *Fat Sci. Technol.* 90: 167-171.

RAHMAN, S.M., Y. TAKAGI and T. KINOSHITA, 1997. Genetic control of high stearic acid content in seed oil of two soybean mutants. *Theor. Appl. Genet.* 95: 772-776.

REBETZKE, G.J., V.R., PANTALONE, J.W. BURTON, B.F.CARVER, and R.F. WILSON, 1996. Phenotypic variation for saturated fatty acid content in soybean. *Euphytica*, 91: 289-295.

RENAUD, S., M. DE LORGERIL, J. DELAYE, J. GUIDOLLET, F. JACQUARD, N. MAMELLE, J.L. MARTIN, I. MONJAUD, P.SALEN, and P. TOUBOL, 1995. Cretan Mediterranean diet for prevention of coronary heart disease. *Am. J. Clin. Nutr.* 61: 1366-1367.

RÖBBELEN, G., 1991. Rapeseed in a changing world: Plant production potential. *In*: GCIRC (Ed). *Proceedings of the 8th International Rapeseed Congress*, Saskatoon, Canada, 9-11 July 1991, GCIRC, Saskatoon Canada. pp. 29-38.

RÖBBELEN, G., and W. THIES, 1980. Biosynthesis of seed oil and breeding improved oil quality of rapeseed. *In*: S. Tsunoda, K. Hinta, and C. Gomez-Campo (eds.), *Brassica* Crops and wild Allies: Biology and Breeding, Japan Scientific Societies Press, Tokyo. pp. 235-283.

RUIZ, M.L., DEL. CASTILLO, G. DOBSON, R. BRENNAN, and S. GORDON, 2002. Genotypic variation in fatty acid content of blackcurrant seeds. *J. Agric. Food Chem.* 50: 332-335.

SAUER, F.D. and J.K.G. KRAMER, 1983. The problems associated with the feeding of high erucic acid rapeseed oils and some fish oils to experimental animals. *In*: J.K.G. Kramer, F.D. Sauer, and W.J. Pigden (eds), High and Low Erucic Acid Rapeseed Oils. Academic Press, New York. pp. 253-292.

SLOVER, H.T. and E. LANZA, 1979. Quantitative analysis of food fatty acids by capillary gas chromatography. *J. Am. Oil Chem. Soc.*, 56:933-943.

TAHOUN, M.K., A.I. GHAZY, and A.E. ALI, 1999. Oil and erucic acid content of five *Brassica* species grown for four successive years in Egypt. $[CD-ROM]$. *In*: N. Wartten and P.A. Salisbury (ed.). *Proceedings of the 10th International Rapeseed*

Congress, September 26-29, Canberra, Austlalia. Groupe Consultatif International de Recherche sur le Colza, Paris.

VELASCO, L., D. FERNANDO, and H.C. BECKER, 1998. Variability for the fatty acid composition of the seed oil in a germplasm collection of the genus *Brassica*. *Genetic Res. and Crop Evol.* 45: 371-382.

VELASCO, L., J.M. FERNÁNDEZ-MARTINEZ, and A. DE HARO, 1997. Induced mutations for C18 unsaturated fatty acids in Ethiopian mustard. *Can. J. Plant. Sci.* 77: 91-95.

VELASCO, L., J.M. FERNÁNDEZ-MARTINEZ, and A. DE HARO, 1995. Isolation of induced mutants in Ethiopian mustard (*Brassica carinata* A. Braun) with low levels of erucic acid. *Plant Breeding* 114: 454-456.

WESTPHAL, A., and R. MARQUARD, 1980. Yield and quality of *Brassica* species in Ethiopia. *Plant Res. Dev.* 13, 114-127.

YITBAREK, S., 1992. Pathological research on noug, linseed, gomenzer and rapeseed in Ethiopia. *In*: Oilseeds research and development in Ethiopia. *Proceedings of the first National Oilseeds Workshop*, 3-5 December 1991, Addis Ababa, Ethiopia. pp. 151-161.

CHAPTER V

GENETIC ANALYSIS OF ETHIOPIAN MUSTARD GENOTYPES USING AMPLIFIED FRAGMENT LENGTH POLYMORPHISM (AFLP) MARKERS

5.1. ABSTRACT

Genetic diversity within Brassica carinata A. Braun has not been extensively examined with molecular markers. The objective of this study was to investigate the genetic relationships among 39 B. carinata genotypes using amplified fragment length polymorphisms (AFLPs). Thirty-nine genotypes of B. carinata were analyzed using six AFLP primer combinations. A total of 189 polymorphic markers were scored, with an average of 32 markers per primer combination. Genetic distance estimates (GDEs) based on AFLPs was calculated to range from 0.346 to 0.639 with a mean of 0.504 ± 0.002. Polymorphic rates ranged from 50 to 80%. The unweighted pair group method of arithmetic averages (UPGMA) cluster analysis revealed seven distinct clusters. These data demonstrate that AFLP is a reliable tool and permits greater insights into the genetic diversity of B. carinata.

Key words: AFLPs, *Brassica carinata* A. Braun, cluster analysis, DNA fingerprinting, genetic distance.

5.2. INTRODUCTION

The genus *Brassica* includes a total of 41 species (Gladis and Hammer, 1990). Six of these are economically important species, namely, *Brassica rapa* (AA), *B. oleraceae* (CC), *B. nigra* (BB), *B. juncea* (AABB), *B. napus* (AACC) and *B. carinata* (BBCC). Ethiopian or Abyssinian mustard (*Brassica carinata)* is an important oil crop in Ethiopia. It is the third most important oil crop next to niger seed (*Guizotia abyssinica* Cass.) and linseed (*Linum usitatissimum* L) (CSA, 1998). It is also eaten as a vegetable, a condiment or salad, and even used as green manure. It is mainly self-pollinating.

Brassica carinata evolved as a natural cross between *B. nigra* (BB) (n=8) and *B. oleracea* (CC) (n=9), in the highlands of the Ethiopian plateau and the adjoining portion of East Africa and the Mediterranean coast and underwent further chromosomal doubling (2n=34) (UN, 1935; Gomez-Campo and Prakash, 1999).

In Ethiopia, *B. carinata* is higher yielding, more resistant to diseases, insect pests, and seed shattering than *B. napus* with the additional agronomic advantages of better tolerance for semi-arid conditions (Knowles *et al.*, 1981; Fereres *et al.*, 1983; Malik, 1990). *Brassica carinata* can serve as an important source of genes, which are rare in other oilseed *Brassicas*.

For an efficient breeding program, information concerning the extent and nature of genetic diversity within a crop species is useful for characterizing individual accessions and cultivars in the selection of parents for hybridization (Rabbani *et al.*, 1998a). There are various techniques for studying the genetic variability of crop germplasm, including morphological traits, total seed proteins, isozymes and various types of molecular markers (Rabbani *et al.*, 1998a).

Different heterotic groups have been assigned on the basis of morphological traits and biochemical parameters such as isozyme analysis and seed storage proteins (Nucca *et al.*, 1978; Arus *et al.*, 1985). However, these characters reveal only limited polymorphisms among closely related genotypes and are influenced by prevailing environmental conditions. Therefore, a new technique that detects more polymorphisms than morphological traits and seed proteins must be explored for genetic characterization. Such information is useful for selection of diverse parents in a breeding program.

DNA marker systems have been effectively used for genetic variation analysis (Lee, 1995). DNA markers can identify many genetic loci simultaneously with an excellent coverage of an entire genome, are phenotypically neutral, and can be applied at any development stage (Jones *et al.*, 1997b). Furthermore, molecular markers are not subject

to environmental change, making them especially informative and superior to traditional methods of genotyping such as the use of morphological traits and biochemical markers (Tanksley *et al.*, 1989; Messmer et *al.*, 1993; Melchinger *et al.*, 1994).

Molecular marker techniques include restriction fragment length polymorphisms (RFLPs) (Beckman and Soller, 1983), simple sequence repeats (SSRs) or microsatellites (Tautz, 1989), random amplification of polymorphic DNA (RAPDs) (Williams *et al.*, 1990; Welsh and McClelland, 1990; Karp *et al.*, 1997) and amplified fragment length polymorphisms (AFLPs) (Vos *et al.*, 1995). These techniques differ in their principles and applications and generate different amounts of data points (Das *et al.*, 1999).

RFLPs are well suited for the construction of linkage maps because of their high specificity (Chyi *et al.*, 1992) and their codominant nature and have also been used for the analysis of genetic diversity (Song *et al.*, 1988a; 1988b). Although RFLPs have been used to study the taxonomic status of a number of *Brassicaceae* species using nuclear and chloroplast DNA (Song *et al.*, 1988a; 1988b; Pradhan *et al.*, 1992; Lanner *et al.*, 1997) this technique is labor intensive, time consuming, and expensive. Therefore, RFLPs is not a technique of choice for studying new or alternative crops where little prior data is available.

With the advent of the polymerase chain reaction (PCR) (Mullis *et al.*, 1986) resulting in the further development of RAPDs, SSRs and AFLPs, most of the problems associated with RFLPs were overcome. The RAPD technique gained importance due to its simplicity and efficiency, requiring no prior sequence knowledge (Karp *et al.*, 1997). However, the RAPD technique has been shown to be not reproducible especially between laboratories as it is highly influenced by experimental conditions (Jones *et al.*, 1997a; Karp *et al.*, 1997; Virk *et al.*, 2000; Devos and Gale, 1992; Staub *et al.*, 1996). In *Brassica*, the RAPD assay has been employed to develop genome specific markers (Quiros *et al.*, 1991) and to resolve taxonomic relationships (Demeke *et al.*, 1992; Rabbani *et al.*, 1998b). It has also been used for cultivar identification (Hu and Quiros,

1991; Kresovich *et al.*, 1992; Das *et al.*, 1999). However, finding stable RAPD polymorphic markers in rapeseed has proven difficult (Mailer *et al.*, 1994).

Simple sequence repeats (SSRs), which are based on microsatellite sequences, have been shown to detect very high levels of polymorphism (Karp *et al.*, 1997). However, prior information about the genome is necessary before SSR markers can be exploited to their fullest potential.

AFLPs combine the advantage of time efficiency of PCR-based markers with the reliability of RFLP markers (Vos *et al.* 1995). The AFLP assay requires no prior sequence knowledge and detects at least 10 times more genetic loci than RFLPs, RAPDs and SSRs in most crops (Tohme *et al.*, 1996; Maughan *et al.*, 1996; Hill *et al.*, 1996). Therefore, the AFLP assay has the ability to detect thousands of independent genetic loci in a short time.

The aim of this study was to determine the feasibility of using AFLP markers to evaluate the genetic relationships between Ethiopian mustard cultivars and advanced breeding lines in order to maximize selection of diverse parents in a breeding program.

5.3. MATERIALS AND METHODS

Plant material

A total of thirty-nine *B. carinata* cultivars and accessions were used in this study (Table 5.1). The material consisted of two nationally released cultivars (Yellow Dodolla and S-67), 18 advanced breeding lines, which were obtained from the Ethiopian national *B. carinata* breeding program and 20 accessions collected from different parts of the country by the Biodiversity Conservation and Research Institute of Ethiopia. Plants were grown from seeds in pots in the greenhouse with 20°C night temperature and 26°C day temperature, during August 2000, at the university of the Free State, Bloemfontein, South Africa.

Table 5.1. List of Ethiopian mustard genotypes used for genetic analysis.

No	Genotypes	No	Genotypes	No	Genotypes	No	Genotypes
1	Yellow Dodolla-03	11	Merawi	21	PGRC/E 21156	31	Yellow Dodolla
2	S-67-02	12	PGRC/E 20059	22	PGRC/E 20112/2	32	S-67
3	PGRC/E 21261-03	13	PGRC/E 21207	23	PGRC/E 200413	33	Adet
4	PGRC/E 203221-03	14	PGRC/E 20080	24	(4DxZem-1)x(Zem-1-AD/88)	34	PGRC/E 21358/2
5	C94-S-67	15	PGRC/E 207929	25	(4DxZem-1)x(Zem-1-F5/10)	35	PGRC/E 207928
6	C94-Dodolla	16	PGRC/E 207975	26	PGRC/E 21162/1	36	PGRC/E 21320/5
7	Yellow Dodolla-01	17	PGRC/E 20168/1	27	PGRC/E 20163/1	37	PGRC/E 20113
8	PGRC/E 203221-01	18	PGRC/E 20095/1	28	PGRC/E 20076/2	38	PGRC/E 20165
9	PGRC/E 21261-02	19	PGRC/E 21356/1	29	PGRC/E 20147/1	39	PGRC/E 21051
10	PGRC/E 20130	20	PGRC/E 21237	30	PGRC/E 208404		

DNA extraction

DNA was extracted from fresh young leaves using a modified monocot extraction protocol (Edwards *et al.*, 1991). Young plant leaf tissue was collected on ice, from one plant of each cultivar and ground into a fine powder in liquid nitrogen. The macerated tissue was incubated in 10 ml of extraction buffer (0.5 M NaCl, 0.1 M Tris-HCl [pH 8], 0.25 M EDTA and 20% SDS) at 65°C for 30 minutes, and with shaking every 10 minutes. One ml of 1% CTAB buffer (w/v) (1 M Tris-CHl, 0.25 M EDTA, and 1% CTAB) and 2 ml 5 M NaCl were added to the homogenate and incubated at 65°C for one hour with mixing every 10 minutes. Chloroform-isoamyl alcohol (24:1 v/v) (10 ml) was added in a 1:1 ratio and mixed gently. Cell debris and proteins were removed by centrifugation for 15 minutes at 10 000 rpm. After centrifugation, the supernatant was retained and the DNA was precipitated by adding two volumes (v/v) of cold absolute ethanol (1:2) (%) followed by overnight incubation at 4°C. The precipitated DNA was spooled with a sterile pasteur pipette and washed twice in 1 ml 70% ethanol (EtOH) to remove residual salt. The DNA pellet was resuspended in 1 ml of sterile PCR grade water and stored at -20°C.

DNA concentration and purity was determined spectrophotometerically at 260 and 280 nm (Hitachi U-2000). The DNA concentration was calculated using the formula,

[DNA]=Optical density x dilution factor x constant (50 μg/ml). Purity was determined using the 260/280 nm ratio. DNA samples were diluted to a working concentration of 250 ng $\mu l L^{-1}$ and stored at 4°C. The integrity and concentration of the DNA was confirmed by 1% agarose electrophoresis for 45 minutes at 80 volts in 0.5X TAE (0.438 g/l Tris, 0.09 ml/L acetic acid, and 0.022 g/L acid EDTA) with visualization under UV light after staining with ethidium bromide.

AFLP procedure

The AFLP reactions were done according to the manufacturers instructions (Gibco BRL, 1996) using commercial adaptor and primer sequences (Table 5.2).

Restriction endonuclease digestion and ligation of adaptors

Genomic DNA (250 ng) was mixed with 5x reaction buffer (50 mM Tris-HCl [pH 7.5], 50 mM Mg-acetate, and 250 mM K-acetae) and 2 μl of the restriction enzyme *EcoR*1 and *Mse*1 and incubated for 2 h at 37°C. The digested DNA fragments were then ligated to *EcoR*1 and *Mse*1 adaptors (Table 5.2) with 1 μl T4 DNA ligase (1 unit/μl in 10 mM Tris-HCl [pH 7.5] for 2 h at 20°C. The ligated DNA template was diluted 10 times with TE buffer (10 mM Tris-HCl [pH 8.0], 0.1 mM EDTA) and stored at -20°C.

Polymerase chain reaction

Pre-selective PCR reaction were performed in 51 μl reactions with 5 μl (1:10) diluted ligation product, 40 μl pre-amp primer mix, 10x PCR buffer (100 mM Tris-HCl [pH 8.3], 15 mM $MgCl_2$, and 500 mM KCl) and Taq DNA polymerase (1 unit/μl) (Gibco BRL). A touchdown Hybaid thermal cycler was used to perform the reaction for 20 cycles with the following profile: 30 s at 94°C, 60 s at 56°C and 60 s at 72°C. Pre-selective PCR products were diluted 1:50 in TE buffer (10 mM Tris-HCl [pH 8.0], and 0.1 mM EDTA) and stored at 4°C.

Selective PCR-reactions were performed in 20 μl reactions containing 5 μl of diluted (1:50) pre-selective template DNA, 4.5 μl of Mse+ primer (6.7 ng/μl, and dNTPs) (Table 5.2), 1 μl Eco+ primer (27.8 ng/μl), 2μl of 10x PCR buffer and 5 U of Ampli Taq

DNA polymerase (unit/μl). Reactions were performed for 30 cycles with the following cycle profile: 30 s at 94°C, 30 s at 65°C and 60 s at 72°C. The 65°C annealing temperature was reduced by 0.7°C per cycle for 12 consecutive cycles and then maintained at 56°C for the remaining 18 cycles. Six primer combinations were tested. *EcoR*1 primers (PE Biosystems) were labeled with NED and FAM, respectively.

Table 5.2. Adaptors and primers used for AFLP preamplification and selective amplification.

Name of Adapters/Primers	Sequence (5'-3')
(1) Adaptor:	
EcoR1 adaptor	5'-CTCGTAGACTGCGTACC-3'
	3'-CATCTGACGCATGGTTAA-5'
Mse1 adaptor	5'-GACGATGAGTCCTGAG-3'
	3'-TACTCAGGACTCAT-5'
(2) Primers:	
Mse1 primer:	GATGAGTCCTGAGTAA
Mse1 +CTT	GATGAGTCCTGAGTAACTT
Mse1 +CAC	GATGAGTCCTGAGTAACAC
Mse1 +CTC	GATGAGTCCTGAGTAACTC
EcoR1 primer	GACTGCGTACCAATTC
EcoR1 + AAC (Ned)	GACTGCGTACCAATTCAAC
EcoR1 + ACT (Fam)	GACTGCGTACCAATTCACT
EcoR1 +ACA (Fam)	GACTGCGTACCAATTCACA

After amplification, 5µl of each selective reaction was added to 24 µl Formamide, and 1 µl Rox standard size marker, denatured at 94°C for 5 minutes, cooled on ice and resolved on a Perkin Elmer ABI Prism 310 Automated capillary sequencer (PE Biosystems).

Statistical analysis

Estimates of genetic distance

Genotypes were scored for presence and absence (1 for presence and 0 for absence) of AFLP bands and the data entered into a binary data matrix as discrete variables. Bands present in all accessions were not scored. Fragments smaller than 80 bp were excluded from the data matrix. Distance matrices for all pairs of genotypes were constructed from the AFLP data matrix using the Euclidean distance method (Kaufman and Rousseeuw, 1990). The Euclidean distances were calculated as follows:

$$GD = \sqrt{\sum[(X_i - Y_i)^2 / N]},$$

Where, GD is the genetic distance between individual X and individual Y; $i = 1$ to N; N is the total number of bands; X_i and Y_i are the i^{th} band scores (1 or 0) for individuals X and Y.

Cluster analysis

Cluster analysis was performed using the genetic distance matrices generated by the Euclidean distance method to reveal the patterns of genetic relationships among genotypes. The unweighted pair group method of arithmetic averages (UPGMA) (Hintze, 2001), which minimizes within-cluster variance, was used. The results of cluster analysis were presented in the form of dendrograms to infer relationships among genotypes. The cophenetic correlation (Kaufman and Rousseeuw, 1990; Hintze, 2001) for each dendrogram was also computed as a measure of 'goodness of fit' for each dendrogram. The two delta goodness of fit statistics, delta (0.5) and delta (1.0) were calculated to determine which clustering configuration fits the data better (Mather, 1976).

5.4. RESULTS AND DISCUSSION

Genetic diversity among *B. carinata* genotypes was assayed with AFLP techniques using six AFLP primer combinations. The number of scorable fragments amplified for each AFLP primer pair varied from 28 to 80 with an average of 46 across all the genotypes (Table 3). The number of polymorphic fragments for each primer pair varied from 14 to 64 with an average of 32 polymorphic fragments. A total of 189 AFLP fragments were polymorphic across all the genotypes for the six primer pairs assayed (Table 3). The 189 polymorphic fragments accounted for 68% of the total amplified fragments.

The fragments generated range in size from 56 to 497 bp. The number of fragments decreased as the size increased. Fragments longer than 400 bp were rarely detected. The number of fragments varied for different primer combinations (Table 3). The primer combination of M-CAC/E-ACA gave the smallest number of fragments (28) while all other primer combinations detected ≥ 30 fragments. The most polymorphic primer combination was, M-CTT/E-ACT, which amplified 80 fragments to each (Table 3).

Table 5.3. Analysis of the level of polymorphism with AFLP primer combinations among 39 *B. carinata* genotypes.

Primer	Total number of fragments (a)	Polymorphic fragments	
		Number (b)	Polymorphism (%) (=b/a x (100))
M-CTT/E-ACC	39	28	72
M-CTT/E-ACT	80	64	80
M-CAC/E-AAC	47	30	64
M-CAC/E-ACA	28	14	50
M-CTC/E-ACC	30	18	60
M-CTC/E-ACT	54	35	65
Total	278	189	68
Average	46	32	

Genetic distances estimates (GDEs)

Genetic distance was determined for all 741 pair wise comparisons of genotypes (Table 4). The range of dissimilarities ranged between 0.346 PGRC/E 20165 and PGRC/E 21051 as well as between PGRC/E 20076/1 and PGRC/E 20147/1 to 0.639 between PGRC/E 21261-03 and PGRC/E 20163/1. The mean genetic distance across all the samples was 0.504±0.002. The frequency of distribution of (Figure 5.2) indicated a normal distribution.

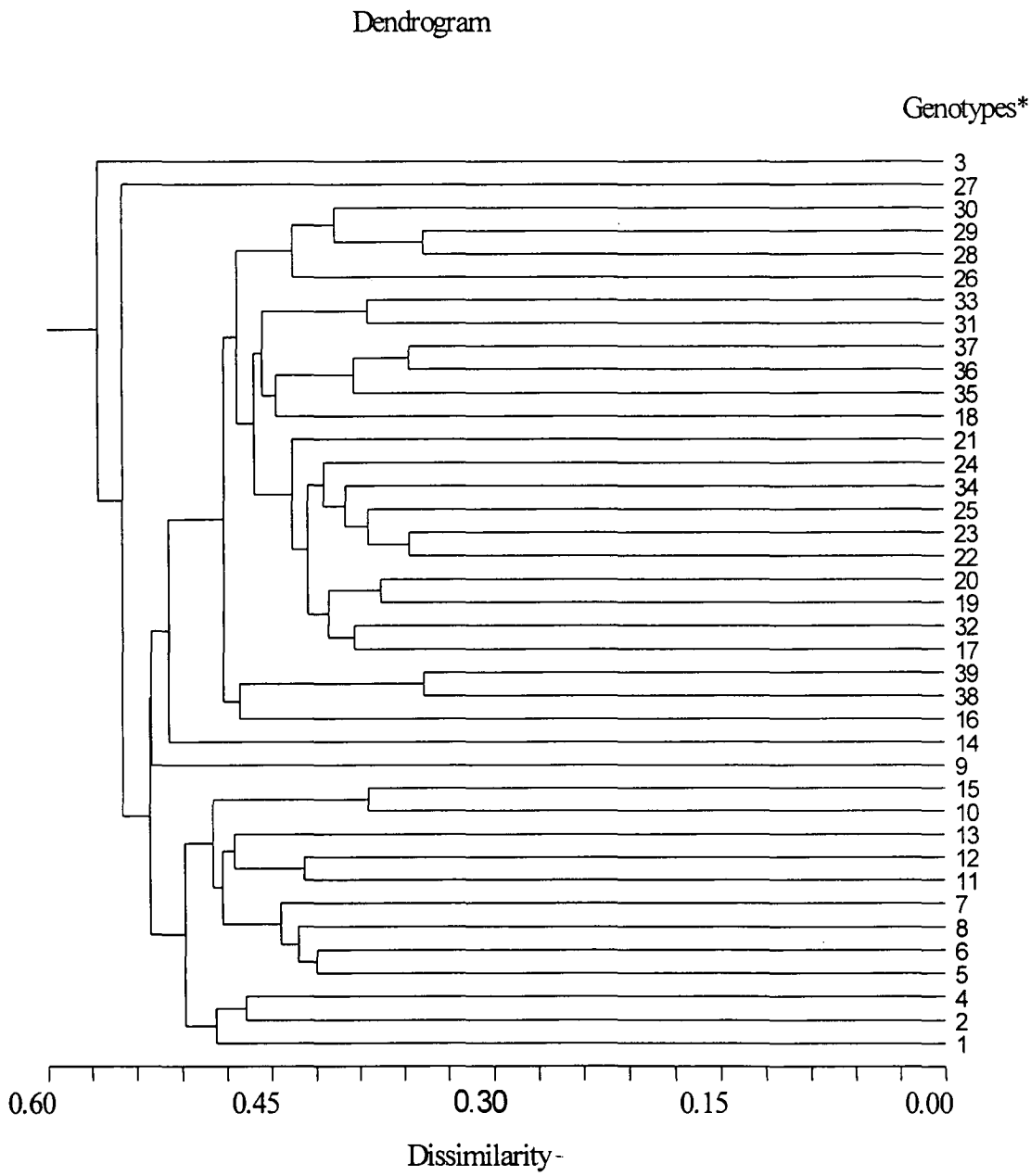
Cluster analysis

The dendrogram from UPGMA clustering method of 39 *B. carinata* genotypes based on Euclidean genetic distances generated from AFLP data matrix is shown in Figure 5.1. At a cut off 0.5 the dendrogram revealed seven major clusters consisting of three clusters containing more than one genotype and four singletons.

The correlation between the cophenetic values was very high (0.82) indicating a very good fit of the cluster analysis performed. The two delta goodness of fit statistics, delta (0.5) and delta (1.0) were (0.04) and (0.06), respectively (Mather, 1976).

Cluster analysis separated the 39 genotypes into seven major clusters. Cluster one consisted of 23 genotypes with a genetic distance value ranging from 0.437 between PGRC/E 20112/2 and PGRC/E 200413 to 0.664 between PGRC/E 21156 and PGRC/E 20165 with a mean value of 0.557±0.002. This cluster further divided into two distinct sub clusters. The first sub cluster consisted of genotypes PGRC/E 207975, PGRC/E 20165, and PGRC/E 21051. The second sub cluster consisted of 20 other genotypes. Cluster two consisted of three genotypes, Yellow Dodolla, S-67-02, and PGRC/E 203221-03. Cluster three had three sub clusters. The first sub cluster had C94-S-67, C94-Dodolla, PGRC/E 203221-01 and Yellow Dodolla-01.

The four genotypes, PGRC/E 21261-03, PGRC/E 21261-02, PGRC/E 20080 and PGRC/E 20163/1 clustered independently (Figure 5.1). Divergent genotypes may have good breeding values. Genotypes in the same cluster may represent members of one heterotic group. Maximum variability for selection in segregating populations may be achieved by utilizing genotypes from different clusters as parents for crosses.



*1=Yellow Dodolla-03, 2=S-67-02, 3=PGRC/E 21261-03, 4=PGRC/E 203221-03, 5=C94-S-67, 6=C94-Dodolla, 7=Yellow Dodolla-01, 8=PGRC/E 203221-01, 9=PGRC/E 21261-02, 10=PGRC/E 20130, 11=Merawi, 12=PGRC/E 20059, 13=PGRC/E 21207, 14=PGRC/E 20080, 15=PGRC/E 207929, 16=PGRC/E 207975, 17=PGRC/E 20168/1, 18=PGRC/E 20095/1, 19=PGRC/E 21356/1, 20=PGRC/E 21237, 21=PGRC/E 21156, 22=PGRC/E 20112/2, 23=PGRC/E 200413, 24=(4DxZem-1) x (Zem-1-AD/88), 25=(4DxZem-1) x (Zem-1-F5/10), 26=PGRC/E 21162/1, 27=PGRC/E 20163/1, 28=PGRC/E 20076/2, 29=PGRC/E 20147/1, 30=PGRC/E 208404, 31=Yellow Dodolla, 32=S-67, 33=Adet, 34=PGRC/E 21358/2, 35=PGRC/E 207928, 36=PGRC/E 21320/5, 37=PGRC/E 20113, 38=PGRC/E 20165, 39=PGRC/E 21051

Figure 5.1. Dendrogram generated based on UPGMA clustering method depicting genetic relationships among 39 *B. carinata* genotypes based on AFLP data.

Table 5.4. Estimates of genetic distances based on AFLP data for pair wise combinations of 39 *B. carinata* genotypes.

No*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38
1	0																																					
2	0.498	0																																				
3	0.545	0.524	0																																			
4	0.476	0.467	0.557	0																																		
5	0.552	0.531	0.575	0.467	0																																	
6	0.476	0.436	0.557	0.46	0.419	0																																
7	0.539	0.489	0.587	0.537	0.46	0.436	0																															
8	0.519	0.467	0.569	0.517	0.436	0.428	0.436	0																														
9	0.595	0.563	0.593	0.581	0.537	0.569	0.562	0.53	0																													
10	0.526	0.489	0.575	0.482	0.46	0.452	0.489	0.451	0.537	0																												
11	0.505	0.467	0.543	0.474	0.496	0.444	0.451	0.46	0.503	0.452	0																											
12	0.559	0.524	0.544	0.531	0.467	0.504	0.496	0.489	0.517	0.496	0.428	0																										
13	0.565	0.524	0.593	0.489	0.467	0.503	0.51	0.503	0.53	0.482	0.474	0.474	0																									
14	0.539	0.524	0.581	0.517	0.524	0.557	0.537	0.489	0.556	0.524	0.53	0.517	0.474	0																								
15	0.533	0.537	0.605	0.517	0.524	0.504	0.524	0.517	0.517	0.384	0.474	0.503	0.517	0.543	0																							
16	0.546	0.562	0.543	0.556	0.496	0.556	0.562	0.531	0.543	0.537	0.569	0.517	0.569	0.503	0.53	0																						
17	0.512	0.531	0.537	0.524	0.46	0.51	0.543	0.467	0.51	0.503	0.524	0.482	0.537	0.51	0.537	0.436	0																					
18	0.583	0.575	0.557	0.543	0.496	0.556	0.537	0.53	0.517	0.524	0.517	0.459	0.569	0.543	0.517	0.503	0.467	0																				
19	0.519	0.482	0.543	0.517	0.452	0.475	0.482	0.459	0.503	0.436	0.489	0.489	0.517	0.489	0.53	0.427	0.419	0.459	0																			
20	0.519	0.537	0.543	0.517	0.451	0.503	0.51	0.459	0.489	0.496	0.503	0.474	0.517	0.489	0.517	0.444	0.402	0.459	0.375	0																		
21	0.539	0.504	0.575	0.537	0.543	0.537	0.557	0.496	0.562	0.503	0.537	0.51	0.537	0.467	0.55	0.524	0.489	0.482	0.402	0.451	0																	
22	0.539	0.544	0.587	0.537	0.489	0.55	0.557	0.482	0.524	0.503	0.537	0.524	0.55	0.496	0.55	0.482	0.411	0.496	0.436	0.436	0.444	0																
23	0.498	0.531	0.575	0.524	0.459	0.51	0.489	0.436	0.524	0.459	0.496	0.496	0.496	0.482	0.51	0.482	0.427	0.467	0.384	0.419	0.411	0.356	0															
24	0.491	0.55	0.593	0.531	0.496	0.543	0.562	0.489	0.53	0.496	0.543	0.517	0.543	0.517	0.543	0.489	0.419	0.489	0.427	0.444	0.436	0.384	0.384	0														
25	0.546	0.537	0.557	0.531	0.51	0.544	0.524	0.46	0.53	0.51	0.53	0.489	0.543	0.489	0.556	0.489	0.436	0.459	0.427	0.427	0.419	0.402	0.365	0.444	0													
26	0.512	0.544	0.563	0.563	0.503	0.524	0.544	0.496	0.524	0.517	0.524	0.467	0.563	0.537	0.537	0.482	0.411	0.467	0.451	0.436	0.489	0.459	0.427	0.451	0.451	0												
27	0.571	0.599	0.639	0.569	0.537	0.581	0.575	0.569	0.581	0.55	0.569	0.543	0.593	0.581	0.569	0.556	0.496	0.556	0.543	0.53	0.562	0.562	0.537	0.543	0.556	0.496	0											
28	0.533	0.537	0.605	0.569	0.51	0.544	0.537	0.517	0.556	0.51	0.556	0.53	0.581	0.543	0.543	0.53	0.467	0.503	0.459	0.489	0.467	0.482	0.436	0.459	0.489	0.419	0.411	0										
29	0.552	0.544	0.587	0.575	0.489	0.524	0.544	0.51	0.537	0.517	0.55	0.524	0.575	0.537	0.55	0.482	0.459	0.51	0.451	0.467	0.503	0.459	0.444	0.467	0.467	0.411	0.467	0.346	0									
30	0.552	0.531	0.599	0.562	0.517	0.524	0.544	0.524	0.524	0.503	0.575	0.562	0.575	0.524	0.524	0.451	0.474	0.537	0.419	0.467	0.489	0.489	0.444	0.496	0.451	0.474	0.524	0.419	0.393	0								
31	0.583	0.575	0.544	0.544	0.524	0.557	0.575	0.517	0.53	0.524	0.543	0.53	0.53	0.53	0.556	0.489	0.467	0.489	0.474	0.459	0.496	0.451	0.467	0.459	0.459	0.482	0.593	0.543	0.524	0.524	0							
32	0.552	0.544	0.562	0.524	0.475	0.524	0.531	0.496	0.524	0.503	0.537	0.482	0.55	0.524	0.524	0.436	0.393	0.451	0.402	0.419	0.444	0.411	0.427	0.419	0.419	0.444	0.524	0.451	0.427	0.427	0.436	0						
33	0.577	0.543	0.51	0.537	0.489	0.55	0.569	0.51	0.524	0.517	0.524	0.524	0.537	0.537	0.55	0.467	0.459	0.467	0.467	0.482	0.503	0.444	0.459	0.467	0.451	0.474	0.562	0.51	0.489	0.503	0.384	0.444	0					
34	0.533	0.51	0.544	0.503	0.482	0.531	0.524	0.444	0.503	0.496	0.503	0.474	0.503	0.474	0.517	0.503	0.436	0.444	0.444	0.444	0.419	0.402	0.402	0.444	0.393	0.451	0.543	0.489	0.482	0.496	0.411	0.402	0.402	0				
35	0.577	0.581	0.575	0.563	0.504	0.563	0.557	0.524	0.562	0.53	0.575	0.55	0.587	0.537	0.575	0.467	0.489	0.451	0.436	0.482	0.459	0.474	0.444	0.467	0.419	0.474	0.562	0.482	0.459	0.459	0.451	0.411	0.427	0.436	0			
36	0.577	0.544	0.537	0.587	0.557	0.573	0.544	0.51	0.537	0.569	0.537	0.51	0.575	0.524	0.587	0.51	0.517	0.467	0.467	0.482	0.459	0.489	0.444	0.51	0.402	0.474	0.575	0.51	0.474	0.503	0.467	0.474	0.474	0.436	0.393	0		
37	0.552	0.531	0.537	0.55	0.531	0.55	0.557	0.524	0.537	0.543	0.51	0.496	0.562	0.51	0.55	0.496	0.503	0.419	0.451	0.451	0.444	0.474	0.444	0.51	0.419	0.427	0.55	0.482	0.474	0.489	0.451	0.489	0.411	0.419	0.393	0.356	0	
38	0.571	0.587	0.543	0.569	0.524	0.569	0.587	0.543	0.53	0.55	0.581	0.556	0.605	0.581	0.581	0.459	0.482	0.489	0.459	0.459	0.524	0.51	0.496	0.53	0.459	0.496	0.543	0.503	0.467	0.467	0.474	0.496	0.436	0.474	0.451	0.467	0.419	0
39	0.552	0.581	0.575	0.563	0.503	0.55	0.557	0.524	0.537	0.53	0.562	0.575	0.575	0.575	0.562	0.482	0.474	0.55	0.451	0.451	0.503	0.489	0.459	0.51	0.467	0.503	0.524	0.467	0.444	0.444	0.51	0.489	0.503	0.496	0.503	0.517	0.489	0.346

*1=Yellow Dodolla-03, 2=S-67-02, 3=PGRC/E 21261-03, 4=PGRC/E 203221-03, 5=C94-S-67, 6=C94-Dodolla, 7=Yellow Dodolla-01, 8=PGRC/E 203221-01, 9=PGRC/E 21261-02, 10=PGRC/E 20130, 11=Merawi, 12=PGRC/E 20059, 13=PGRC/E 21207, 14=PGRC/E 20080, 15=PGRC/E 207929, 16=PGRC/E 207975, 17=PGRC/E 20168/1, 18=PGRC/E 20095/1, 19=PGRC/E 21356/1, 20=PGRC/E 21237, 21=PGRC/E 21156, 22=PGRC/E 20112/2, 23=PGRC/E 200413, 24=(4DxZem-1) x (Zem-1-AD/88), 25=(4DxZem-1) x (Zem-1-F5/10), 26=PGRC/E 21162/1, 27=PGRC/E 20163/1, 28=PGRC/E 20076/2, 29=PGRC/E 20147/1, 30=PGRC/E 208404, 31=Yellow Dodolla, 32=S-67, 33=Adet, 34=PGRC/E 21358/2, 35=PGRC/E 207928, 36=PGRC/E 21320/5, 37=PGRC/E 20113, 38=PGRC/E 20165, 39=PGRC/E 21051

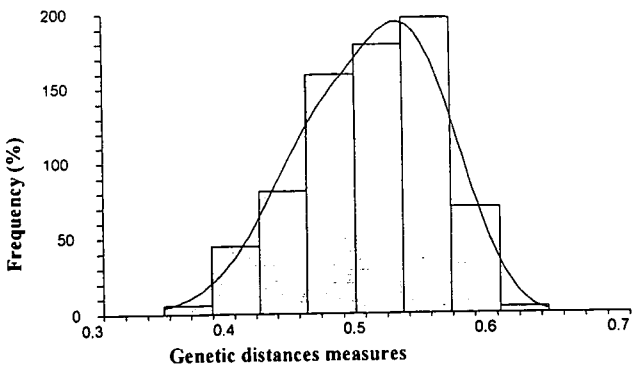


Figure 5.2. Frequency distribution of pair wise AFLP based genetic distance estimates (GDEs) among 39 *B. carinata* genotypes. GDEs were calculated for all combinations (n=741).

Knowledge of genetic relationships among genotypes is useful in plant breeding programs because it permits the organization of germplasm and provides information for more efficient parental selection. The breeder can use genetic distance information to make informed decisions regarding the choice of genotypes to cross for the development of populations, or to facilitate the identification of diverse parents to cross in hybrid combinations in order to maximize the expression of heterosis (Smith *et al.*, 1990). Furthermore, fingerprinting genotypes offers the opportunity for the removal of any duplicate germplasms.

In *B. carinata*, information on genetic diversity and/or genetic relationships among genotypes is currently limited. Therefore, the AFLP DNA pattern, or fingerprint should be a valuable tool for breeders.

Lombard *et al.* (2000) in their study of the genetic relationships and fingerprinting of 83 rapeseed (*B. napus* L.) cultivars by AFLP showed that AFLP markers functioned well in the assessment of genetic relatedness between rapeseed cultivars in the context of plant registration and protection. A total of 324 polymorphic fragments were generated with 17

primer combinations. The number of markers per primer combination ranged from 12 to 30, with an average of 19.1. The most polymorphic primer combinations were M-CAA+E-AAC, M-CTT+E-AAC, and M-CTT+E-AAG, which produced 30 markers each. The fragment sizes ranged from 100 to 750 bp.

Das *et al.* (1999) assessed the genetic variation of *B. campestries* cultivars with AFLP and RAPD markers. RAPDs generated a total of 125 bands using 13 decamer primers (an average of 9.6 bands per assay) of which nearly 80% were polymorphic. The percent polymorphism ranged from 60-100%. AFLPs, on the other hand generated a total of 319 markers, an average of 64 bands per assay. Of these, 213 were polymorphic in nature (66.8%). AFLP methodology detected polymorphism more efficiently than the RAPD approach due to greater numbers of loci assayed per reaction. However, the level of polymorphism was higher with RAPD (80 %) than with AFLP (66.8 %). Although AFLPs do not reveal a high percentage of polymorphism but are more efficient as they produce more number of bands per reaction as compared to RAPD. Thus, AFLP has a higher marker index, an overall measure of marker efficiency (Powell *et al.*, 1996; Nakajima *et al* 1998). AFLP has also been reported to be highly reproducible with low error rates, which provides a definite advantage over RAPD (Jones *et al.*, 1997a).

In conclusion, AFLP markers exhibited a high level of efficiency for detecting DNA polymorphisms among *B. carinata* genotypes. Knowledge of diversity patterns and specific GDEs may increase the efficiency of *B. carinata* genetic improvement in Ethiopia among adapted parents used for cultivar development, and providing predictive measure of such important parameters as population genetic variance and heterosis.

5.5. REFERENCES

ARUS, P., C.R. SHIELDS and J.J. ORTON. 1985. Applications of isozyme electrophoresis for purity testing and cultivar identification of F1 hybrids in *Brassica oleraceae*. *Euphytica* 34: 651-657.

BECKMAN, J.S. and M. SOLLER. 1983. Restriction fragment length polymorphism in plant genetic improvement: methodologies, mapping and costs. *Theor. Appl. Genet.* 67: 35-43.

CENTRAL STATISTICAL AUTHORITY OF ETHIOPIA (SCA). 1998. Agricultural sample survey for 1997/98: Area and production of belg season crops for private peasant holdings. Volume 5 Addis Ababa, Ethiopia.

CHYI, S., M.E. HOENCKE and J.L. SERNYK. 1992. A genetic linkage map of restriction fragment length polymorphism loci for *Brassica rapa* (syn. *campestries*). *Genome* 35: 746-757.

DAS, S., J. RAJAGOPAL, S. BHATIA, P.S. SRIVASTAVA and M. LAKSHMIKUMARAN. 1999. Assessment of genetic variation within *Brassica campestris* cultivars using amplified fragment length polymorphism and random amplification of polymorphic DNA markers. *J. Biosciences* 24(4): 433-440.

DEMEKE, T., R.P. ADAMS and R. CHIBBAR. 1992. Potential taxonomic use of random amplified polymorphic DNA (RAPD): a case study of *Brassica*. *Theor. Appl. Genet.* 84: 990-994.

DEVOS, K.M. and M.D. GALE. 1992. The use of random amplified polymorphic DNA in wheat. *Theor. Appl. Genet.* 84: 567-572.

EDWARDS, K., C. JOHNSTONE and C. THOMPSON. 1991. A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acid Res.* 19: 1349.

FERERES, E., J. FERNÁNDEZ-MARTÍNEZ, I. MINGUEZ and J. DOMÍNGUEZ. 1983. Productivity of *B. juncea* and *B. carinata* in relation to rapeseed. *In:* Int. Rapeseed Conf., 6th, 17-19 May 1983. GCIRC, Paris, France. pp. 293-298.

GIBCO BRL. 1996. AFLP analysis system I. Technical Bulletin 34.

GOMEZ-CAMPO, C. and S. PRAKASH. 1999. Origin and domestication of the *Brassica*. *In*: GOMEZ-CAMPO, C. (ed.). Biology of Brassica Coenospecies. Elsevier, Amsterdam, pp 33-58.

GLADIS, T., and K.H. HAMMER, 1990. Die Gaterslebener Brassica-kollektion eine Einführung, *Kulturpflanze* 38:121-156.

HILL, M., H. WITSENBOER, M. ZABEAU, P. VOS, R. KESSLI and R. MICHELMORE. 1996. PCR-based fingerprinting using AFLPs as a tool for studying genetic relationships in Lactuca. *Theor. Appl. Genet.* 93: 1202-1210.

HINTZE, J.L., 2001. NCSS 2001 Statistical System for Windows. Number Cruncher Statistical Systems. Kaysville, Utah.

HU, J. and C.F. QUIROS. 1991. Identification of broccoli and cauliflower cultivars with RAPD markers. *Plant Cell Rep.* 10: 505-511.

JONES, C.J., K.J. EDWARDS, S. CASTAGLIONE, M.O. WINFIELD, F. SALA, C. VAN DE WIEL, G. BREDEMEIJER, B. VOSMAN, M. MATTHES, A. DALY, R. BRETTSCHNEIDER, P. BRTTINI, M. BUIATTI, E. MAESTRI, A. MALCEVSCHI, N.

MARMIROLI, R. AERT, G. VOLCKAERT, J. RUEDA, R. LINACERO, A. VAZQUEZ and A. KARP. 1997a. Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. *Mol. Breeding* 3: 381-390.

JONES, N., H. OUGHAM and H. THOMAS. 1997b. Markers and mapping: we are all geneticists now. *New Phytol.* 137: 165-177.

KARP, A., S. KRESOVICH, K.V. BHAT, W.G. AYAD and T. HODGKIN. 1997. Molecular tools in plant genetic resources conservation: a guide to the technologies; *In: IPGRI Technical Bulletin* No. 2. International Plant genetic Resources Institute, Rome Italy.

KAUFMAN, L. and P.J. ROUSSEEUW. 1990. Finding groups in data: An introduction to cluster analysis. A Wiley-Interscience Publication, New York.

KNOWLES, P.F., T.E. KEARNEY and D.B. COHEN. 1981. Species of rapeseed and mustard as oil crops in California. *In*: Pryde, E.H. (ed). New sources of fats and oils. Am. Oil Chem. Soc. Champain, IL, pp. 255-268.

KRESOVICH, S., J.G.K. WILLIAMS, J.R. MCFERSON, E.J. ROUTMAN and B.A. SCHAAL. 1992. Characterization of genetic identities and relationships of *Brassica oleraceae* L. via a random polymorphic DNA assay. *Theor. Appl. Gent.* 85: 190-196.

LANNER, C., T. BRYNGELSSON and M. GUSTAFSON. 1997. Relationship of wild Brassica species with chromosome number 2n=18, based on RFLP studies. *Genome* 40: 302-308.

LEE, M. 1995. DNA markers and plant breeding programs. *Adv. Agron.* 55: 265-344.

LOMBARD, V., C.P. BARIL, P. DUBREUIL, F. BLOUET and D. ZHANG. 2000. Genetic relationships and fingerprinting of rapeseed cultivars by AFLP: consequences for varietal registration. *Crop Sci.* 40: 1417-1425.

MAILER, R.J., R. SCARTH and B. FRISTENSKY. 1994. Discrimination among cultivars of rapeseed (*Brassica napus* L.) using DNA polymorphism amplified from arbitrary primers. *Theor. Appl. Genet.* 87: 697-704.

MALIK, R.S. 1990. Prospects of *Brassica carinata* as an oilseed crop in India. *Exp. Agric.* 26: 125-129.

MATHER, P. 1976. Computational methods of multivariate analysis in physical geography. Jhon Wiley & Sons.

MAUGHAN, P.J., M.A. SAGHAI MAROOF, G.R. BUSS and G.M. HUESTIS. 1996. Amplified fragment length polymorphism (AFLP) in soybean: species diversity, inheritance, and near-isogenic line analysis. *Theor. Appl. Genet.* 93: 392-401.

MESSMER, M.M., A.E. MELCHINGER, R.G. HERRMANN and J. BOPPENMAIER 1993. Relationships among early European maize inbreeds: II. Comparison of pedigree and RFLP data. *Crop Sci.* 33: 944-950.

MELCHINGER, A.E., A. GRANER, M. SINGH and M.M. MESSMER. 1994. Relationships among European barley germplasm: I. Genetic diversity among winter and spring cultivars revealed by RFLPs. *Crop Sci.* 34: 1191-1199.

MULLIS, K., F. FALOONA, S. SCHARE, R. SAIKI, G. HORN and H. ERLICH 1986. Enzymatic amplification of DNA in vitro: the polymerase enzyme chain reaction. *Cold Spring Harbor Symp. Quant. Biol.*, 51: 263-273.

NAKAJIMA, Y., K. OEDA and T. YAMAMOTO. 1998. Characterization of genetic diversity of nuclear and mitochondrial genomes in *Daucus* varieties by RAPD and AFLP. *Plant Cell Rep.* 17: 848-853.

NUCCA, R., C. SOARE, M. MOTTO and T. SALAMINI. 1978. Taxonomic significance of the zein isoelectric focusing pattern. *Maydica* 23: 239-249.

POWELL, W., M. MORGANTE, C. ANDRE, M. HANAFEY, J. VOGEL, S. TINGEY and A. RAFALASKI. 1996. The comparison of RFLP, RAPD, AFLP, and SSR (microsatellite) markers for germplasm analysis. *Mol. Breeding* 3: 225-238.

PRADHAN, A., S. PRAKASH, A. MUKHOPADHYAY and D. PENTAL. 1992. Phylogeny of *Brassica* and allied genera based on variation in chloroplast and mitochondrial DNA patterns: molecular and taxonomic classifications are incongruous. *Theor. Appl. Genet.* 85: 331-340.

QUIROS, C.F., J. HU, P. THIS, A.M. CHEVRE and M. DELSENY. 1991. Development and chromosomal localization of genome-specific markers by the polymerase chain reaction. *Theor. Appl. Genet.* 82: 627-632.

RABBANI, M.A., A. IWABUCHI, Y. MURAKAMI, T. SUZUKI and K. TAKAYANAGI. 1998a. Phenotypic variation and the relationships among mustard (*Brassica juncea* L.) germplasm from Pakistan. *Euphytica* 101(3): 357-366.

RABBANI, M.A., A. IWABUCHI, Y. MURAKAMI, T. SUZUKI and K. TAKAYANAGI. 1998b. Genetic diversity in mustard (*Brassica juncea* L.) germplasm from Pakistan as determined by RAPDs. *Euphytica* 103: 235-242.

SMITH, O.S., J.S.C. SMITH, S.L. BOWEN, R.A. TENBORG and S.J. WALL. 1990. Similarities among a group of elite maize inbreeds as measured by pedigree, F1 grain yield, grain yield, heterosis, and RFLPs. *Theor. Appl. Genet.* 80: 833-840.

SONG, K.M., T.C. OBSORN and P.H. WILLIAMS. 1988a. *Brassica* taxonomy based on nuclear restriction fragment length polymorphism (RFLPs). I. Genome evolution of diploid and amphiploid species. *Theor. Appl. Genet.* 75: 784-794.

SONG, K.M., T.C. OBSORN and P.H. WILLIAMS. 1988b. *Brassica* taxonomy based on nuclear restriction fragment length polymorphism (RFLPs). Preliminary analysis of sub-species within *Brassica rapa* (syn campestries) and *Brassica oleraceae*. *Theor. Appl. Genet.* 76: 593-600.

STAUB, J., J. BACHER and K. PEOTTER. 1996. Sources of potential errors in the application of random amplified polymorphic DNAs in cucumber. *HortScience* 31: 262-266.

TANKSLEY, S.D., N.D. YOUNG, A.H. PATERSON and M.W. BONIERVAL. 1989. RFLP mapping in plant breeding: new tools for an old science. *Biotechnology* 7: 257-264.

TAUTZ, D. 1989. Hypervariability of simple sequences as general sources for polymorphic DNA markers. *Nucleic Acids Res.* 17(16): 6463-6471.

TOHME, J., D.O. GONZALEZ, S. BEEBE and M.C. DUQUE, 1996. AFLP analysis of gene pools of a wild bean core collection. *Crop Sci.* 36: 1375-1384.

UN. 1935. Genome analysis in *Brassica* with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. *Jap. J. Bot.* 9: 389-452.

VIRK, P.S., J. ZHU, H.J. NEWBURY, G.J. BRYAN, M.T. JACKSON and B.V. FORD-LLOYD. 2000. Effectiveness of different classes of molecular markers for classifying and revealing variations in rice (*Oryza sativa*) germplasm. *Euphytica* 112: 275-284.

VOS, P., R. HOGERS, M. BLEEKER, M. REIJANS, T. VAN DE LEE, A. FRIJTERS, J. POT, J. PELEMAN, M. KIPER and M. ZABEAU. 1995. AFLP: A new technique for DNA fingerprinting. *Nucleic Acids. Res.* 23(21): 4407-4414.

WELSH, J. and M. McCLELLAND. 1990. Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Res.* 18: 7213-7218.

WILLIAMS, G.K., A.R. KUBELIKE, K.J. LIVAK, J.A. RAFALSKI, and S.V. TINGEY. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res.* 18: 6531-6535.

CHAPTER VI

GENOTYPE X ENVIRONMENT INTERACTIONS AND STABILITY ANALYSES OF ETHIOPIAN MUSTARD

6.1. ABSTRACT

The genotype x environment (GE) interaction and the stability of 14 B. carinata genotypes for seed yield were evaluated at four rain fed locations for three years in north-western Ethiopia. The objectives were to estimate the component of variance associated with the first and second ordered interactions and to determine their effects and to measure the genotypes' stability using different stability statistics, i.e., mean ($\bar{x}$); regression coefficient (b_i); deviation from regression (s^2d_i); coefficient of variability (CV_i); stability variance ($\sigma^2{}_i$); coefficient of determination (r^2); ecovalence (W_i); cultivar superiority measure (P_i); variance of ranks (S1); and AMMI stability value (ASV) for each genotype and to compare the stability parameters.

A standard multi-factor analysis of variance test revealed that the main effects due to years, locations, and the first order interactions (year x location) as well as second order interactions (genotype x year x location) were highly significant. The genotype x location interaction was not significant. The non-significant genotype x location effect and significant genotype x year effect indicate the necessity for more emphasis on multiple years than a series of locations. The estimates of variance components involving first order interaction of genotype x year ($\sigma^2{}_{gy}$), genotype x location ($\sigma^2{}_{gl}$) and the error variance ($\sigma^2{}_e$) were very small. The variance components of the main effects of the genotype ($\sigma^2{}_g$) and second order interaction of genotype x location x year ($\sigma^2{}_{gly}$) were very high. Mean yield over all locations and years ranged from 1304 to 1541 kg/ha by PGRC/E 210406 and PGRC/E 21163 respectively. The regression of genotype mean yield on the environmental index resulted in a b_i value ranging from 0.7488-1.2288. The, s^2d_i were non significant for all the test genotypes. The first two interaction principal component axes of the additive-main-

effects-and-multiplicative-interaction (AMMI) model together accounted between 35.7% and 54.3% of the total genotype-environment interaction of sum of squares for seed yield. The stability analyses identified PGRC/E 20017 and PGRC/E 20112 as more stable genotypes, while PGRC/E 21169 and (4DxZEM-1) x (ZEM-1-AD/88) were specifically adapted to some environments. Rank correlation coefficient was performed to study the relationship of stability parameters. The statistic, S^2d_i was highly and significantly rank correlated with σ^2_i, W_i, $S1$, and ASV. The statistic σ^2_i was also highly significantly rank correlated with W_i, S^2d_i, and $S1$. The b_i statistic was also highly significantly rank correlated with CV_i. There was highly significant rank correlation with $S1$ and σ^2_i, S^2d_i, and W_i. The mean seed yield was highly significantly rank correlated with P_i. These highly significant rank correlations among stability statistics indicated their effectiveness in detecting stable genotypes over a range of environments. To reduce the impact of GE interaction effects further, genotypes and environments were stratified by cluster analysis. The hierarchical clustering of AMMI adjusted mean yield using unweighted pair group method with arithmetic average (UPGMA) clustering method grouped the genotypes into five clusters and the environments into three heterogeneous clusters.

Key words: *Brassica carinata*, GE interaction, Ethiopian mustard, stability statistics

6.2. INTRODUCTION

Genotype x environment (GE) interaction, which is associated with the differential performance of genetic materials, tested at different locations and in different years, and its influence on the selection and recommendation of cultivars has long been recognized (Lin *et al.*, 1986; Geletu *et al.*, 1995; Annicchiarico, 1997a). Selection of genotypes is based on the assessment of their phenotypic value in varying environments. Evaluation of genotypic performance at a number of locations provides useful information to determine their adaptation and stability (Crossa, 1990). Measuring GE interactions helps to determine an optimum breeding strategy of either to breed for specific or wide adaptation, which depends on the expression of stability

under a limited or a wide range (Romagosa and Fox, 1993; Annicchiarico, 1997b; Yue *et al.*, 1997).

The wide range of methods available for the analysis of GE interactions and stability analyses can be divided into four groups: 1) the analysis of components of variance, 2) stability analysis, 3) qualitative methods and 4) multivariate methods (Hohls, 1995).

The method of studying GE interactions by components of variance has been widely used, and is well documented (Comstock and Moll, 1963; Sprague and Federer, 1951). Variance components are used to separate the effects of genotypes, environments and their interactions by equating the observed and expected mean squares in the analysis of variance.

GE interactions are classified into two types; quantitative (without change in ranks) and qualitative (with changes in ranks) and it is the latter type that presents a challenge to breeders (Romagosa and Fox, 1993). In the presence of interaction, mean yield is inadequate, as it does not sufficiently indicate consistency of performance (Francis and Kannenberg, 1978).

Stability refers to non-erratic performance with respect to agronomic traits and stable cultivars, which shows minimal or low interactions (Allard and Bradshaw, 1964). Stability also denotes consistency in rank relative to other cultivars in a given set of environments (Yue *et al.*, 1997). Becker and Léon (1988) stated that varieties with high and stable performance are desirable. Instability is caused by the differential expression of traits across environments or GE interaction, which decreased progress from selection since they reduce the association between genotypic and phenotypic values. This makes the selection of widely adapted and stable cultivars, which overcome the problem of interaction very complicated (Romagosa and Fox, 1993).

The concept of stability of a genotype in an evaluation and breeding program is, however, ambiguous (Lin *et al.*, 1986), often used in different senses (Becker and Léon, 1988) and based on different statistical determinations and analyses (Crossa, 1990; Hohls, 1995). Lin *et al.* (1986) identified three concepts of stability: Type I is

defined as a genotype that is stable if its variance over a range of environments is small; Type II is defined as a genotype that is stable if its response to environments is parallel to the mean response of all genotypes in the trial. This type of stability is primarily based on the interpretation of the regression coefficient in joint linear regression. Type III stability is defined as a genotype that is stable if the residual mean squares from the regression model on the environment index is small (Eberhart and Russell, 1966).

The concept of stability has been defined in several ways and several biometrical methods including univariate and multivariate ones have been developed to assess stability (Hill, 1975; Lin *et al.*, 1986; Westcott, 1986; Becker and Léon, 1988; Crossa, 1990). Joint regression analysis is the most popular among the univariate methods because of simplicity of calculation and applications (Becker and Léon, 1988).

Becker and Léon (1988) stated that all stability procedures based on quantifying GE interaction effects belong to the dynamic concept. Included are procedures partitioning GE interaction, such as Wricke's ecovalence (Wricke, 1962), and Shukla's stability of variance (Shukla, 1972) procedure using the regression approach such as proposed by Finlay and Wilkinson (1963), Eberhart and Russell (1966) as well as non-parametric statistics (Nassar and Hühn, 1987).

Lin and Binns (1988a; 1988b) also proposed the cultivar performance measure (P_i), and defined P_i of genotype i as the mean square of distance between genotype i and the genotype with the maximum response. The smaller the value of P_i, the less its distance to the genotype with maximum yield and thus the better the genotype.

The main problem with stability statistics is that they do not provide an accurate picture of the complete response pattern (Hohls, 1995). The reason for this is that a genotype's response to varying environment is multivariate (Lin *et al.*, 1986) whereas the stability indices are usually univariate. Through multivariate analysis, genotypes with similar responses can be clustered, and thus the data can be summarized and analysed more easily (Gauch, 1982b; Crossa, 1990). Methods of classification have the advantage over stability statistics of having a broader inferential base.

Characterization of the response patterns of genotypes to environmental change enables extrapolation to a much wider range of environments than those tested (Hohls, 1995).

One of the multivariate techniques is the additive main effects and multiplicative interaction method, known as the AMMI model. It combines the analysis of variance of genotype and environment main effects with principal component analysis of the GE interaction into a unified approach (Gauch, 1988; Zobel *et al.*, 1988; Gauch and Zobel, 1996). The results can be graphically represented in an easily interpretable and informative biplot that shows both main effects and GE interactions. The AMMI model has been used extensively and with success over the past few years to analyse and understand various crop GE interactions (Crossa, 1990; Yau, 1995; Yan and Hunt, 1998).

The objectives of this study were to estimate the component of variance associated with the first and second ordered interactions and to determine their effects and to measure the genotypes' yield stability of Ethiopian mustard in the north-western Ethiopian using different stability statistics.

6.3. MATERIALS AND METHODS

Plant material

Fifteen genotypes including Yellow Dodolla (standard check) and local check were studied (Table 6.1). These varieties were selected based on good agronomic performance and disease resistance from the national collections from Plant Genetic Resources of Ethiopia (now Biodiversity Conservation and Research Institute) by the standard pedigree method. The local check was the cultivar of the farmers, which has been under production at each location, while Yellow Dodolla was one of the improved varieties released in 1986 by Holleta Research Centre from the National yield trial.

Table 6.1. Description of test materials used in this study.

No	Genotypes	Origin/description
1	PGRC/E 21163	1[a]
2	PGRC/E 21169	1
3	PGRC/E 21184	1
4	PGRC/E 21258	1
5	PGRC/E 20017	1
6	PGRC/E 20021	1
7	PGRC/E 20112	1
8	PGRC/E 20156	1
9	PGRC/E 210406	1
10	PGRC/E 20165	1
11	M-1	1
12	(4DXZEM-1) X (ZEM-1-AD/88)	2
13	(4DXZEM-1) X (ZEM-1-F5/10)	2
14	Yellow Dodolla	Standard check
15	Local check	Farmer's cultivar per location

[a]1=selection from landrace collections by pedigree method

2=crosses between high and low erucic acid

Experimental layout and testing sites

This experiment was conducted at four rain-fed locations (Adet, Mota, Finote Selam and Debre Tabor) (Table 6.2) representing the major crop growing agro-ecologies of mid and high altitude areas of the north-western Ethiopia. The experiment layout was a randomised complete block design (RCBD) with three replications. Each plot

consisted of six rows spaced 30 cm apart and 5m long. The net plot harvested was 6 m^2.

Table 6.2. Location, coordinates and their agro-climatic conditions of trial sites in north-western Ethiopia, 1997-1999.

Trial Sites	Mean annual Rainfall (mm)	Temperature (°C)		Soil types	Altitude (Meter)	Latitude	Longitude
		(Min)	(Max)				
Adet	1244.93	8.87	26.06	Cambisol	2240	11° 16' N	37° 29'E
Mota	1679.47	9.49	23.60	Nitosol	2470	11° 20'N	37° 88'E
D/Tabor	1235.63	9.58	21.87	Luvisol	2630	11° 89'N	38° 09'E
F/Selam	988.73	12.81	28.35	Nitosol	1935	10° 84'N	37° 36E

Date of planting varied from late May to early June for these locations. Fertilizer (both N and P_2O_5) was applied at the rate of 46/69 kg/ha each at the time of planting. A seed rate of 10 kg/ha was applied for six rows. Seed and fertilizer were drilled uniformly by hand. Weeding was carried out 30 days after emergence and again 45 days after the first weeding. Neither herbicides nor insecticides were applied. The plots were harvested individually, dried for 15-30 days to 8% moisture threshed and cleaned.

Measurements

Data on seed yield and agronomic and disease traits were taken from the center middle four rows of each plot. Days to flowering and days to maturity were separately taken when each plot reached 50% flowering, and 75% maturity respectively. The days were calculated beginning from the date of sowing. Plant height (cm) was taken at full maturity by measuring five randomly selected plants from the ground level to the top of the plant. Scores for disease reaction were recorded at a peak infection period (usually between flowering and maturity). Leaf and pod spot (*Alternaria brassicae*) and white rust (*Albugo candida*) were recorded on a 0-5 scoring scale (0=nil; 5=severe). Disease scores were transformed by square root method before analysis. Stand percentage was taken by counts at the time of maturity and square root transformed before analysis. Seed yield data was recorded on clean, dry samples; plot yields were converted to kilogram per hectare. Then, combined analyses of variance were performed using data across locations and years.

Statistical analyses

The following statistical analyses were conducted using appropriate statistical packages: Agrobase^{TM} 2000 (Agronomix Software Inc.,), SAS (SAS Institute, 1996), NCSS 2000 (Hintze, 2000), and MSTAT-C (Michigan State University, 1991), to determine:

1. Analysis of variance (ANOVA) estimation of variance components

Bartlett's test was used to establish the homogeneity of variances between environments to determine the validity of the combined analyses of variance on the data. Various transformations were conducted on the data set. For combined analysis of variance 14 entries out of 15 were considered, as the local check (farmer's varieties) were different from location to location. The mean values of each trait were subjected to combined analyses of variance using the PROC GLM procedure of SAS computer program (SAS Institute Inc., 1996) to determine the effects of genotypes, locations and years as well as their first and second order interactions. Genotypes were assumed to be fixed, and year and location effects random. Significance levels were determined as suggested by McIntosh (1983). The ANOVA method for estimating variance components consists of equating mean squares to their expectations and solving the resulting set of simultaneous equations. The observed mean squares were equated to their expectations (Comstock and Moll, 1963) and solved to obtain estimates of the variance components. Approximate standard errors of the estimates were also calculated.

The form of the analysis of variance and the expected mean square composition and method of determining the individual estimates of the variance components are shown in Tables 6.3 and 6.4 and are based on the model provided by Allard (1960), which was developed by Comstock and Moll (1963) for the determination of interaction variance components.

Table 6.3. Form of variance analysis and mean square expectations for GE interaction.

Source	DF	MS	Expected mean square
Years (Y)	(Y-1)		
Locations (L)	(L-1)		
YxL	(Y-1) (L-1)		
Replicates in locations and years	LY (R-1)		
Genotypes (G)	(G-1)	M5	$\sigma^2_e + r\sigma^2_{gly} + rl\sigma^2_{gy} + ry\sigma^2_{gl} + rly\sigma^2_g$
GxL	(G-1) (L-1)	M4	$\sigma^2_e + r\sigma^2_{gly} + ry\sigma^2_{gl}$
GxY	(G-1) (Y-1)	M3	$\sigma^2_e + r\sigma^2_{gly} + rl\sigma^2_{gy}$
GxLxY	(G-1) (L-1)(Y-1)	M2	$\sigma^2_e + r\sigma^2_{gly}$
Error	LY (G-1) (R-1)	M1	σ^2_e

Where, Y, L, G and R are number of years, locations, genotypes and replications, respectively. The σ^2_e, σ^2_g are components of variance of error and genotypes respectively. Combinations of the subscript identify the components, for the interactions. M1 to M5 are the observed values of the various mean squares.

Table 6.4. Estimates of variance components and method of determination for GE interaction.

Variance component	Method of determination
Genotypes (σ^2_g)	$M5 + M2 - M3 - M4/rly$
Genotypes x locations (σ^2_{gl})	$M4 - M2/ry$
Genotypes x years (σ^2_{gy})	$M3 - M2/rl$
Genotypes x locations x years (σ^2_{gly})	$M2 - M1/r$
Error (σ^2_e)	$M1$

Where, M1 to M5 are the values of the appropriate mean squares as indicated in Table 6.3; r, l, and y are numbers of replicates, locations and years, respectively, in which the cultivars were evaluated.

All the following analyses of the stability models and AMMI were performed using Agrobase 2000^{TM} (Agrobase, 2000) while the Spearman's rank correlation coefficient was calculated by MSTAT-C (Michigan State university, 1991).

2. Coefficient of variability (CV_i) (Francis and Kannenberg, 1978)

3. Shukla's procedure of stability variance ($\sigma^2{}_i$) (Shukla, 1972)

4. Lin and Binns' cultivar performance measure (P_i) (Lin and Binns, 1988a)

5. Finlay and Wilkinson's joint regression analysis (b_i) (Finlay and Wilkinson, 1963).

6. Eberhart and Russell's joint regression analysis ($s^2 d_i$) (Eberhart and Russell, 1966).

7. Coefficient of determination (r^2) (Pinthus, 1973)

8. Wricke's ecovalence (W_i) (Wricke, 1962)

9. Mean absolute difference (S1) and variance of ranks (Nassar and Hühn, 1987).

10. AMMI model (Gauch, 1988)

Since the AMMI model does not make provision for a specific stability measure to be determined, and as such a measure is essential for this study in order to rank genotypes in terms of stability, such a measure is proposed by Purchase (1997) as:

$$ASV = \sqrt{\left[\frac{IPCA1SS}{IPCA2SS}(IPCA1)\right]^2 + \left[IPCA2score\right]^2}$$

Where, ASV=AMMI stability value, SS=sum of squares and IPCA=Interaction of principal components analysis.

11. Combined comparison of stability analysis procedures

To statistically compare the 10 stability analysis procedures, Spearman's coefficient of rank correlation (r_s) was employed (Steel and Torrie, 1980). All the genotypes evaluated were respectively assigned stability values according to the procedure and definitions used, and were then ranked in order to determine Spearman's rank correlation coefficient between the different procedures. Assume n genotypes are arranged in the same following order to two stability parameters X_i indicates the

ranking order (or ranking number) of the i^{th} genotype for the first parameter, Y_i, indicates the ranking number of the i^{th} genotype of the second parameter, then $d_i = X_i - Y_i$ (i=1,2...n) and Spearman's rank correlation coefficient (r_s) can be described as:

$$r_s = \frac{6\Sigma d^2{}_i}{n(n^2-1)}$$

Ranking numbers are whole numbers and when two or more equal ranking numbers occur, the average of the ranking numbers that they otherwise would have received are ascribed to each genotype.

12. Cluster analysis of genotypes and environments based on AMMI adjusted means was performed using the NCSS 2000 software computer program (Hintze, 2000). For the purpose of classification, cluster analysis with the unweighted pair group method with arithmetic average (UPGMA) clustering method has been applied with the NCSS software program (Hintze, 2000). The Euclidean distance matrix was computed. Before computing the distance between locations (genotypes), the data were standardized as recommended by Fox and Rosielle (1982). By standardizing the data, each environment and genotype will have a mean of zero and unit variance, and the effect of variability in phenotypic variance (as well as the mean) should be reduced.

6.4. RESULTS AND DISCUSSION

1. Analysis of variance (ANOVA) estimation of variance components

The relative performance of genotypes based on mean seed yield and other agronomic characters over years and locations are presented in Table 6.5. As shown in Table 6.7, significant differences among genotypes were observed for all traits. Table 6.7, also includes the mean squares relevant to the study of the genotype-environment interactions from the analysis of variance over years and locations.

The partitioning of the sum of squares of the components (Table 6.6) indicated that 51.81% of the total variability was due to locations, 14.35% due to error, 9.91 due to years, 5.72% due to genotype x year x location interactions and the genotype was only 1.18% of the total variation.

The genotypes x year interaction mean squares were significant for all traits except alternaria leaf spot (ALS) (Table 6.7). This would suggest that testing over years is of the same value for determining relative performance among genotypes. Genotypes had considerable larger mean squares than the other sources of variance. So, adequate data for these traits could be obtained from minimum test environments over years. That is, evaluation over years would probably not be necessary.

Main effects due to year, location, year x location, genotype x year x location were highly significant ($P<0.01$) for seed yield. Main effects due to genotypes, and genotype x year were significant ($P<0.05$). The genotype x location interaction was not significant (Table 6.6).

When individual estimates of variances for seed yield (Table 6.8) were expressed as a percent of the total variation ($\sigma^2_g+\sigma^2_{gy}+\sigma^2_{gl}+\sigma^2_{gly}+\sigma^2_e$) the σ^2_g component accounted for 0.47 % of the total. The σ^2_e component accounted 84.13 % of the total variation. The estimate of σ^2_{gl} accounted negative which will be 0 % of the total variation and the corresponding mean square was not significant which suggested that the performance of genotypes relative to each other were similar at each location over

years and further indicates a certain degree of redundancy in this sample of locations. The consistent performance of *B. carinata* genotypes at four locations implies that with respect to seed yield the recommendation of specific genotypes from among the samples tested in this study for the specific region in north-western Ethiopia is probably not warranted.

The σ^2_{gy} component comprised 2.08% of the total variation. It was significant ($P<0.05$) and greater than the σ^2_{gl} component (Table 6.7). The performance of the genotypes were therefore, less consistent over years than over locations. The relative importance of the σ^2_{gy} component suggests a need for testing over more years than more locations. The σ^2_{gly} component made up 13.32% of the total variation. It was highly significant ($P<0.01$) and larger relative to the σ^2_{gl} and σ^2_{gy} components. This suggests that there is an important differential response to environmental variations that is not accounted for by either the year or by location groupings. The large σ^2_{gly} component may suggests that factors such as precipitation, temperature, and disease incidence can result in conditions unique to each year-location combination, and that the genotypes respond differently to these conditions.

The variance component estimates and their standard errors for genotypes and their interactions with locations and years are given in Table 6.8. They were calculated from the mean squares in Table 6.7. Some estimates were negative which associates some error of estimation. Since variances theoretically cannot be negative, one must assume in the cases where such estimates are obtained that the quantity being estimated is either zero or a relatively small positive number. In this analysis negative estimates were assumed to be zero (Rao, 1997). The standard errors are informative in this respect. When the standard errors were smaller than the components, adequate precision was attained.

Error variances for each trait were smaller than the components. The Interaction variance components were relatively small in comparison to the genotypes variance component for all traits. The first and second order interaction components (σ^2_{gy}, σ^2_{gly}) were higher compared to the genotypes (σ^2_{g}) for seed yield.

The significance of environmental effects and their interactions warrant further analyses of the stability and adaptation of the genotypes.

Table 6.5. Mean performance of genotypes across years and over locations for characters.

No	Genotypes	Yield	DF	DM	PH	SP	WR	ALS
1	PGRC/E 21163	1541	82	153	173	88	1.11	1.18
2	PGRC/E 21169	1357	94	168	195	87	1.25	1.29
3	PGRC/E 21184	1388	93	165	182	87	1.23	1.28
4	PGRC/E 21258	1360	86	156	183	87	1.19	1.19
5	PGRC/E 20017	1363	90	163	187	88	1.22	1.28
6	PGRC/E 20021	1527	91	162	194	90	1.24	1.28
7	PGRC/E 20112	1431	90	164	194	89	1.22	1.29
8	PGRC/E 20156	1403	92	167	197	87	1.23	1.26
9	PGRC/E 210406	1304	78	157	178	86	1.17	1.18
10	PGRC/E 20165	1370	92	164	197	86	1.23	1.28
11	M-1	1347	105	182	187	82	1.08	1.17
12	(4DXZEM-1) X (ZEM-1-AD/88)	1427	76	149	170	85	1.08	1.14
13	(4DXZEM-1) X (ZEM-1-F5/10)	1339	77	149	168	85	1.10	1.12
14	Yellow Dodolla	1349	85	154	183	84	1.12	1.14

DF=days flowering; DM=days maturity; PH=plant height; SP=stand percent; WR=white rust; ALS=alternaria leaf spot

Table 6.6. Sum of squares and its percentage (out of total) contribution of the combined analysis of seed yield of 14 *B. carinata* genotypes tested over 12 environments.

Source	DF	SS	MS	%SS
Year	2	18028896.8	9014448.4**	9.91
Location	3	94252632.8	31417544.3**	51.81
Y*L	6	17430611.9	2905102.0**	9.58
Rep (Y*L)	24	5686835.7	236951.5**	3.13
Genotypes	13	2143669.2	164897.6*	1.18
G*Y	26	3852722.4	148181.6*	2.12
G*L	39	4026803.9	103251.4	2.21
G*Y*L	78	10410552.2	133468.6**	5.72
Error	312	26098062	83648	14.34
Total	503	181930787		
		CV (%)=20.7	R^2=0.856	

*,** =Significant at 0.05 and 0.01 probability levels respectively

Table 6.7. Mean squares and variance components relevant to the study of GE interaction.

Sources	DF	Kg/ha	DF	DM	PH	SP	WR	ALS
Genotypes	13	164897.6*	2220**	2756**	3540**	165**	0.159**	0.155**
G*Y	26	148181.6*	112*	144*	408**	102**	0.058**	0.050
G*L	39	103251.4	114**	186**	690**	50	0.090**	0.068**
G*Y*L	78	133468.6**	151**	100	235	89	0.044**	0.073**
Error	312	83648	68	91	215	46	0.023	0.033
R^2		0.86	0.91	0.85	0.86	0.78	0.93	0.88
CV (%)		21	9	6	8	8	13	15

*,** =Significant at 0.05 and 0.01 probability levels respectively.

Table 6.8. Estimates of variance components and their standard errors for genotypes and their interactions with locations and years.

Components[a]	Yield	DF	DM	PH	SP	WR	ALS
σ^2_g (s.e)	464 (2128)	58 (24)	70(30)	74 (39)	1.7 (1.9)	0.002 (0.002)	0.002 (0.001)
σ^2_{gy} (s.e)	2066 (3679)	0	4 (3)	14 (10)	2.2 (2.5)	0.001(0.001)	0
σ^2_{gl} (s.e)	0	0	10 (5)	51 (18)	0	0.01 (0.002)	0.0002 (0.002)
σ^2_{gly} (s.e)	13249 (5 822)	22 (6)	3 (6)	7 (14)	4 (3)	0.01 (0.002)	0.011 (0.003)
σ^2_e (s.e)	83648	68	91	215	46	0.023	0. 034

[a]G, L and Y refers to genotypes, locations and years respectively.

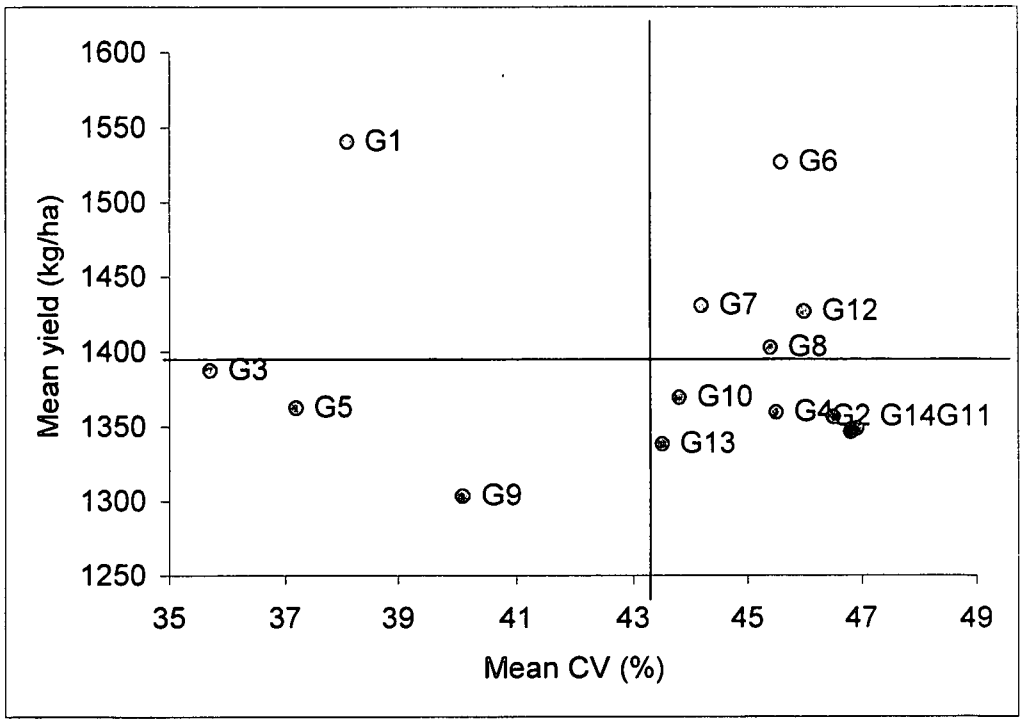
This result shows that the most important source of variation was locations and these indicated the wide and divergent genetic response. Recently, Adugna and Labuschagne (2002) in GE interaction and stability study on linseed in Ethiopia indicated the impacts of environmental variability, especially rainfall, which differed greatly across the locations and years along with other edaphic, climatic and biotic factors. Tesfaye *et al.* (1998) have also reported a similar GE and stability study, indicating the impacts of agro-ecological diversity of Ethiopian teteraploid wheat.

Basford and Cooper (1998) in their recent review of GE interactions in wheat in Australia distinguished between two major categories of GE interactions: (1) defined causes and (2) the undefined causes. The defined causes of GE interactions are due to diseases, soil borne constraints (nutritional, deficiencies, toxicities, pathogens, etc.,), crop phenology, drought, and poor experimental methods. Similarly, for the undefined causes of GE interactions they emphasized inadequacy of detailed characterization of the environments, limited time to invest in detailed investigation of interactions, and the lack of a general frame work for investigating the causes of GE interactions that accommodates the influence of both abiotic and biotic stresses on plant adaptation.

2. Francis and Kannenberg's coefficient of variability (CV_i)

The mean-CV analysis introduced by Francis (1977) was designed primarily to aid in studies on the physiological basis for yield stability. It is a simple graphical approach to assess performance and stability concurrently. It measures the performance and CV for each genotype over all environments and the mean yield plotted against the CV. It was found to characterize genotypes on a group basis rather than individually (Francis and Kannenberg, 1978).

Mean yield is plotted against CV in Figure 6.1. High yield and small variation group of genotypes appear the most desirable using any approach. A stable genotype is the one that provides high and consistent performance. According to this definition, only PGRC/E 21163 from high yield and low variation group can be considered as stable.



Note: G1=PGRC/E 21163; G2=PGRC/E 21169; G3=PGRC/E 21184; G4=PGRC/E 21258; G5=PGRC/E 20017; G6=PGRC/E 20021; G7=PGRC/E 20112; G8=PGRC/E 20156; G9=PGRC/E 210406; G10=PGRC/E 20165; G11=M-1; G12=(4DxZEM-1) x (ZEM-1-AD/88); G13=(4DxZEM-1) x (ZEM-1-F5/10); G14=Yellow Dodolla.

Figure 6.1. Mean yield (kg/ha) plotted against CV (%) from data collected on 14 *B. carinata* genotypes in 12 environments.

3. Lin and Binn's cultivar superiority performance (P_i)

According to Lin and Binn's method (1988a), the cultivars superiority performance (P_i) was estimated by the square of differences between a genotype and maximum genotype mean at a location, summed and divided by twice the number of locations. Genotypes with lower P_i values are considered the most stable. Table 6.9 presents the cultivars superiority measure (P_i) for seed yield of 14 *B. carinata* genotypes tested in 12 environments of north-western Ethiopia during the period of 1997-1999.

Table 6.9. Lin and Binn's cultivar performance measure (P_i) for 14 genotypes included in GE trials in the north-western Ethiopia, 1997-1999.

No	Cultivars	P_i	Rank	Mean yield	Rank
1	PGRC/E 21163	50237.60	2	1541	1
2	PGRC/E 21169	105278.66	9	1357	10
3	PGRC/E 21184	110743.32	11	1388	6
4	PGRC/E 21258	104101.35	8	1360	9
5	PGRC/E 20017	96044.31	6	1363	8
6	PGRC/E 20021	42884.11	1	1527	2
7	PGRC/E 20112	61830.55	3	1431	3
8	PGRC/E 20156	86055.47	4	1403	5
9	PGRC/E 210406	129117.58	14	1304	14
10	PGRC/E 20165	96235.80	7	1370	7
11	M-1	117541.69	12	1347	12
12	(4DXZEM-1) X (ZEM-1-AD/88)	95400.34	5	1427	4
13	(4DXZEM-1) X (ZEM-1-F5/10)	110236.60	10	1339	13
14	Yellow Dodolla	118110.74	13	1349	11

From this analysis, PGRC/E 20021, PGRC/E 21163 and PGRC/E 20112 had the lowest P_i values and thus the best stability, while M-1, Yellow Dodolla and PGRC/E 210406 had the poorest stability. In this analysis the ranks of cultivar's superiority measure were in agreement with that of the overall mean yield (Table 6.9).

4. Wricke's ecovalence analysis (W_i)

Wricke (1962) defined the concept of ecovalence as the contribution of each genotype to the genotype x environment interaction sum of squares. The ecovalence (W_i) or the stability of i^{th} genotype is its interaction with environments, squared and summed across environments. Genotypes with low ecovalence have smaller fluctuations from the mean across different environments and are therefore stable.

In view of this principle, Wricke's ecovalence was determined for each of 14 genotypes evaluated over 12 environments in the north-western Ethiopia (Table 6.10).

Table 6.10. Wricke's ecovalence value (WEV) for 14 genotypes over 12 environments of north-western Ethiopia, 1997-1999.

No	Cultivars	Ecovalence	Rank	Mean yield	Rank
1	PGRC/E 21163	529551.64	9	1541	1
2	PGRC/E 21169	739497.74	14	1357	10
3	PGRC/E 21184	409430.63	7	1388	6
4	PGRC/E 21258	661044.96	13	1360	9
5	PGRC/E 20017	128521.38	1	1363	8
6	PGRC/E 20021	315760.67	5	1527	2
7	PGRC/E 20112	169988.33	2	1431	3
8	PGRC/E 20156	540402.16	10	1403	5
9	PGRC/E 210406	613393.20	11	1304	14
10	PGRC/E 20165	220283.69	3	1370	7
11	M-1	477082.75	8	1347	12
12	(4DXZEM-1) X (ZEM-1-AD/88)	656500.60	12	1427	4
13	(4DXZEM-1) X (ZEM-1-F5/10)	307687.07	4	1339	13
14	Yellow Dodolla	336472.48	6	1349	11

According to Wricke's ecovalence analysis, the genotypes PGRC/E 20017, PGRC/E 20112, and PGRC/E 20165 were the most stable genotypes. Whereas, PGRC/E 21169, PGRC/E 21258 and (4DxZEM-1)X(ZEM-1-AD/88) were found to be unstable.

5. Shukla's stability variance parameter (σ^2_i)

Shukla's stability variance (1972), mean yield and ranking order of genotypes to the values were given in Table 6.11. According to this stability parameter, entries with minimum stability variance are considered more stable. Accordingly, PGRC/E 20017, PGRC/E 20112, and PGRC/E 20165 are the most stable genotypes, while PGRC/E 21169, PGRC/E 21258, and (4DxZEM-1) x (ZEM-1-AD/88) were classified as the least stable ones.

Table 6.11. Shukla's stability variance, for 14 *B. carinata* genotypes tested across 12 environments of north-western Ethiopia, 1997-1999.

No	Cultivars	Stability variance with no covariate	Rank	Mean yield	Rank
1	PGRC/E 21163	157819.55	9	1541	1
2	PGRC/E 21169	224620.58	14	1357	10
3	PGRC/E 21184	119599.23	7	1388	6
4	PGRC/E 21258	199658.33	13	1360	9
5	PGRC/E 20017	30219.01	1	1363	8
6	PGRC/E 20021	89795.15	5	1527	2
7	PGRC/E 20112	43413.04	2	1431	3
8	PGRC/E 20156	161271.99	10	1403	5
9	PGRC/E 210406	184496.41	11	1304	14
10	PGRC/E 20165	59416.11	3	1370	7
11	M-1	141124.90	8	1347	12
12	(4DXZEM-1) X (ZEM-1-AD/88)	198212.40	12	1427	4
13	(4DXZEM-1) X (ZEM-1-F5/10)	87226.28	4	1339	13
14	Yellow Dodolla	96385.27	6	1349	11

6. Finlay and Wilkinson joint regression analysis

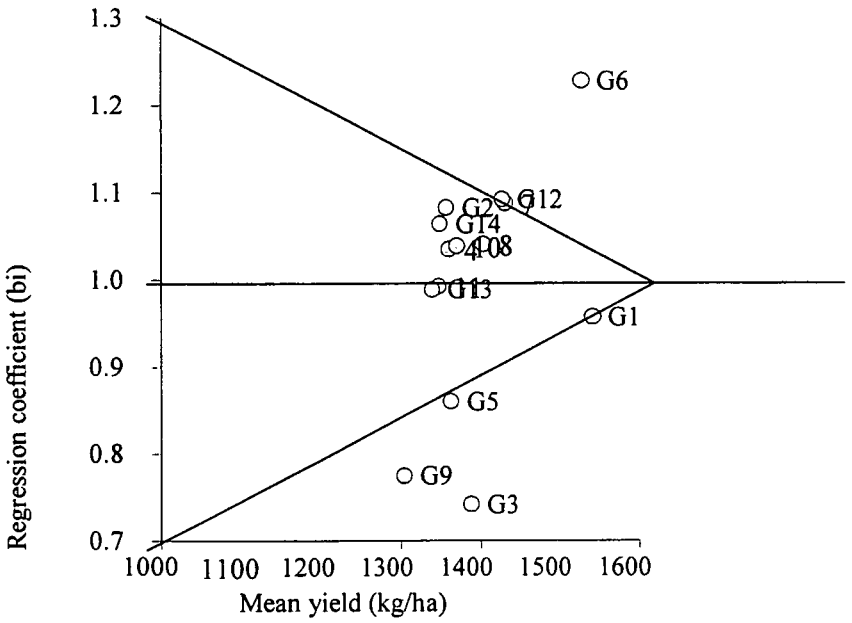
According to Finlay and Wilkinson (1963), regression coefficients approximating to 1.0 indicate average stability. When this is associated with high mean yield, genotypes have general adaptability. When this is associated with low mean yield, genotypes are poorly adapted to all the environments. Regression coefficients above 1.0 indicate genotypes with increasing sensitivity to environmental change, thus below average stability and great specific adaptability to high-yielding environments. Regression coefficients decreasing below 1.0 provide a measure of greater resistance to environmental change, (above average stability), and therefore, increasing specificity of adaptability to low yielding environments. The second index, the genotype mean yield over all environments, provides a comparative measure of performance of the individual genotypes.

The mean seed yield over all locations and years ranged from 1304 to 1541 kg/ha by PGRC/E 210406 and PGRC/E 21163 respectively. The regression of genotype mean yield on the environmental index resulted in a regression coefficient, b_i ranging from 0.7428-1.2288. Based on the regression coefficient, b_i the genotypes tested could be categorized as those having more than average stability, average stability, and below average stability in the ranges of environments they encountered.

In line with the above concepts and principles the three genotypes PGRC/E 20021, PGRC/E 20112, and (4DXZEM-1) X (ZEM-1-AD/88) showed below average stability, but with specifically good adaptation to favourable conditions. Genotypes PGRC/E 20017, PGRC/E 210406, and PGRC/E 21184 indicated above average stability, but also specific adaptability to unfavourable environments (See Figure 2.5.1 and 6.2). On the other hand, the genotypes M-1, (4DXZEM-1) X (ZEM-1-F5/10), PGRC/E 21258, PGRC/E 20165, PGRC/E 20156, and PGRC/E 21169 all showed average stability and increasing adaptability to all environments.

Generally, the trend (Table 6.13) was that genotypes with high b_i values tended to have higher coefficient of variation (CV) and vice versa. The coefficient of determination (r^2), which is a measure of the explanatory power of the regression

model, also measures variation. The values of r^2 for seed yield were mostly very high indicating environmental effects as the main determinants of phenotypic performance.



Note: G1=PGRC/E 21163; G2=PGRC/E 21169; G3=PGRC/E 21184; G4=PGRC/E 21258; G5=PGRC/E 20017; G6=PGRC/E 20021; G7=PGRC/E 20112; G8=PGRC/E 20156; G9=PGRC/E 210406; G10=PGRC/E 20165; G11=M-1; G12=(4DxZEM-1) x (ZEM-1-AD/88); G13=(4DxZEM-1) x (ZEM-1-F5/10); G14=Yellow Dodolla.

Figure 6.2. Regression coefficients plotted against genotype mean yield.

7. Eberhart and Russell joint regression analysis

Eberhart and Russell (1966) procedure involves the use of joint regression where the yield of each genotype is regressed on the environmental mean yield. The appropriate analysis of variance for the regression model is presented in Table 6.12. With this model the sum of squares due to environments and genotype x environments are partitioned into environments (linear), genotypes x environments (linear) and deviations from the regression model. The genotype's performance across environments is generally expressed in terms of three parameters, mean yield, the regression coefficient (b) and the deviations from the regression ($s^2\overline{d}$). A stable

genotype should have a high mean yield ($\bar{\chi}$), unit regression coefficient ($b = 1$) and deviations from regression as small as possible (s^2d =0). Therefore, the definition of a stable genotype, according to Eberhart-Russell regression model, is the one with $b = 1$ and $s^2d = 0$. It is however, specifically the deviation from regression, s^2d that is used as a measure of genotype stability across environments. The result of analysis of variance (Table 6.13) according to Eberhart and Russell (1966) shows non-significant differences among the genotypes and genotype x environment interaction.

In Table 6.12, the GE (linear) sum of squares was not a large portion of the GE interaction when compared with the environment (linear) sum of squares and residual sum of squares. Hence, only the deviation mean square was considered important. Genotypes with the highest s^2d in ranked order are PGRC/E 21169, PGRC/E 21258, (4DXZEM-1) X (ZEM-1-AD/88), PGRC/E 21163, and PGRC/E 20156. The three genotypes PGRC/E 20021 (b_i=1.2288), (4DXZEM-1) X (ZEM-1-AD/88) (b_i=1.0935), and PGRC/E 20112 (b_i=1.0886) had a regression coefficient greater than 1.0, which indicates instability. They are however, ranked 2nd, 4th, and 3rd in mean seed yield respectively. The best stability was shown by (4DXZEM-1) X (ZEM-1-F5/10), since its regression coefficient (b_i) was nearly unity and it had the smallest deviation from the regression line (Table 6.13). Becker and Léon (1988) indicated that the coefficient of regression could be used to describe the general response of genotypes to environmental conditions, while the deviations from regression measure the yield stability. So, genotypes (4DXZEM-1) X (ZEM-1-AD/88), PGRC/E 21258 and PGRC/E 21169 were found poorly adaptable as they significantly deviated from linearity (Table 6.13), and had high coefficients of variation.

Table 6.12. Analysis of variance for linear regression of cultivar means on environmental mean yield according to joint regression model.

Source	DF	SS	MS	F-value	Pr>F
Total	503	49802834.49			
Genotypes	13	735559.60	56581.51	1.46	0.1397
E in GxE	154	49067274.89	318618.67		
E (linear)	1	42961657.60			
GxE (linear)	13	682025.64	52463.51	1.35	0.1893
Pooled deviation	140	5423591.65	38739.94		
Residual	336	10177165.63	30289.18		

G=Genotypes; E=Environment

Table 6.13. Mean yield (kg/ha), b_i, s^2d_i, r_i^2, and CV for 14 genotypes evaluated across 12 environments, north-western Ethiopia, 1997-1999.

No	Cultivars	Mean	R	b_i	s^2d_i	R	r_i^2	CV%
1	PGRC/E 21163	1541	1	0.9584	22134	9	0.995	38.1
2	PGRC/E 21169	1357	10	1.0838*	41507	14	1.000	46.5
3	PGRC/E 21184	1388	6	0.7428	-9652	4	0.995	35.7
4	PGRC/E 21258	1360	9	1.0362*	35413	13	0.978	45.5
5	PGRC/E 20017	1363	8	0.8614	-23330	10	1.000	37.2
6	PGRC/E 20021	1527	2	1.2288	-14773	5	0.996	45.6
7	PGRC/E 20112	1431	3	1.0886	-15700	7	0.997	44.2
8	PGRC/E 20156	1403	5	1.0422	23204	11	0.997	45.4
9	PGRC/E 210406	1304	14	0.7760	15650	6	0.997	40.1
10	PGRC/E 20165	1370	7	1.0397	-8743	3	0.992	43.8
11	M-1	1347	12	0.9938	17407	8	0.983	46.8
12	(4DXZEM-1) X (ZEM-1-AD/88)	1427	4	1.0935*	32677	12	1.000	46.0
13	(4DXZEM-1) X (ZEM-1-F5/10)	1339	13	0.9901	449	1	0.964	43.5
14	Yellow Dodolla	1349	11	1.0648	2068	2	0.932	46.9
	R^2=0.8895 C.V. = 21 %							

* =Significant difference from 1 for b_i

8. Nassar and Hühn's variance of ranks

Nassar and Hühn's (1987) non-parametric measures of stability based on ranks provide a viable alternative to existing parametric measures. Table 6.14 presents the non-parametric measure for seed yield of 14 genotypes of *B. carinata* in 12 environments of north-western Ethiopia.

This non-parametric test is based on ranks of the genotypes across the locations. This gives equal weight to each location or environment. Genotypes with less change in rank are expected to be more stable. The mean absolute rank difference (S1) estimates are all possible pair wise rank differences across locations for each genotype. The S1 estimates are simply the variances of the ranks for each genotype across locations (Nassar and Hühn, 1987; Hühn, 1990). For S1, entries may be tested as significantly less stable or more stable than the average stability/instability. For the variance of ranks (S2), smaller estimates may indicate relative stability. Often, S2 has less power for detecting stability than S1. The S1 may lose power when genotypes are similar in their interactions with the environments.

For several reasons and for practical applications the use of S1 is preferred, because this stability parameter is very easy to compute and allows a clear and relevant interpretation (mean absolute rank difference between the environments). Furthermore, an efficient test of significance is available (Hühn, 1990).

The two overall chi-square calculated stabilities (Z1=19.56 and Z2=15.64) were less than the tabulated chi-square values ($\chi^2_{0.05, 14}$=23.68); hence, there was no evidence for significant differences in stability among the 14 genotypes (Table 6.14).

Table 6.14. Mean absolute rank difference (S1) and variance of ranks (S2) for seed yield (kg/ha) of 14 *B. carinata* genotypes tested in 12 environments of north-western Ethiopia, 1997-1999.

No	Cultivars	Mean yield	Rank	S1	Z1	S2	Z2
1	PGRC/E 21163	1541	1	5.212	0.665	17.722	0.102
2	PGRC/E 21169	1357	10	4.758	0.027	14.889	0.087
3	PGRC/E 21184	1388	6	5.288	0.854	19.076	0.374
4	PGRC/E 21258	1360	9	5.697	2.280	20.917	1.021
5	PGRC/E 20017	1363	8	3.182	4.380	7.00	4.011
6	PGRC/E 20021	1527	2	4.394	0.127	13.333	0.399
7	PGRC/E 20112	1431	3	3.348	3.437	8.243	3.005
8	PGRC/E 20156	1403	5	4.667	0.001	16.722	0.010
9	PGRC/E 210406	1304	14	5.894	3.211	23.076	2.184
10	PGRC/E 20165	1370	7	4.545	0.019	14.917	0.083
11	M-1	1347	12	5.182	0.596	17.556	0.080
12	(4DXZEM-1) X (ZEM-1-AD/88)	1427	4	5.955	3.530	23.854	2.710
13	(4DXZEM-1) X (ZEM-1-F5/10)	1339	13	4.348	0.178	12.243	0.753
14	Yellow Dodolla	1349	11	4.288	0.259	12.076	0.816

The overall chi-square for stability = 19.5637, 14 df individual Z(1) distributed as single df chi-squares; Overall chi-square for stability = 15.6354, 14 df individual Z(2) distributed as single df chi-squares.

Tabulated $\chi^2_{0.05}$, 14=23.68

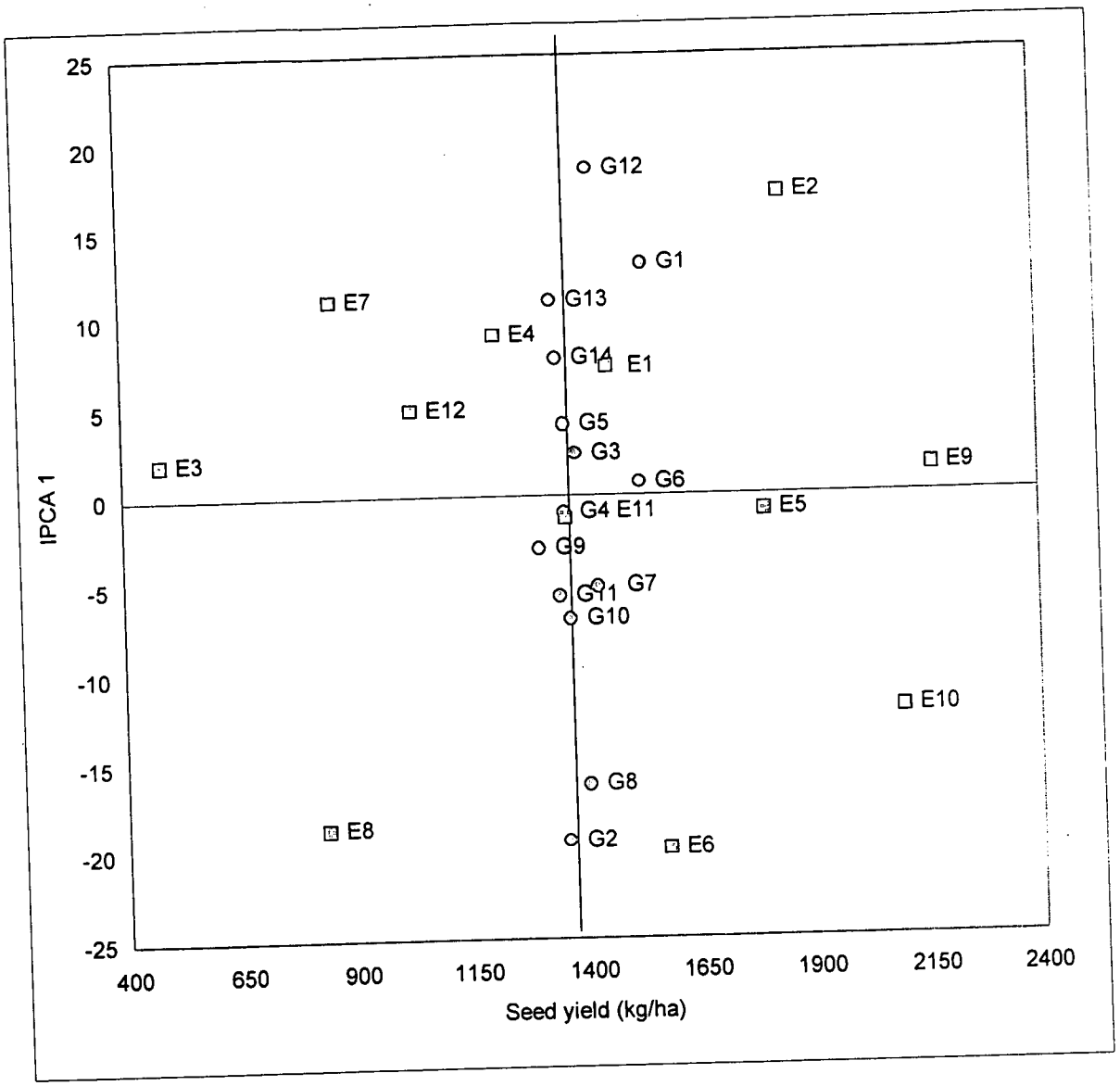
9. Additive Main Effects and Multiplicative Interactions (AMMI)

The GE interaction was partitioned into 11 Interaction Principal Component Axis (IPCA) using AMMI analysis. The IPCA are ordered according to decreasing importance (Table 6.15). The F-test was highly significant for the first 3 IPCA axes and significant for the 4^{th} IPCA axis at 1 and 5% probability levels respectively. All together the first four out of the 11 IPCA axes explained 83.23% of the GE interactions. The first IPCA captured 35.7% of the total interaction sum of squares in 6.1% of the interaction degrees of freedom. The second IPCA also explained 18.33% of the interaction sum of squares in 15% of the interaction degrees of freedom.

Partitioning of the interaction SS by AMMI is quite effective. The mean square for the first IPCA axis was 16X higher than the mean square of the residual. Although the first four axes were taken as adequate dimensions for the data, only the two IPCA axes were plotted against one other to help investigating the GE interactions pattern of each genotype.

Both the genotype and environment were plotted on the same graph to show associations between genotypes and environments (Figures 6.3 and 6.4). Genotypes and/or environments that appear almost on a horizontal line have a similar interaction patterns for the IPCA 2, and those that fall along the perpendicular line or close to it have similar interaction pattern along IPCA 1. The sign of the interaction is associated with the location of the quadrant. The bottom left of the quadrant is associated with a negative interaction along the axes and the top right of the quadrant with a positive interaction. The other two quadrants have different signs of interactions for both axes.

The IPCA scores of a genotype in the AMMI analysis are an indication of the stability of a genotype over environments. The greater the IPCA scores, either negative or positive as it is a relative value, the more specifically adapted a genotype is to certain environments. The more IPCA scores approximate to zero (0), the more stable the genotype is over all environments sampled.



Note: G1=PGRC/E 21163; G2=PGRC/E 21169; G3=PGRC/E 21184; G4=PGRC/E 21258; G5=PGRC/E 20017; G6=PGRC/E 20021; G7=PGRC/E 20112; G8=PGRC/E 20156; G9=PGRC/E 210406; G10=PGRC/E 20165; G11=M-1; G12=(4dxZEM-1) x (ZEM-1-AD/88); G13=(4DxZEM-1) x (ZEM-1-F5/10); G14=Yellow Dodolla.

E1=Adet 1997; E2=Mota 1997; E3=D/Tabor 1997; E4=F/Selam 1997; E5=Adet 1998; E6=Mota 1998; E7=D/Tabor 1998; E8=F/Selam 1998; E9=Adet 1999; E10=Mota 1999; E11=D/Tabor 1999; E12=F/Selam 1999

Figure 6.3. AMMI-1 model for seed yield (kg/ha) showing means of genotypes (○) and environments (▣) plotted against their respective scores of the first interaction principal component (IPCA-1).

Since the contribution of IPCA 1 is larger than IPCA 2 in explaining interactions, genotypes with considerable absolute scores along IPCA 1 have stronger interaction than that along IPCA 2. As a result genotypes PGRC/E 21169, M-1, (4DXZEM-1) X (ZEM-1-AD/88) and PGRC/E 210406 have large interactions along both axes.

Genotypes PGRC/E 21169 and (4DXZEM-1) X (ZEM-1-AD/88) also have large interactions along IPCA 1 and are unstable.

In the AMMI 2 biplot graph (Figure 6.4), points near the origin (PGRC/E 20021) had little interaction. These points closely fitted the additive part of the model. Points near each other (PGRC/E 21258, PGRC/E 20021) were similar and had similar interaction pattern. Conversely, points distant from each other (PGRC/E 21169, (4DXZEM-1) X (ZEM-1-AD/88) are dissimilar and had completely different values for the interaction effects.

Based on the AMMI stability value (ASV) (Table 6.16) the most stable genotypes are PGRC/E 21258, PGRC/E 20021, and PGRC/E 20017 in that order.

AMMI biplot IPCA 1 vs IPCA 2 for seed yield (kg/ha) for genotypes and environments (Figure 6.4) shows stable genotypes (PGRC/E 21258, PGRC/E 20021, and PGRC/E 20112) clustered close to the centre of the biplot. The unstable genotypes such as PGRC/E 21169, M-1, and (4DXZEM-1) X (ZEM-1-AD/88) were far from the centre. Similarly, environments like E3, E9, and E12 are more stable than E6, E8, and E2. This figure also illustrates the dominant genotypes and environments with negative IPCA 1 scores that strongly influence the GE interaction such as (4DXZEM-1) X (ZEM-1-AD/88) and E6, and PGRC/E 20112 and E8.

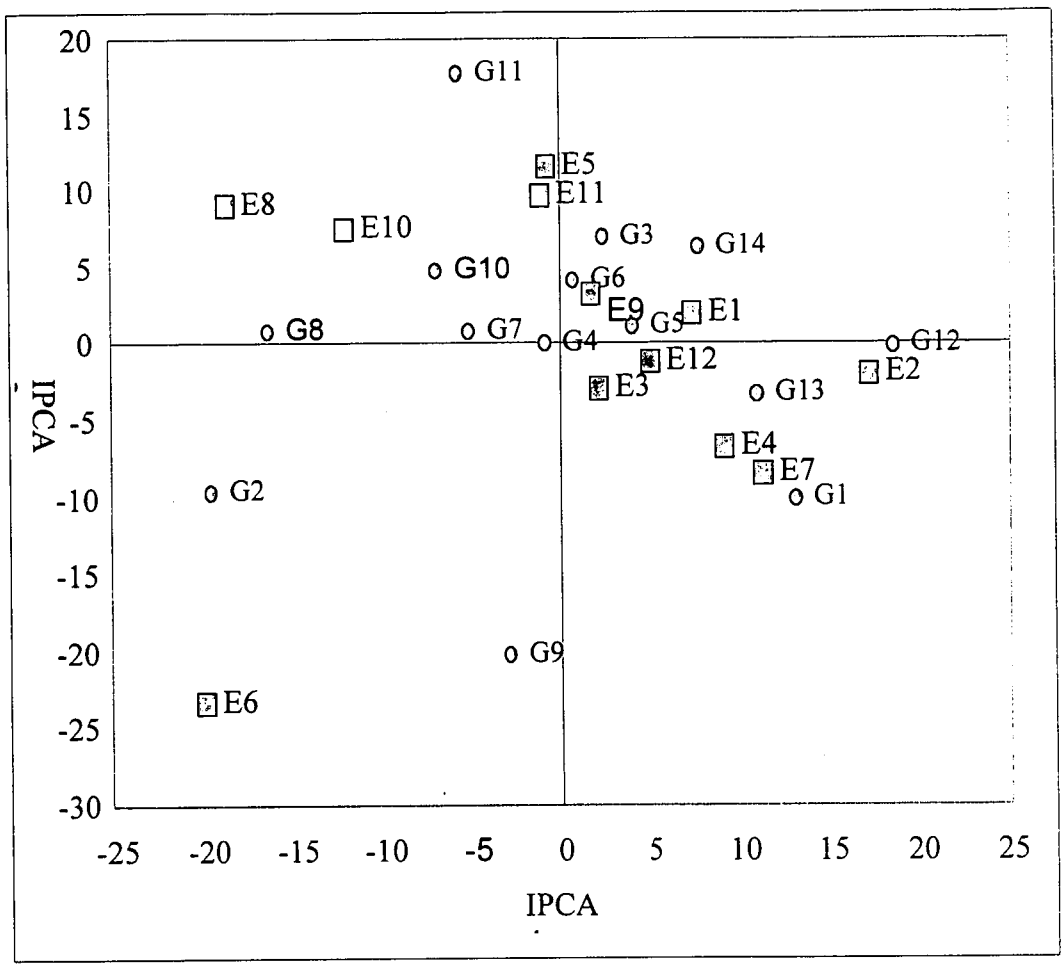
Table 6.15. Additive main effects and multiplicative interaction analysis of variance for seed yield (kg/ha) for 12 environments including the first four interaction principal component analysis (IPCA) axes.

Source	DF	SS	MS	F-value	Pr>F
Total	503	179940000			
Environments	11	128884972	11716815	49.58	0.0000
Reps within Env.	24	5671364	236306		
Genotype	13	2206678	169744	1.33	0.1965
Genotype x Env.	143	18316851	128089	1.61	0.0003**
IPCA 1	23	6539291	284317	3.57	0.0000**
IPCA 2	21	3356721	159843	2.01	0.0061**
IPCA 3	19	2972852	156465	1.96	0.0100*
IPCA 4	17	2376324	139783	1.75	0.0331*
IPCA Residual (Pooled error)	63	3071661.9	48756.5		
Residual	312	24860132	79679		

*,**=Significant at 0.05 and 0.01 probability levels respectively.

Table 6.16. Mean seed yield (kg/ha), rank, IPCA1 and IPCA 2 scores and an AMMI stability value (ASV) of 14 genotypes tested in 12 environments in north-western Ethiopia, 1997-1999.

No	Cultivars	Mean	R	IPCA1score	IPCA2 score	ASV	Rank
1	PGRC/E 21163	1541	1	13.079	-10.044	27.43	11
2	PGRC/E 21169	1357	10	-19.494	-9.630	39.21	14
3	PGRC/E 21184	1388	6	2.350	6.964	8.34	4
4	PGRC/E 21258	1360	9	-0.947	-0.009	1.85	1
5	PGRC/E 20017	1363	8	4.002	1.104	7.88	3
6	PGRC/E 20021	1527	2	0.686	4.225	4.43	2
7	PGRC/E 20112	1431	3	-5.115	0.873	10.01	5
8	PGRC/E 20156	1403	5	-16.259	0.909	31.72	12
9	PGRC/E 210406	1304	14	-2.946	-20.132	20.94	9
10	PGRC/E 20165	1370	7	-6.924	4.839	14.34	6
11	M-1	1347	12	-5.648	17.723	20.87	8
12	(4DXZEM-1) X (ZEM-1-AD/88)	1427	4	18.534	-0.035	36.14	13
13	(4DXZEM-1) X (ZEM-1-F5/10)	1339	13	10.978	-3.241	21.65	10
14	Yellow Dodolla	1349	11	7.703	6.456	16.35	7



Note: G1=PGRC/E 21163; G2=PGRC/E 21169; G3=PGRC/E 21184; G4=PGRC/E 21258; G5=PGRC/E 20017; G6=PGRC/E 20021; G7=PGRC/E 20112; G8=PGRC/E 20156; G9=PGRC/E 210406; G10=PGRC/E 20165; G11=M-1; G12=(4DxZEM-1) x (ZEM-1-AD/88); G13=(4DxZEM-1) x (ZEM-1-F5/10); G14=Yellow Dodolla.

E1=Adet 1997; E2=Mota 1997; E3=D/Tabor 1997; E4=F/Selam 1997; E5=Adet 1998; E6=Mota 1998; E7=D/Tabor 1998; E8=F/Selam 1998; E9=Adet 1999; E10=Mota 1999; E11=D/Tabor 1999; E12=F/Selam 1999

Figure 6.4. AMMI-2 model for seed yield (kg/ha) showing the plotting of respective scores of the interaction principal component analysis (IPCA) of genotypes (o) tested across environments (□).

10. Comparison of stability analyses

Table 6.17 shows ranking orders for the stability of 14 *B. carinata* genotypes evaluated across 12 environments in north-western Ethiopia according to the ten stability analysis procedures. The parameter of Shukla's (1972) stability variance, Eberhart and Russell's (1966) deviation parameter, Wricke's (1962) ecovalence, and Nassar and Hühn (1987) variance of ranks unanimously ranked PGRC/E 20017 as the most stable genotype. It was ranked 3rd by the AMMI stability value (ASV). The second stable genotype PGRC/E 20112 was also ranked in harmony by Shukla's (1972) stability variance, Eberhart and Russell's (1966) deviation parameter, Wricke's (1962) ecovalence, and Nassar and Hühn (1987) variance of ranks, whereas, the AMMI stability value placed it in the 5th rank. The over all ranks of the stability statistics shows that PGRC/E 20017 and PGRC/E 20112 were the most stable genotypes.

Spearman's coefficient of rank correlation was then determined for each of the possible pair wise comparisons of the different stability statistics (Table 6.18). Finlay & Wilkinson's procedure shows significant positive rank correlations with CV_i ($P<0.01$) and non-significant positive rank correlations with ASV, Shukla's procedure, W_i and s^2d_i. Whereas, negative rank correlation was observed with r_i^2, P_i, S1, and with the mean. This negative rank correlation suggests that the genotypes that were highly responsive to high yielding environments were less responsive to low yielding environments, and vice versa (Jalaluddin and Harrison, 1993). Brown *et al.* (1983) observed in rice (*Oryza sativa* L.) that regression responses are expressed more clearly (with high r_i^2) in high yielding environments than in low yielding environments. Finlay & Wilkinson's procedure defines stability as the sensitivity of a genotype to changing environments, and this is measured and reflected by the regression coefficient (b_i) of joint analysis. This definition is similar to the static concepts as defined by Becker and Léon (1988), as well as to Type I stability as defined by Lin *et al.* (1986).

Eberhart and Russell procedure shows highly significant correspondence ($P<0.01$) with the procedure of Shukla (r=0.9209**), Wricke (r=0.9209**), S1 (r=0.6703**)

and ASV (0.6000**), but to a lesser degree with CV_i (%) and with the Mean. It has also a negative rank correlation with coefficient of determination, r_i^2. Their definition of stability is based on a genotype's average sensitivity to environmental fluctuations and is determined by using joint linear regression analysis in which the average deviation from the regression, or response to environments, is determined. The Eberhart and Russell's definition of a stable genotype is one unit of regression coefficient (b_i =1.0) and deviations from regression as small as possible (s^2d_i=0). Lin and Binn's procedure (P_i) shows the highest positive rank correlations (P<0.01) with mean seed yield (r=0.8945**), and fairly low rank correlations with r_i^2, S1, and Shukla (σ^2_i). It also shows the negative correlations with s^2d_i (r=-0.4769). The highly significant rank correlation of seed yield with P_i indicates that selection for yield would change yield stability by increasing P_i leading to development of genotypes that are specially adapted to environments with optimal growing conditions. The Lin and Binn procedure defines stability as the deviations of a specific genotype's performance from the performance of the best cultivar. In other words, a stable cultivar is the one that performs in tandem with the environment. A genotype with an inherent high yield would be classified as stable if its yield over locations is simillar to that of the top performer, over the respective locations. Thus, the Lin and Binn procedure appears to be more of a genotype performance measure, rather than a stability measure over sites.

The Wricke's procedure of stability statistic shows the highest significant positive correlation (P<0.01) with Shukla (1.0000**) and s^2d_i (0.9209**). According to Table 6.18, the rank correlation coefficient of 1.0 was found between Shukla's and Wricke's procedures. This indicates that the two procedures are equivalent for ranking purposes. Shukla's stability variance is a linear combination of the ecovalence and therefore, both W_i and σ^2_i are equivalent for ranking purposes (Wricke and Weber, 1980). Purchase (1997) has also reported similar correspondence between the two stability parameters. This relationship is expected since these stability parameters have biometrical relationships between coefficient of determination, r_i^2 and deviations from regression, s^2d_i. Negative and non-significant correlations were

observed between coefficient of determination, r_i^2 and deviations from regression, $s^2 d_i$. Becker (1981) reported consistent and strong correlations between environmental variance and the regression coefficient on one hand and ecovalence and deviation mean square on the other.

Table 6.17. Mean yield (kg/ha) and various stability measurements and their ranking (R) orders of 14 *B. carinata* genotypes evaluated across 12 environments in north-western Ethiopia.

No	Genotypes	$\bar{x}$	R	CV_i (%)	R	$\sigma^2{}_i$	R	b_i	R	s^2d_i	R	r_i^2	R	W_i	R	P_i	R	S1	R	ASV	R	Overall Rank
1	PGRC/E 21163	1541	1	38.1	3	157819.55	9	0.9584	4	22134	11	0.995	8.5*	529551.64	9	50237.60	2	5.212	10	27.43	11	6
2	PGRC/E 21169	1357	10	46.5	12	224620.58	14	1.0838	11	41507	14	1.000	2	739497.74	14	105278.66	9	4.758	8	39.21	14	14
3	PGRC/E 21184	1388	6	35.7	1	119599.23	7	0.7428	1	-9652	4	0.995	8.5	409430.63	7	110743.32	11	5.288	11	8.34	4	4
4	PGRC/E 21258	1360	9	45.5	9	199658.33	13	1.0362	7	35413	13	0.978	12	661044.96	13	104101.35	8	5.697	12	1.85	1	11.5
5	PGRC/E 20017	1363	8	37.2	2	30219.01	1	0.8614	3	-23330	1	1.000	2	128521.38	1	96044.31	6	3.182	1	7.88	3	1
6	PGRC/E 20021	1527	2	45.6	10	89795.15	5	1.2288	14	-14773	3	0.996	7	315760.67	5	42884.11	1	4.394	5	4.43	2	3
7	PGRC/E 20112	1431	3	44.2	7	43413.04	2	1.0886	12	-15700	2	0.997	5	169988.33	2	61830.55	3	3.348	2	10.01	5	2
8	PGRC/E 20156	1403	5	45.4	8	161271.99	10	1.0422	9	23204	10	0.997	5	540402.16	10	86055.47	4	4.667	7	31.72	12	8
9	PGRC/E 210406	1304	14	40.1	4	184496.41	11	0.7760	2	15650	8	0.997	5	613393.20	11	129117.58	14	5.894	13	20.94	9	9.5
10	PGRC/E 20165	1370	7	43.8	6	59416.11	3	1.0397	8	-8743	5	0.992	10	220283.69	3	96235.80	7	4.545	6	14.34	6	5
11	M-1	1347	12	46.8	13	141124.90	8	0.9938	6	17407	9	0.983	11	477082.75	8	117541.69	12	5.182	9	20.87	8	11.5
12	(4DXZEM-1)X(ZEM-1-AD/88)	1427	4	46.0	11	198212.40	12	1.0935	13	32677	12	1.000	2	656500.60	12	95400.34	5	5.955	14	36.14	13	13
13	(4DXZEM-1)X (ZEM-1-F5/10)	1339	13	43.5	5	87226.28	4	0.9901	5	449	6	0.964	13	307687.07	4	110236.60	10	4.348	4	21.65	10	7
14	Yellow Dodolla	1349	11	46.9	14	96385.27	6	1.0648	10	2068	7	0.932	14	336472.48	6	118110.74	13	4.288	3	16.35	7	9.5

Note: R=ranks; $\bar{x}$ =Mean yield; *CV (%)*=Francis and Kannenberg's (1978) Coefficient of variability; $\sigma^2{}_i$ = Shukla's (1972) Stability variance; b_i =Finlay & Wilkinson's (1963) regression coefficient; s^2d_i =Eberhart & Russell's (1966) deviation parameter; r_i^2 =Coefficient of determination; W_i = Wricke's (1962) ecovalence; P_i =Lin and Binn's (1988) cultivar superiority performance; S1=Nassar & Hühn (1987) variance of ranks; ASV=Purchase (1997) AMMI stability value.

* For Spearman's rank correlation coefficients equal rank values are averaged.

Table 6.18. Spearman's rank correlation coefficients among mean yield and other stability statistic indices across all environments and years

	$\bar{x}$	CV_i (%)	$\sigma^2{}_i$	b_i	$s^2 d_i$	r_i^2	W_i	P_i	S1
CV_i (%)	0.1604								
$\sigma^2{}_i$	0.1385	0.3495							
b_i	-0.3890	0.7231**	0.0769						
$s^2 d_i$	0.1429	0.4374	0.9209**	0.1341					
r_i^2	0.3241	0.1487	-0.1665	-0.1909	-0.0178				
W_i	0.1385	0.3495	1.0000**	0.0769	0.9209**	-0.1665			
P_i	0.8945**	0.1121	0.1912	-0.4769	0.1297	0.3774	0.1912		
S1	0.0374	0.0022	0.8022**	-0.1956	0.6703**	-0.0932	0.8022**	0.2088	
ASV	0.0989	0.2220	0.4857	0.1297	0.6000**	-0.3019	0.4857	0.0681	0.2088

*,**=Significant at p=0.05 and 0.01 probability levels respectively.

11. Cluster analysis of genotypes and environments

Cluster analysis is the most widely used technique for classifying environments or genotypes into homogeneous groups. It operates on a matrix of dissimilarity (or similarity) indexes for all possible pairs of genotypes or pairs of environments, depending on which is being clustered (Ghaderi *et al.*, 1980)

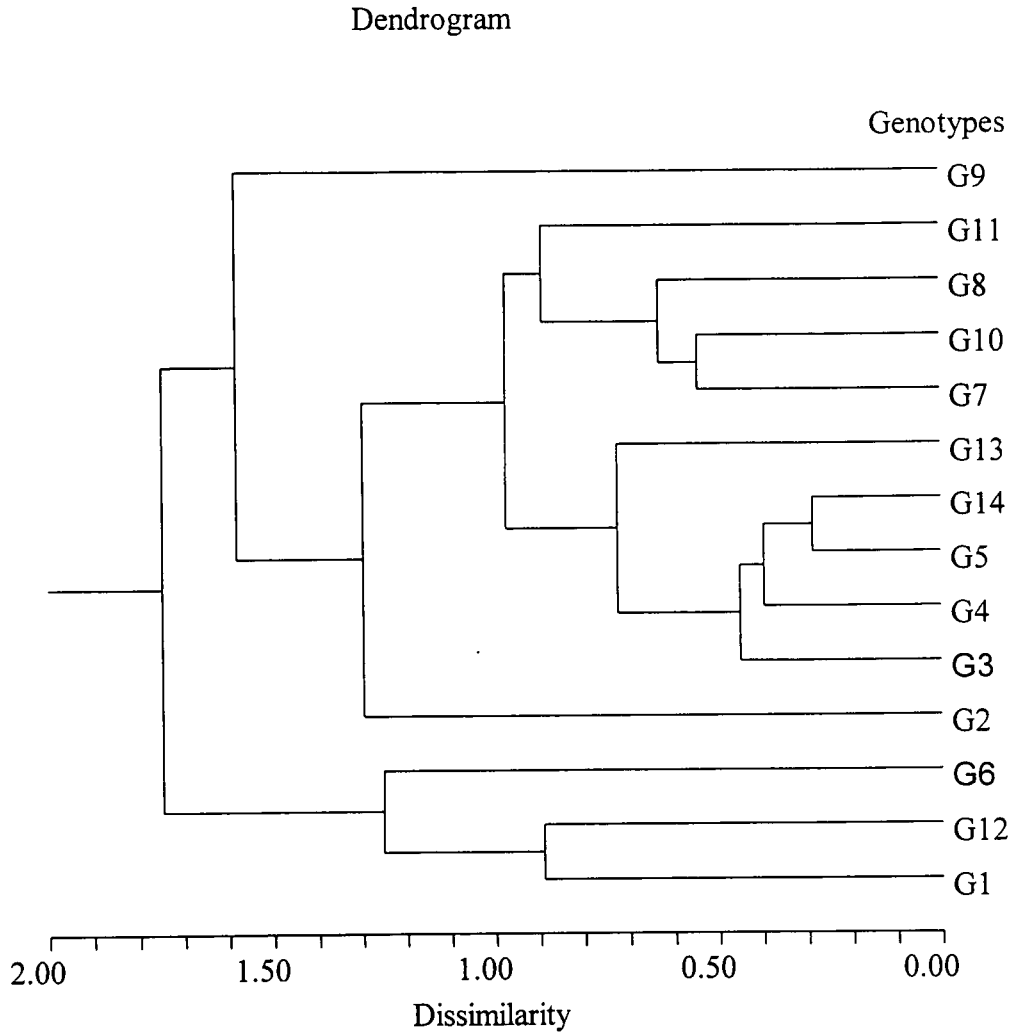
Cluster analysis was performed to study the patterns of groupings of genotypes and environments. The dendrograms (Figures 6.5 & 6.6) were generated from the UPGMA clustering method of genotypes and environments based on Euclidean distances using AMMI adjusted mean yields of environment and genotypes respectively. Clustering of genotypes at a cut-off value of 1.0 produced five clusters of which three are singletons (Figure 6.5). Cluster one consists of PGRC/E 21163 and (4DXZEM-1) X (ZEM-1-AD/88). Cluster two consists of nine genotypes (PGRC/E 21184, PGRC/E 21258, PGRC/E 20017, PGRC/E 20112, PGRC/E 20156, PGRC/E 20165, M-1, (4DXZEM-1) X (ZEM-1-F5/10), and Yellow Dodolla). Cluster three; four and five consists of G2, G6 and G9 respectively.

The two genotypes in cluster one (PGRC/E 21163) and (4DxZEM-1) x (ZEM-1-AD/88) were characterized with high yield. This was also clearly depicted from the biplot of yield means and AMMI-1 (Figure 6.3). Cluster two consists of PGRC/E 21184, PGRC/E 21258, PGRC/E 20017, PGRC/E 20112, PGRC/E 20156, PGRC/E 20165, (4DXZEM-1) X (ZEM-1-F5/10) and Yellow Dodolla. Genotypes in cluster two are grouped around the positive and negative perpendicular axis of IPCA 1 (Figure 6.3).

Cluster analysis groups genotypes according to their phenotypic similarity and stability. The first cluster contains genotypes PGRC/E 21163 and (4DXZEM-1) X (ZEM-1-AD/88) with high yield, high coefficient of variability, similar s^2d_i and ASV.

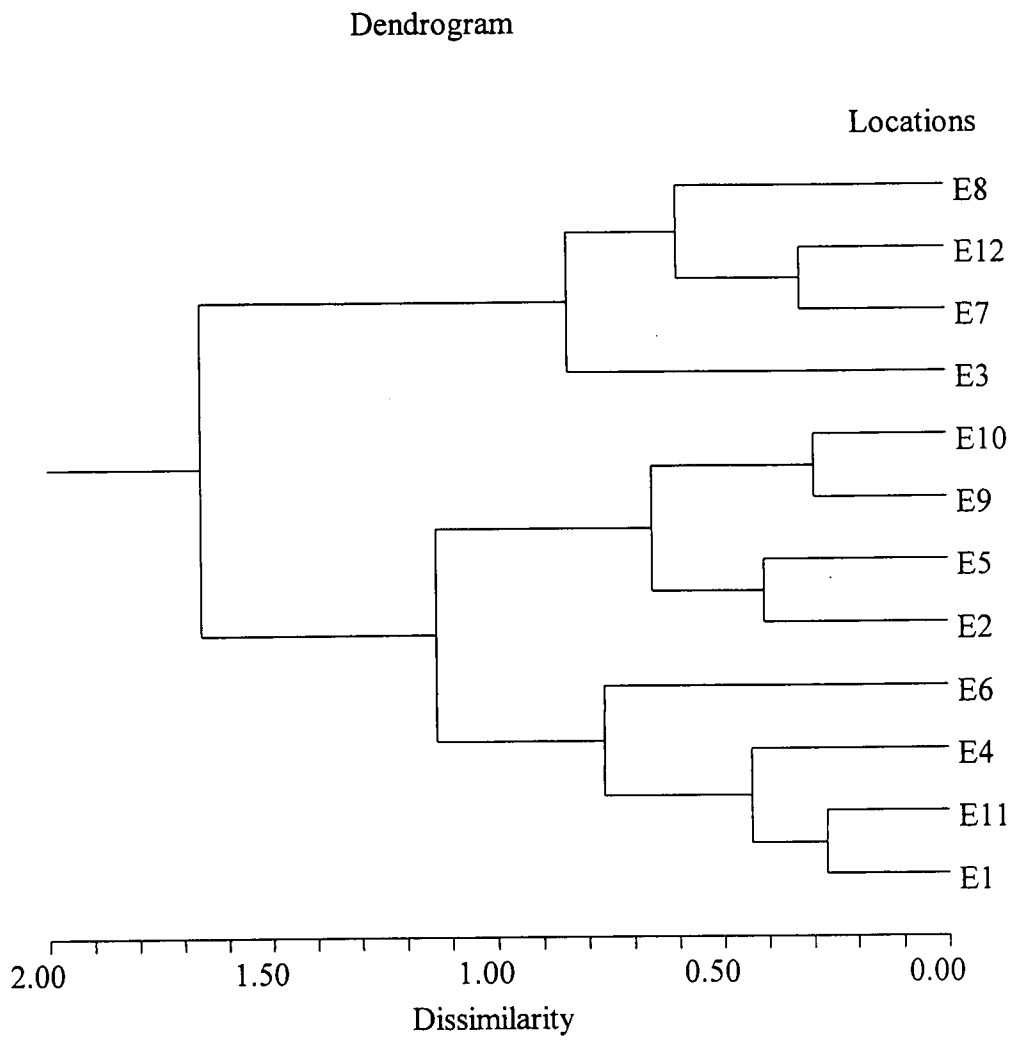
Cluster analysis of environments at cut-off point 1.0 produced three clusters each containing four environments. Cluster one consists of E1, E5, E9, and E10. Cluster

two E1, E4, E6, and E11. Cluster three E3, E7, E8, and E12.



Note: G1=PGRC/E 21163; G2=PGRC/E 21169; G3=PGRC/E 21184; G4=PGRC/E 21258; G5=PGRC/E 20017; G6=PGRC/E 20021; G7=PGRC/E 20112; G8=PGRC/E 20156; G9=PGRC/E 210406; G10=PGRC/E 20165; G11=M-1; G12=(4dxZEM-1) x (ZEM-1-AD/88); G13=(4DxZEM-1) x (ZEM-1-F5/10); G14=Yellow Dodolla.

Figure 6.5. Dendrogram depicting the clustering of 14 genotypes using AMMI adjusted means of 12 environments.



Note: E1=Adet 1997; E2=Mota 1997; E3=D/Tabor 1997; E4=F/Selam 1997; E5=Adet 1998; E6=Mota 1998; E7=D/Tabor 1998; E8=F/Selam 1998; E9=Adet 1999; E10=Mota 1999; E11=D/Tabor 1999; E12=F/Selam 1999

Figure 6.6. Dendrogram depicting the clustering of 12 environments using AMMI adjusted means of 14 genotypes.

6.5. REFERENCES

ADUGNA, W., and M.T. LABUSCHAGNE, 2002. Genotype-environment interactions and phenotypic stability analyses of linseed in Ethiopia. *Plant Breeding* 121: 66-71.

AGROBASE^{TM} , 2000. Agronomic Software Inc., 71 Waterloo St. Winnipeg, Manitoba, R3N0S4, Canada.

ALLARD, R.W., 1960. Principles of plant breeding. 1st ed. John Wiley & Sons, Inc., N.Y.

ALLARD, R.W., and A.D. BRADSHAW, 1964. Implication of genotype-environment interactions in plant breeding. *Crop Sci.* 4: 503-508.

ANNICCHIARICO, P., 1997a. Joint regression vs AMMI analysis of genotype-environment interactions for cereals in Italy. *Euphytica* 94: 53-62.

ANNICCHIARICO, P., 1997b. Additive main effects and multiplicative interaction (AMMI) analysis of genotype-location interaction in variety trials repeated over years. *Theor. Appl. Genet.* 94: 1072-1077.

BASFORD, K.E., and M. COOPER, 1998. Genotype x environmental interactions and some considerations of their implications for wheat breeding in Australia. *Aust. J. Agric. Res.* 49: 154-174.

BECKER, H.C., 1981. Correlations among some statistical measures of phenotypic stability. *Euphytica* 30: 835-840.

BECKER, H.C., and J. LÉON, 1988. Stability analysis in plant breeding. *Plant Breeding* 101: 1-23.

BROWN, K.D., M.E. SORRELLS, and W.R. COFFMAN, 1983. A method for

classification and evaluation of testing environments. *Crop Sci.* 23: 889-893.

COMSTOCK, R.E., and R.H. MOLL, 1963. Genotype-environment interactions. *Symposium on Statistical Genetics and Plant Breeding*. NAS-NRC Pub. 982. 164-196. W.D. Hanson and H.F. Robinson (eds.). Washington.

CROSSA, J., 1990. Statistical analyses of multilocation trials. *Advances in Agronomy* 44: 55-85.

EBERHART, S.A., and W.A. RUSSELL, 1966. Stability parameters for comparing varieties. *Crop Sci.* 6: 36-40.

FINLAY, K.W., and G.N. WILKINSON, 1963. The analysis of adaptation in plant breeding programme. *Aust. J. Agric. Res.* 14: 742-754.

FOX, P.N., and A.A. ROSIELLE, 1982. Reducing the influence of environmental main-effects on pattern analysis of plant breeding environments. *Euphytica* 31: 645-656.

FRANCIS, T.R., 1977. Yield stability studies in short-season maize (*Zea mays* L.). Ph.D. Thesis, University of Guelph, Guelph, Ont.

FRANCIS, T.R., and L.W. KANNENBERG, 1978. Yield stability studies in short-season maize. I. A descriptive method for grouping genotypes. *Can. J. Plant Sci.* 58: 1029-1034.

GAUCH, H.G. 1988. Model selection and validation for yield trials with interaction. *Biometrics* 44: 705-715.

GAUCH, H.G., 1982b. Multivariate analysis in community ecology, 1st ed. Cambridge Univ. Press, London and New York.

GAUCH, H.G., and R.W. ZOBEL, 1996. AMMI analysis of yield trials. *In*: Genotype

by environment interaction. M. S. Kang and H.G. Gauch, Jr. (eds). pp. 85-122.

GELETU, B., S. TESGAYA and A. TULU, 1995. Performance and stability of seed yield of lentil varieties (*Lens culinaris* Medik) grown in the Ethiopian highlands. *Crop Res.* 9(3): 337-345.

GHADERI, A., E.H. EVERSON, and C.E. CRESS, 1980. Classification of environments and genotypes in wheat. *Crop Sci.* 20: 707-710.

HILL, L., 1975. Genotype-environment interactions-a challenge for plant breeding. *J. Agric. Sci.* 85: 477-493.

HINTZE, J.L., 2000. NCSS 2000 Statistical System for Windows. Number Cruncher Statistical Systems. Kaysville, Utah.

HOHLS, T., 1995. Analysis of genotype environment interactions. *S. Afr. J. Sci.* 91: 121-124.

HÜHN, M., 1990. Nonparametric measure of phenotypic stability. Part 2: Applications. *Euphytica* 47: 189-194.

JALALUDDIN, M.D., and S.A. HARRISSON, 1993. Repeatability of stability estimators for grain yield in wheat. *Crop Sci.* 33: 720-725.

LIN, C.S. and M.R. BINNS, 1988a. A superiority measure of cultivar performance for cultivar x location data. *Can. J. Plant Sci.* 68: 193-198.

LIN, C.S., and M.R. BINNS, 1988b. A method of analysing cultivar x location x year experiments: A new stability parameter. *Theor. Appl. Genet.* 76: 425-430.

LIN, C.S., M.R., BINNS, and L.P. LEFKOVICH, 1986. Stability analysis: where do we stand? *Crop Sci.* 26: 894-900.

McINTOSH, M.S., 1983. Analysis of combined experiments. *Agron. J.* 75: 153-155.

MICHIGAN STATE UNIVERSITY, 1991. MSTAT-C, a software program for the design, management, and analysis of agronomic research experiments. Michigan State University, East Lansing, USA.

NASSAR, R., and M. HÜHN, 1987. Studies on estimation of phenotypic stability: Test of significance for nonparametric measures of phenotypic stability. *Biometrics* 43: 45-53.

PINTHUS, M.J., 1973. Estimates of genotypic value: A proposed method. *Euphytica* 22: 121-123.

PURCHASE, J.L., 1997. Parametric analysis to describe Genotype x environment interaction and yield stability in winter wheat. Ph.D. Thesis, Department of Agronomy, Faculty of Agriculture of the Orange Free State, Bloemfontein, South Africa.

RAO, P.S.R.S., 1997. Variance components estimation: mixed models, methodologies and applications. Chapman & Hall, London.

ROMAGOSA, I., and P. N. FOX, 1993. Genotype x environment interaction and adaptation. *In*: M.D. Hayward, N.O. Bosemark & I. ROMAGOSA (Eds.), Plant breeding: principles and prospects, Chapman and Hall, London, pp. 373-390

SAS Institute, 1996. SAS Institute, Cary, NC.

SHUKLA, G.K., 1972. Some statistical aspects of partitioning genotype-environmental components of variability. *Heredity* 29: 237-245.

SPRAGUE, G.F., and W. T. FEDRER, 1951. A comparison of variance components in corn yield trials. II. Error, year, year x variety, location x variety, and variety

components. *Agron. J.* 43: 535-541.

STEEL, R.G. and J.H. TORRIE, 1980. Principles and procedures of statistics. McGraw-Hill, New York.

TESFAYE, T., S. TSEGAYE, G. BELAY, E. BECHERE and D. MITIKU, 1998. Stability of performance of teteraploid wheat landraces in the Ethiopian highland. *Euphytica* 102: 301-308.

WESTCOTT, B., 1986. Some methods of analysing genotype x environment interaction. *Heredity* 56: 243-253.

WRICKE, G., 1962. Über eine Methode zur Erfassung der Ökologischen Streuberite in Feldversuchen. *Z. Pflanzenzüchtg.* 47: 92-96.

WRICKE, G., and W.E. WEBER, 1980. Erweiterte analyse von wechselwirkungen in versuchsserien. In: Biometrie-heute und morgen. Eds Köpcke & Überla. Springer-Verlag, Berlin.

YAN, W., and L.A. HUNT, 1998. Genotype-by-environment interaction and crop yield. *Plant Breed. Rev.* 16: 35-178.

YAU, S.K., 1995. Regression and AMMI analyses of genotype x environment interactions: an empirical comparison. *Agron. J.* 87: 121-126.

YUE, G.L., K.L. ROOZEBOOM, W.T. SCHAPAUGH, and G.H. LIANG, 1997. Evaluation of soybean cultivars using parametric and nonparametric stability estimates. Plant Breeding 116: 271-275.

ZOBEL, R.W., M.J. WRIGHT, and H.G. GAUCH, 1988. Statistical analysis of a yield trial. *Agron. J.* 80: 388-393.

CHAPTER VII
SUMMARY

Ethiopian or Abyssinian mustard (*Brassica carinata* A. Braun) is one of the most important oil crops in Ethiopia. Apart from its phylogenetic relationships with other members of the genus *Brassica*, little information is available for this species. This study explored the genetic diversity and genotype x environment interaction for its possible use in genetic conservation and breeding. Geographical patterns of morphological variations were studied on 258 accessions of *B. carinata* randomly sampled from different parts of the country using Shannon Weaver diversity index (H') to estimate the phenotypic diversity index for six morphological traits. The results revealed that the mean diversity index ($\bar{H}'$) across regions of the country was 0.72±0.06. Polymorphism was high for some of the traits. The mean diversity index ($\bar{H}'$) for the country across altitudinal classes was $\bar{H}' = 0.71 \pm 0.07$. Highest mean diversity index was recorded from areas ranging from 2600-280 masl. The centres of diversity appears to be Shewa, Bale and Arsi provinces with due consideration of altitudes. Further exploration and *in situ* conservation efforts should be undertaken in these regions.

Brassica carinata possesses a number of agronomic advantages over other oilseed crops with similar ecological adaptation. However, the oil is considered low quality. Although efforts have been made to improve its quality, much has to be done to use natural variations that might exist within the species for fatty acid composition. The objectives of this study were to determine the oil content and fatty acid composition to determine the range of genetic diversity for these traits, and to select genotypes that can be used in crosses to improve these traits. Ninety-eight Ethiopian grown genotypes were analyzed by CGC to study the variability of fatty acid composition. The analysis revealed a wide variation in fatty acid composition; twenty-six fatty acids were identified. The predominant fatty acids were erucic (6.91-35.05%), linoleic (17.21-28.53%), α-linolenic (10.50-22.52%), and oleic (8.91-24.64%), followed by gadoleic (1.22-10.24%) and palmitic (3.38-16.42%). Palmitic acid had significant positive correlations with stearic acid. Erucic acid showed significant and negative linear correlations with palmitic, stearic, vaccinic, oleic, linoleic, α-linolenic acids and

positive correlation with eicosenoic acid. The high oil yielding genotypes were relatively richer in erucic acid content. Oil content ranged from 25 to 48%.

Genetic diversity analysis within *B. carinata* has not been examined with molecular markers. The objective of this study was to investigate the genetic relationships among 39 *B. carinata* genotypes by the AFLP markers. This study explored the amount of genetic variation for possible use in breeding. Thirty-nine genotypes of *B. carinata* were analysed using six AFLP primer combinations. A total of 189 polymorphic markers were scored, with an average of 32 markers per primer combination. GDE ranged from 0.346 to 0.639, with a mean of 0.504 ± 0.002. Polymorphism rates ranged from 50 to 80%. UPGMA cluster analysis revealed seven genetically distinct groups of genotypes.

The GE interaction and the stability of 14 *B. carinata* genotypes for seed yield were evaluated at four rain-fed locations, over three years in north-western Ethiopia. The objectives were to estimate the component of variance associated with the first and second order interactions and to determine their effects and to measure the genotypes' stability using different stability statistics; and to compare the stability statistics.

The analysis of variance test revealed that the main effects due to years, locations, and the first order interactions (year x location) as well as second order interactions (genotype x year x location) were highly significant. The genotype x location interaction was not significant. This suggests the necessity for increased emphasis on multiple years of testing than more locations. The estimates of variance components involving first order interaction of genotype x year, genotype x location and the error variance were very small. The variance components of the main effects of the genotype and second order interaction of genotype x location x year were very high.

Mean yield over all locations and years ranged from 1304 to 1541 kg/ha by PGRC/E 210406 and PGRC/E 21163 respectively. The regression of yield on the environmental index resulted in a regression coefficient value ranging from 0.7488-1.2288. The deviation from regression was non significant for all genotypes. The first two interaction principal component axes of the additive main effects and multiplicative interaction (AMMI) model together accounted between 35.7% and

54.3% of the total genotype x environment interaction of sum of squares for seed yield. The stability analyses identified PGRC/E 20017 and PGRC/E 20112 as more stable genotypes, while PGRC/E 21169 and (4DxZEM-1) x (ZEM-1-AD/88) were specifically adapted to some environments.

Spearman's coefficient of rank correlation was performed for each of the possible pair wise comparisons of different stability statistics. Significant rank correlations were observed among some of the stability statistics. These significant rank correlations indicated their effectiveness in detecting stable genotypes over a range of environments. The hierarchical clustering of AMMI adjusted mean yield using UPGMA clustering method grouped the genotypes into five genetically distinct clusters and the environments into three heterogeneous clusters.

OPSOMMING

Etiopiese of Abisiniese mosterd (*Brassica carinata* A. Braun) is een van die belangrikste olie gewasse in Etiopië. Afgesien van die filogenetiese verwantskappe met ander lede van die genus, is daar min inligting beskikbaar oor hierdie spesie. Hierdie studie het genetiese diversiteit ondersoek en genotipe x omgewings interaksie vir moontlike gebruik in genetiese bewaring en teling. Geografiese patrone van morfologiese variasie is ondersoek vir 258 genotipes van *B. carrinata* wat toevallig gemonster is van verskillende dele van die land, met die gebruik van die Shannon Weaver diversiteits indeks ($\bar{H}'$) om die fenotipiese diversiteits indeks vir ses kwalitatiewe eienskappe te bepaal. Die resultate het aangetoon dat die gemiddelde genetiese diversiteits indeks ($\bar{H}'$) oor areas van die land 0.72±0.06 was. Polimorfisme was ook hoog vir sommige eienskappe. Die gemiddelde diversiteits indeks ($\bar{H}'$) vir die land oor hoogte klasse was $\bar{H}'$=0.71±0.07. Die hoogste gemiddelde diversiteits indeks is gevind in areas wat gewissel het van 2600-2800 meter bo seevlak. Die sentrum van genetiese diversiteit is skynbaar Shewa, Bale en Arssi provinsies, met inagneming van hoogte. Verdere ondersoek en *in situ* bewarings pogings moet gedoen word in hierdie gebiede.

Brassica carrinata het 'n aantal agronomiese voordele bo ander oliesaad gewasse, met soortgelyke ekologiese aanpassing. Die olie word egter as van swak kwaliteit beskou. Alhoewel pogings aangewend is om kwaliteit te verbeter, moet baie nog gedoen word om natuurlike variasie te gebruik wat bestaan binne die spesie vir vetsuur samestelling. Die doelwitte van hierdie studie was om olie inhoud en vetsuur samestelling te bepaal, met die reeks van genetiese variasie wat bestaan binne die material, en om genotipes te selekteer wat gebruik kan word om hierdie eienskappe te verbeter. Agt en negentig Etiopiese genotipes is gebruik om vetsuur samestelling variabiliteit te bepaal met CGC. Die oorheersende vetsure was erusien (6.91-35.05%), linoleïen (17.21-28.53%), α-linoleen (10.50-22.52%) gevolg deur gadoleïen (1.22-10.24%) en palmitien (3.38-16.42%). Palmitien suur was betekenisvol positief gekorreleer met stearien suur. Erusien suur was betekenisvol negatief gekorreleer met palmitien, stearien, vasien, oleïen, linoleïen en α-linoleen suur en positief gekorreleer

met eikosenosien suur. Die hoë olie opbrengs genotipes was relatief ryker aan erusien suur. Olie inhoud het gewissel van 25 tot 48%.

Genetiese diversiteit analise binne *B. carrinata* is nog nooit ondersoek met molekulêre merkers nie. Die doel van hierdie studie was om die genetiese verwantskappe tussen 39 *B. carrinata* genotipes te bepaal met die gebruik van AFLP merkers. Die studie het die hoeveelheid genetiese variasie ondersoek wat bruikbaar is vir teling. Die 39 genotipes van *B. carrinata* is geanaliseer met ses AFLP priemstuk kombinasies. 'n Totaal van 189 polimorfiese merkers is geëvalueer met 'n gemiddeld van 32 merkers per priemstuk kombinasie. GDE het gewissel van 0.346 tot 0.639 met 'n gemiddeld van 0.504±0.002. Polimorfisme persentasies het gewissel van 50 tot 80%. UPGMA tros analise het sewe duidelike groepe genotipes identifiseer.

Die genotipe omgewing interaksie en die stabiliteit van 14 *B. carrinata* genotipes vir saad opbrengs is geëvalueer by vier lokaliteite onder veld toestande vir drie jare in noord wes Etiopië. Die doel was om die komponent van variasie te vind wat geassossieer is met die eerste en tweede orde interaksies en hulle effekte, om genotipe stabiliteit te bepaal met verskillende stabiliteits statistiese modelle en om stabiliteits statistieke te vergelyk.

Die variansie analise het aangetoon dat hoof effekte was a.g.v. jare en lokaliteite, en eerste orde (jaar x lokaliteit) en tweede orde interaksies (genotipe x jaar x lokaliteit) was hoogs betekenisvol. Die genotipe x lokaliteit interaksie was nie betekenisvol nie. Die nie-betekenisvolle genotipe x lokaliteit interaksie en betekenisvolle genotipe x jaar effek het aangetoon dat meer klem geplaas moet word op meer jare eerder as 'n reeks lokaliteite. Die bepaling van variansie komponente wat eerste orde interaksies wat genotipe x jaar, genotipe x lokaliteit en die fout insluit was baie klein. Die variansie komponente van die hoof effekte van die genotipe en die tweede orde interaksie van genotipe x lokaliteit x jaar was baie hoog.

Gemiddelde opbrengs oor alle lokaliteite en jare het gewissel van 1304 tot 1541 kg/ha vir PGRC/E 210406 en PGRC/E 21163 onderskeidelik. Die regressie van genotipe gemiddelde opbrengs op die omgewings indeks het 'n regressie koeffisiënt waarde gegee wat wissel van 0.7488-1.2288. Die afwyking van die regressie was nie

betekenisvol vir alle genotipes. Die eerste twee interaksie hoof komponent asse van die AMMI (additive-main-effects-and multiplicative-interaction) saam het tussen 35.6% en 54.3% van die totale genotipe omgewings interaksie som van kwadrate van saad opbrengs verklaar. Die stabiliteits analise het PGRC/E 20017 en PGRS/E 20112 meer stabiel gevind, terwyl PGRC/E 21169 en (4DxZEM-1) x (ZEM-1-AD/88) vir spesifieke omgewings aangepas was.

Spearman se koeffisiënt van rangorde korrelasie is uitgevoer vir elke moontlike paarsgewyse vergelyking van verskillende stabiliteits statistieke. Betekenisvolle rang korrelasies is gevind tussen sekere stabiliteits statistieke. Hierdie betekenisvolle rang korrelasies het hulle effektiwiteit aangetoon om stabiele genotipes oor 'n reeks omgewings te bepaal. Die hierargale groepering van AMMI aangepaste gemiddelde opbrengste met die UPGMA tros analise metode het die genotipes ingedeel in vyf geneties verskillende groepe, en die omgewings in drie heterogene groeperings.

CHAPTER VIII

CONCLUSION AND RECOMMENDATION

Based on the characters studied across the regions, the highest diversity index was found in Shewa, Bale and Arssi provinces. The highest genetic diversity is concentrated in areas between altitudinal brackets of 2600 and 2800 masl. This altitudinal range includes the major *B. carinata* growing areas in the country. Further exploration and *in situ* conservation missions of *B. carinata* germplasm as a source of diversity should take account of the distribution of polymorphisms. Priorities of germplasm collection should focus on areas with relatively large variation with due consideration to the cause of genetic erosion, drought and environmental degradation.

The *B. carinata* germplasm were found to contain considerable variation for oil content and fatty acid compositions. The major fatty acids in the seed oil on average basis were erucic>linoleic>α-linolenic>oleic>gadoleic>palmitic. Gamma linolenic, lignoceric, dimo-γ-linolenic, margaric, palmitoleic, myristic were also rarely detected and only in small number of accessions. The high oil yielding genotypes were relatively richer in erucic acid content. There was significant and negative correlation between erucic acid and palmitic, stearic, vaccinic, linoleic, α-linolenic acids and positive correlations with eicosenoic acid. This result suggested that selection for a particular fatty acid profile would tend to increase those positively correlated, and decrease the negative ones. Selection and hybridization techniques could then be applied to modify the oil content and fatty acid composition considering the variability observed.

The genetic diversity analysis based on AFLP markers exhibited a high level of efficiency for detecting DNA polymorphisms among *B. carinata* genotypes. Knowledge of diversity patterns and specific GDEs may increase the efficiency of *B. carinata* genetic improvement in Ethiopia among adapted parents used for cultivar development. The knowledge of genetic distance is also used for organization of germplasm resources and identification of parents for crossing blocks to maximize genetic gain and heterosis.

The GE factorial analysis of variance identified the non-significant genotype x location and significant genotype x year interactions, which indicate that testing over years is important than over locations in north-western Ethiopia.

The significant GE interactions and changes in ranks of genotypes across environments suggest a breeding strategy of specifically adapted genotypes in homogeneously grouped environments.

The above ten stability statistics identified PGRC/E 20017 as the most stable genotype, and is recommended for official release in north-western Ethiopia. This can help Gomenzer farmers who have been growing mixed varieties to maintain stable productivity under diversified agro-ecologies.

The Wricke's procedure of stability statistic shows the highest significant positive correlation with Shukla's stability variance and deviation mean square. The rank correlation coefficient of 1.0 was found between Shukla's and Wricke's procedures. This indicates that the two procedures are equivalent for ranking purposes. Shukla's stability variance is a linear combination of the ecovalence and therefore, both Wricke's ecovalence and Shukla's stability variances are equivalent for ranking purposes.

AMMI combines the analysis of variance and principal component analysis in one model. It appears to describe more accurately both the GE interaction and stability analysis by means of response patterns, which can be easily elucidated from the biplot or the scattergram of IPCA1 and IPCA2 scores. The AMMI model provides more information not only in terms of stability measure, but also in terms of describing response and spatial patterns, clarifying genotypic effects and having inherent predictive value. Hence, the AMMI model should be the principal statistical method to be used to describe GE interaction and stability of *B. carinata* genotypes grown in north-western Ethiopia.

Cluster analysis was performed to study the pattern of groupings of genotypes and environments. The dendrograms were generated from the UPGMA clustering method of genotypes and environments based on Euclidian distances using AMMI adjusted

mean yield of environments and genotypes. UPGMA clustering method grouped the genotypes into five genetically distinct clusters and the environments into three clusters.