Genetic diversity in wheat breeding populations
by Chhandak Basu

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Agronomy
Montana State University

Abstract:
Genetic diversity is an essential tool for the crop breeders. Diversity can help the crop breeders to improve the crop species in several ways. Breeders continuously introduce new lines in the population which brings diversity in the population. Besides this, breeders look for a desired agronomic trait. If they do not get it in present population then they look for the desired trait in the parental population. All these methodologies use the genetic diversity principle. Genetic diversity among spring and winter wheat breeding population was tried to estimate. 47 varieties of spring and winter wheats were used in this research. They were grown in greenhouse and DNA was extracted from the six-week old leaves. Agarose gel electrophoresis was performed with the DNAs from the plants to check the concentration of DNA. PCR (Polymerase Chain Reaction) was performed afterwards with the DNA. After that AFLP (Amplified Fragment Length Polymorphism) was performed with DNAs. Polymorphisms were observed among all the varieties of plants and the bands were scored manually and the software NTSYS and KIN were used to analyze the data. A nonsignificant correlation ($r = 0.13$) was obtained when the data obtained from the pedigree information (Coefficient of Parentage or COP) was compared with the AFLP data. Solid stem varieties (Glenman, rampart, Vanguard, rescue, Fortune and MT9433) were found to be unrelated with each other and observed to be scattered throughout the tree diagram generated from the NTSYS. We also found out that there was no significant difference in the mean genetic similarity (GS) between the winter wheat population with the rest of the wheat data set. Spring and winter wheats were not found to be clustered separately based on the AFLP data set. The tree diagram generated from NTSYS based on the COP data showed the separation of spring and winter wheats, although the data generated from AFLP showed no separation of spring and winter wheats in the tree diagram.

GENETIC DIVERSITY IN WHEAT BREEDING POPULATIONS

by

Chhandak Basu

A thesis submitted in partial
fulfillment of the requirements for the degree

of

Master of Science

in

Agronomy

MONTANA STATE UNIVERSITY
Bozeman, Montana

August, 1998

APPROVAL

of a thesis by

Chhandak Basu

This thesis has been read by each member of the thesis committee and has been found satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the College of Graduate Studies.

Dr. Luther Talbert — Signature — Date 8/20/98

Approved for the Department of Plant, Soil, and Environmental Sciences

Dr. Luther Talbert — Signature — Date 8/20/98

Approved for the College of Graduate Studies

Dr. Joseph Fedock — Signature — Date 8/20/98

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a master's degree at Montana State University-Bozeman, I agree that the Library shall make it available to borrowers under rules of the Library.

If I have indicated my intention to copyright this thesis by including a copyright notice page, copying is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for permission for extended quotation from or reproduction of this thesis in whole or in parts may be granted only by the copyright holder.

Signature Chhandak Basu

Date 08/08/98

To my parents (Baba and Maa)

ACKNOWLEDGMENTS

I wish to express my sincere thanks to my major advisor Dr. Luther Talbert for his encouragement, patience and guidance throughout my stay with him and giving me an opportunity to work for this project. I also wish to express my gratitude to my other committee members Dr. John Martin and Dr. Phil Bruckner for their guidance, valuable advice and for helping me every respect to work for the project.

I wish to thank everyone in my lab and in my department including Laura Smith, Nancy Blake, Dr. Roy J. Martens, Ben Lehfeldt, Gail Sharp, Amber Hemphill, Susan Lanning, Dr. E. Sivamani, Jacob Anderson, Xueyan Shan, Dr. Mike Giroux, Arunrat Vanichanon and Dr. Jamie Sherman for their friendly behavior, encouragement and helping me learning new techniques. Without their cooperation and help it was not possible for me to work for this project.

I am really missing my friends and well wishers of India like Dadai, Barda, Sumanta, Shankuda, Babua, Lopa, Tumpa, my friends of Surendranath Cooperative and my 'mashis' of Maniktala Girls' High School, who are thousands of miles away from me. I remember their good wishes at this moment.

Thanks a lot to my wife Jhuma (Jayati), not for making delicious food or staying awake in lab nights after nights with me, but for encouragement, appreciation and true cooperation throughout my stay in Montana.

Lastly, I want to express my gratefulness to my parents who sacrificed their entire life for me and spent their last farthing to send me to U.S.A. Without their 'ashirbad' (blessings) I have nothing in my life.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Genetic diversity is an essential tool for the crop breeders. Diversity can help the crop breeders to improve the crop species in several ways. Breeders continuously introduce new lines in the population which brings diversity in the population. Besides this, breeders look for a desired agronomic trait. If they do not get it in present population then they look for the desired trait in the parental population. All these methodologies use the genetic diversity principle. Genetic diversity among spring and winter wheat breeding population was tried to estimate. 47 varieties of spring and winter wheats were used in this research. They were grown in greenhouse and DNA was extracted from the six-week old leaves. Agarose gel electrophoresis was performed with the DNAs from the plants to check the concentration of DNA. PCR (Polymerase Chain Reaction) was performed afterwards with the DNA. After that AFLP (Amplified Fragment Length Polymorphism) was performed with DNAs. Polymorphisms were observed among all the varieties of plants and the bands were scored manually and the software NTSYS and KIN were used to analyze the data. A nonsignificant correlation ($r = 0.13$) was obtained when the data obtained from the pedigree information (Coefficient of Parentage or COP) was compared with the AFLP data. Solid stem varieties (Glenman, rampart, Vanguard, rescue, Fortuna and MT9433) were found to be unrelated with each other and observed to be scattered throughout the tree diagram generated from the NTSYS. We also found out that there was no significant difference in the mean genetic similarity (GS) between the winter wheat population with the rest of the wheat data set. Spring and winter wheats were not found to be clustered separately based on the AFLP data set. The tree diagram generated from NTSYS based on the COP data showed the separation of spring and winter wheats, although the data generated from AFLP showed no separation of spring and winter wheats in the tree diagram.

CHAPTER 1

INTRODUCTION

Common wheat(*Triticum aestivum*) is the most important cultivated food crop superseding all other food crops from an economic point of view. Wheat is cultivated in almost every part of the world. In the year of 1996 the United States produced 2.28 billion bushels of wheat (http//:www.usda.gov/nass/aggraphs/awprod.htm). Several types of foods are made from wheat, including bread, pasta, cookies, cakes and muffins. Besides food use, wheat starch is used in laundering, in textile industries, for making adhesives and for wall paper and billboard paste. Wheat is also used as animal feed. Montana, whose economy is largely agricultural, ranks fourth in wheat production in the U.S.(Montana Agricultural Statistics Service, 1997, Helena) and produces 7.7% of total wheat production in the nation(M.A.S.S.,1997,Helena). Wheat and wheat products contribute about 8.8% of the state's agricultural exports (M.A.S.S.,1997,Helena).

Evolution of Wheat

There has been an intimate association between man and wheat since prehistoric time. The exact place of origin of wheat is still unknown. Diploid and tetraploid wheat were found to exist before 8000 B.C. in the Fertile Crescent area, nowadays known as the drainage basins of the Tigris and Euphrates rivers in the present day Syria and Iraq (Smith,1995). Hexaploid wheat is believed to have evolved before 7000 B.C. in the south of the Caspian Sea in the northern Iran(Smith,1995). Wheat kernels were found in the Egyptian pyramids and tombs in Mesopotamia (Christie et al.,1930). Common wheat's tetraploid ancestor, emmer wheat(*T. turgidum* ssp. *dicoccon,* BBAA) was domesticated about 9000 years ago (Cox,1998).

Origin of Wheat Genome

According to Kimber et al.(1987) an ancestral diploid wheat had differentiated into other diploids, each of which were given unique genome symbols (A,B,C,D,M,U...). Hybridization of diploid wheats with A and B genomes, followed by chromosome doubling formed tetraploid wheat AABB. Another diploid wheat with genome D combined with the tetraploid to form modern cultivated hexaploid wheat AABBDD. Sakamuara(1918), Kihara(1919,1924)and Sax(1922) showed that cultivated wheat constitutes an allopolyploid genome series, diploid through hexaploid (Kimber et al., 1987).

Importance of Genetic Diversity in Crop Breeding

Genetic diversity among breeding material is essential for effective selection. Breeders strive to introduce variability in the genetic pool. Variation may be negative, so breeders want only those variations which have desired agronomic characteristics. Breeders continuously introduce new lines into their germplasm base. In autogamous crop species, like wheat, the pedigree method of breeding is mostly used (Snape et al.,1983). Selection can be of two types - artificial and natural. Both of these types of selection can produce diversity among a crop

species. The first step of a selection process is to choose parents with desired agronomic traits. The next step is to cross them. Then the breeders select from the offspring and introduce new lines. These new lines may produce enough diversity for breeders to select again from these lines and cross them, introducing another set of new lines. This cycle has continued since the birth of plant breeding. At the same time breeders have practiced selection for a specific set of agronomic traits.

Gene flow can play a vital role in changing the characteristics of a crop species. The best example of this fact is production of disease resistant plants by genetic manipulation and using genetic diversity. One of the main objectives of plant breeders is to produce disease resistant plant. If the desired disease resistance gene is not available in the adapted germplasm, then the breeder needs to explore other populations in order to find the favorable gene.

The introduction of dwarf variety of wheat was a triumph for the wheat breeders. The shorter stem increased the harvest index. The dwarfing gene most commonly present in the U.S.A originally came from Japanese variety Daruma, which had the dwarfing genes *Rht 1* and *Rht 2*(Holden *et al.*,

1993). Daruma's derivative Shirodaruma was crossed with the North American variety Fultz in 1917 (Holden *et al.*, 1993). Ultimately, we got the most important dwarfing gene carrying variety Norin-10. The introduction of diversity in crop populations, increase chances of successful selection and obtaining new crop cultivars with new useful agronomic traits. So, genetic diversity is very essential in a crop breeding program. The successful utilization of genetic diversity in a breeding project will lead to selection of useful characteristics from a population which will eventually lead to introduction of new varieties.

CHAPTER 2

LITERATURE REVIEW

Determination of Diversity

There are different types of molecular marker methods used nowadays to determine DNA polymorphisms among cultivars. The molecular marker methods are : RFLP (Restriction Fragment Length Polymorphism), RAPD(Random Amplified Polymorphic DNA), PCR(Polymerase Chain Reaction), AFLP(Amplified Fragment Length Polymorphism) etc. Out of these methods AFLP(Vos etal.,1995) is considered to be most highly efficient compared to other methods(Barrett et al., 1998). In AFLP large number of polymorphic bands are observed compared to other methods. So, AFLP is very useful in determination of genetic diversity.

Determination of genetic diversity and utilization of genetic diversity to improve crop species are very interesting research topic among the breeders. Quite a few studies have been done on this aspect.

Souza et al.(1994) worked on spring wheat diversity in Mexico and Pakistan based on pedigree method. They found out the genetic diversity in the Yaqui valley in Mexico in a given year was 20% less than in Punjab, Pakistan. But on the other hand, the rate of change in germplasm was 22% lower in Pakistani Punjab than Yaqui valley, Mexico. Talbert et al.(1994) found that genetic similarity among the hard red spring wheat was 0.88(percentage of shared restriction fragment length polymorphism). They concluded that the breeding pool for hard red spring wheat was narrow compared to levels of diversity in hexaploid wheat. Bryan et al. (1994) found out that genome of hexaploid bread wheat exhibits very low levels of restriction fragment length polymorphism (RFLP) perhaps due to relatively recent origin of hexaploid wheat. Tsunewaki et al.(1993) evaluated wheat germplasm by RFLP. They found that nucleotide diversities in both einkorn and emmer wheats are two to three times greater than the diversities of common wheat and nucleotide diversity between einkorn species is two times than those of emmer wheat. They also found that nucleotide diversity between einkorn species is about five times larger than that of common wheat. They concluded that commomn wheat is the least diversified whereas einkorn

is the most. It should be mentioned here that, Turkey and Hard Red Calcutta are considered to be the major ancestors of hard red winter and hard red spring wheats respectively.

Comparison Between Pedigree Method and Marker Technology to Determine Genetic Diversity

Both the COP (Coefficient of Parentage) method and molecular marker technology have been used to determine genetic diversity among crop species. Some of the researchers got very good agreement between the COP values and marker analysis, while the others not. The following table(Table 1) may help us to get some ideas about the effectiveness of the two methods.

Table 1: Comparison between genetic similarity indicated by COP method and genetic similarity indicated by molecular marker method (Source: Burkhamer et al.,1998; Barrett et al.,1998; Schut et al., 1997)

Name of the researchers	**Name of the crop**	**Did the COP values were in good agreement with marker analyzed values ?**
Moser et al.,1994	Oat	Not very much (r= 0.63)
Barbosa-Neto et al., 1996	Wheat	Positive(comparison between diversity and COP)
Plaschke et al., 1996	Wheat	Moderate (r=0.55)

From the above table we can see that in most of the cases the COP and molecular marker data did not match considerably. The reason behind this mismatch may be the assumptions that are made at the time of calculating COP values. COP calculation assumes that if two genotypes are not related by pedigree then they do not carry homologous DNA fragments. Barrett et al.(1998)pointed out that some drawbacks for this assumption including the fact that 55 landraces (which were unrelated by pedigree) from Afganistan, Iran and Turkey were found to have high levels of genetic similarity(mean=0.91)(Kim and Ward,1997). They

also mentioned, that the assumption in case of COP that, with the presence of selection pressure and drift each parent contribute 50% of the genetic material to the offspring had been proved to be invalid by Siedler et al.(1994). Besides these the COP calculation assumes that there will be no mutation, migration or selection pressure at the population. This may not be true in practical cases.

Barrett et al.(1998) worked on AFLP based genetic diversity assessment among wheat cultivars from Pacific Northwest. They found out that mean genetic diversity estimates(based on data from AFLP) were highest(0.58) for spring vs. winter type, intermediate(0.53) within winter wheats and lowest(0.49) within spring type. Cox et al.(1986),on the basis of pedigree analysis, found that genetic diversity has increased in hard red winter(HRW) wheat germplasm and decreased slightly in soft red winter (SRW) wheat germplasm during this century. Barrett et al.(1998) stated that determination of genetic diversity is more effective based on AFLP data than from pedigree analysis. They found that mean of 903 genetic diversity estimate on the basis of pedigree was 0.96. The genetic diversity estimate, on the basis of AFLP data, was normally distributed(mean=0.54).

Contribution of Germplasm From Different Countries to U.S. Wheat Gene Pool

Beuningen et al.(1997) studied genetic diversity among North American spring wheat cultivars. The wheat variety Hard Red Calcutta(from India) contributed 23% Canadian Western Red Spring(CWRS) wheat cultivars and 21% to the hard red spring wheat group. Approximately 124 ancestors from 32 countries on five continents significantly contributed to the North American spring wheat gene pool. Russia-Ukraine contributed the most germplasm, followed by Poland and India(Beuningen et al., 1997). Ukraine was considered to be the origin of some ancestral winter wheat introductions such as 'Turkey', 'Cheyenne' and 'Rio'(Cox, 1991). Russia-Ukraine contributed 23, 11 and 8% to the CWRS, HRS and WS(white spring wheat) group(Beuningen et al.,1997). India contributed 25, 14, 4 and 9% to the CWRS, HRS(Hard red spring), CIMMYT and WS group. 'Fife', 'Hard Red Calcutta' and 'Turkey Red' contributed the most to the hard red spring wheats released from 1901 to 1991(Mercado et al.,1996). 'Kenya 324'(from Kenya) contributed 0% to the HRS wheat cultivars from 1901-1940 but contributed 5.5% from 1941-1991(Mercado et al.,1996). Other countries also contributed significantly to the North American hard spring

wheat cultivars, including Kenya(7.1 %), Brazil(6.6%), Egypt (1.6%) and Morocco(2.0%)(Mercado et al.,1996).

CHAPTER 3

MATERIALS AND METHODS

Plant Materials

(***Note:*** *The entire work was done using 47 varieties of wheat plants. But at the time of data analysis, it was found out that 13 of them did not work very well in AFLP. So, for the accuracy of the experiment, those varieties were discarded from the data set at the time of data analysis and result interpretation. The final data analysis was done by using 34 varieties of wheat plants.)*

Forty-seven varieties of spring and winter wheats were grown in separate pots in greenhouse. Within each growth habit, varieties were chosen in two distinct classes. 'Inputs' are varieties or lines that were used as parents in the breeding programs, usually in order to incorporate a specific trait. Members of the input (Table 2) group are part of the pedigrees for the output group. The outputs (Table 3) are all varieties released or scheduled for release from the Montana wheat breeding programs. Out of the 47 varieties of spring and winter wheats, we have one variety called Turkey selection which is a selection from the landrace variety Turkey.

Table 2:

Varieties released before 1975 (INPUT)

(Note: HRW= Hard red winter, HRS= Hard red spring, WW= winter wheat, SW= spring wheat)

Variety	Growth Habit	Release Date	Traits of Interest
Marquis	WW	1913	Stripe rust (in adults)resistance
Minturki	WW	1919	Winterhardiness
Ceres	HRS	1926	Stem rust resistance, drought resistance
Marquillo	HRS	1928	Stem rust resistance
Yogo	HRW	1932	Winterhardiness
Fronteira	SW	1932	High protein content, leaf and stem rust resistance
Cheyenne	HRW	1933	Shatter resistance, high yield, stiff straw
Norin 10	WW	1935	Dwarf nature and high yield
Pilot	HRS	1939	Stem rust resistance
Comanche	HRW	1942	Earliness, stiff Straw, excellent baking and milling qualities
Mentana	WW	1945	Stripe rust and lodging resistance
Kenya 58	WW	1945	Lodging and Shatter resistance
Rescue	HRS	1946	Solid stem and saw fly resistance
Thatcher	SW	1948	Stem rust resistance, shatter resistance, lodging resistance
Brevor	WW	1949	Good bread making quality

Table 2 (Continued):

Variety	Growth Habit	Release Date	Traits of Interest
Winalta	HRW	1962	Winterhardiness and excellent milling quality
Fortuna	HRS	1966	Solid stem
Froid	WW	1968	Winterhardiness
Centurk	WW	1971	Moderately resistance to stripe rust, lodging and straw break

Table 3:

Varieties released since 1975 (OUTPUT)

Variety	Growth Habit	Release Date	Traits of Interest
Neeley	HRW	1980	Stripe rust (inadults) resistance
Cree	HRW	1982	Semi dwarf, Winterhardiness
Norwin	HRW	1984	Winterhardiness
Glenman	HRS	1985	Stripe rust resistance
Tiber	HRW	1988	Winterhardiness, shatter resistance
Hiline	HRS	1991	Stem rust resistance
McGuire	HRW	1966	High protein
Vanguard	HRW	1996	Solid stem
Rampart	HRW	1996	Solid stem

DNA Extraction

DNA was extracted from 47 genotypes of wheats by the Extraction buffer was prepared by following the recipe cited below.

Stock Solution	150µl
1M Tris	15 µl
4M NaCl	3.75 µl
10% SDS	15 µl
14.4M Mercaptoethanol	104 µl
Sterile H_2O	101.14 µl

This extraction buffer is sufficient for genomic DNA extraction from at least eight plants.

About three leaves per six-week old wheat plant were placed in wet paper towels and kept at 4°C. Leaf tissue was ground in mortar and pestle with 15 ml of extraction buffer. The crushed leaf tissue with the buffer was transferred to Oakridge tubes and incubated at 65°C for 10 minutes. Five ml of 5M potassium acetate was added to each tube and tubes were incubated at -20°C for 20 minutes. The supernatant was poured through a miracloth filter into another 30 ml tube which had 10 ml cold isopropanol and 1ml 5M ammonium acetate. The tubes were gently mixed and incubated at -20°C in ice for at least 20 minutes. The tubes were spun again at 20,000 X gravity for 15 minutes to pellet the DNA.

The supernatant was poured off gently and pellets were dried by inverting tubes on paper towels for 5-10 minutes. DNA pellets were redissolved in 0.7 ml TE. The solutions were transferred into a microcentrifuge tube and 75µl 3M sodium acetate (pH 7.0) and 500 µl of cold isopropanol were added. The tubes were gently mixed and after 30 seconds of mixing, the tubes were spun and the clot of DNA was observed. The supernatant was poured off and the DNA pellets were dried thoroughly. DNA was redissolved again in 100-200 µl TE. The DNA samples were stored at -20^{0}C.

Agarose Gel Electrophoresis

Agarose gel electrophoresis was performed to check the concentration of DNA. The agarose gel was made by combining 90ml of water, 0.8 gm of agarose, 10 ml of TBE and 10 µl of ethidium bromide in a conical flask and heated to boiling until agarose dissolved. Then the solution was allowed to cool for a few minutes and poured into a gel electrophoresis tray. Standard comb was used to make wells in the tray. After the gel solidified, 2.5 µl of DNA, 8.0 µl of sterile water and 3 µl of dye(dye was diluted to 1:10 when mixing) were added to each well. Samples were run for about 45 minutes at 85 volts. After taking the photograph in UV light, the concentration of each sample was

determined by comparison to a sample of known concentration.

Polymerase Chain Reaction

Polymerase chain reaction(PCR) amplification were performed using 50 µl reactions in each tube. The reactions for a single sample consisted of 28.85 µl sterile water, 5 µl 10X reaction buffer, 8 µl dNTP(1.25 mM),), 3 µl $MgCl_2$ (25 mM), 1.5 µl of each of left and right primers, 0.15 µl *Taq* polymerase(5 u/µl) and 2 µl(about 50 ng) of DNA. The following primers were used- D14, H8, E8, G49 and F11. The above reactions were prepared for 48 samples. Reactions were performed 0.5 ml microfuge tubes and overlaid with two drops of mineral oil. PCR was performed with the following temperature conditions: initial denaturation at 94^{0}C for 4 minutes, 30 cycles of 94^{0}C for 1 minute, 45^{0}C for 1 minute, 72^{0}C for 1.2 minutes and again 72^{0}C for 7 minutes followed by a hold at 4^{0}C. Contents of each reaction tube were equally divided into two parts (25 µl each) and added to two assay plates. Restriction enzymes *HhaI* and *HinfI* (0.2 µl *Hha/Hinf* and 4.8 µl sterile water for one well in the assay plate) were added to each of the assay plates separately and incubated at 37^{0}C for 1 hour. Then the reactions were run in a polyacrylamide gel(22 ml ddH_2O, 3 ml 10X TBE, 7 ml

30% acrylamide, 150 µl 20% ammonium persulfate and 15 µl Temed- this is a recipe for 1 polyacrylamide gel at 0.75 mm spacing) with 0.5% Tris-Borate EDTA running buffer (22mM Tris-HCl, 22mM Boric Acid and 0.5 mM EDTA) with 30 volts current per gel. The gels were run for approximately 2 hours and 30 minutes, stained with ethidium bromide for about 10 minutes and observed under UV light and photographed. Five primers were used to perform the PCR. The objective of the PCR was to be sure of the fact that each DNA sample was of sufficient quality.

Amplified Fragment Length Polymorphism

Amplified Fragment Length Polymorphism (AFLP) were performedafter the PCR. AFLP is one of the latest technologies to fingerprint genomic DNA. It was used in this research project to visualize DNA polymorphism among different genotypes of wheat following the protocols of Vos et al.(1995). The AFLP was performed by the following steps:

Restriction Digestion of Genomic DNA

Master mix was prepared by this recipe: 5 µl 5X reaction buffer, 2.0 µl *EcoI* and *MseI* primers, 15.5 µl sterile H_2O. 22.5 µl of reaction mixture was aliquoted to each of the 48 microfuge tubes. Then 2.5 µl of DNA (about

50 ng) was added to each tube and incubated at 37°C for 70 minutes.

Ligation of Adapters

Each tube had 25 µl reaction from the previous step. Now, 25 µl of ligation solution was added to each tube and incubated at 20°C for 2 hours in thermocycler. The tubes were spun down in a centrifuge. Now a 1:10 dilution of the ligated product was made by taking 10 µl of ligated product and 90 µl of TE. They were then stored at -20°C.

Preamplification Reaction

Strip tubes (0.2 ml., in a batch of eight) were used for this purpose. The tubes were numbered from 1-48 and placed in ice. Five µl of 1:10 diluted (digested and ligated from the previous step) DNA was added to each tube. Then the master mix was prepared by the following recipe: 40 µl pre-amplification primer mix, 5 µl 10X PCR buffer for AFLP, 1 µl *Taq* DNA polymerase (5 u/µl). Then 46 µl of this master mix was added to the 5 µl of DNA in each tube. Tubes were placed in Perkin-Elmer thermocycler and the preamplification program(20 cycles of 94°C for 30 seconds, 56°C for 60 seconds and 72°C for 60 seconds)was run.

Primer Labelling

First the primer combination was selected for the day (10 primer combinations were used for this research project). The names of the primer combinations are-

1. E-AGC
   M-CTG
2. E-ACC
   M-CTA
3. E-ACC
   M-CTG
4. E-AGG
   M-CAT
5. E-AGC
   M-CAG
6. E-AGC
   M-CAT
7. E-ACG
   M-CAC
8. E-ACG
   M-CAC
9. E-AAG
   M-CTC
10. E-ACT
    M-CAC

Master mix was prepared by this recipe (enough for 54 reactions):

	(µl)
*Eco*RI primer	14.4
AFLP grade water	8.0
5X kinase buffer	8.0
^{33}P	8.0
T4 polynucleotide kinase	1.6
Total volume	40.0

The reaction mixture(in 2ml tube) was placed in 37°C water bath for 1 hour and then 70°C for 15 minutes. The second step in water bath in 70°C for 15 minutes was performed to deactivate the kinase.

Selective Amplification

Strip tubes(0.2 ml., in a batch of eight) were used for this step. 5 µl of preamplification product of DNA was added to each tube. Two types of mixture were prepared for this step. Those are as follows:

Mix 1	
MseI Primer	270 ul
Labelled *EcoRI*	30 ul
Total	300 ul
Mix 2	
AFLP Grade water	474 ul
10X PCR buffer for AFLP	120 ul
Taq DNA polymerase	6 ul
Total	600 ul

Five µl of Mix 1 and 10 µl of Mix 2 were added to each tube and run in Perkin-Elmer thermocycler using the selective amplification program:

1. One cycle at 94°C for 30 seconds; 65°C for 30 seconds and 72°C for 60 seconds

2. Annealing temperature was lowered each cycle 0.7°C during 12 cycles
3. Twenty three cycles at 94°C for 30 seconds; 65°C for 30 seconds and 72°C for 60 seconds.

Preparation of AFLP Gel

Gel was prepared by this recipe:

Urea	31.5 gm.
10x Sequencing buffer	7.7 ml.
40% Acrylamide for sequencing	11.25 ml.

These reagents were added in a 150 ml beaker along with 75 ml sterile water. Then the mixture was mixed thoroughly with an electric stirrer. Sequencing gel running apparatus(glass plates, clamps, casting tray, combs) were arranged accordingly. First 15 ml of the above reaction mixture was poured(with 60 µl 25% ammonium persulphate and 60 µl of Temed) in the casting tray to seal the bottom portion of the glass plates. After waiting for 30 minutes minutes the remaining gel mixture was poured between the two glass plates(with 50 µl 25% ammonium persulphate and 50 µl Temed). A 48 well comb was placed at the top of the gel and entire apparatus was laid down horizontally. Wet paper towel was pressed between the two plates at the top and covered with plastic wrap for at

least two hours. After two hours the gel was ready to be run.

Running the Gel

Before loading the samples in the gel, the gel was heated to 50^{0}C with 0.5X TBE buffer with appropriate electric current(approximate 50 volt). The samples were denatured in Perkin-Elmer thermocyler (90^{0}C for 3 minutes and hold at 4^{0}C) and loaded in each well in the gel. The gel was run for about 3 hours at about 50-65 volts (current was altered accordingly to keep the gel temperature constant at 50^{0}C). After the electrophoresis was completed, the gel was dried in gel drier for about 40 minutes and exposed to Kodak Bio Max Film for 2-3 days. After 2-3 days the film was developed and the bands were scored manually. Ten primer combinations were used in AFLP analysis and 118 bands were scored manually('1' was used to symbolize 'presence of a band', '0' was used to symbolize absence of a band and '9' was used to denote missing data).

CHAPTER 4

DATA ANALYSIS

Calculation of Coefficient of Parentage (COP)

Coefficient of Parentage(COP) values were calculated from the pedigree records, by following the methods of Cox et al.(1985) and by using the software KIN(Tinker et al., 1993).

Tinker et al.(1993) defined the kinship coefficient (*r*) as probability of two alleles at a locus are identical by descent. The kin software helps us to find the *r* value (they also called *r* as 'coefficient of coancestry'). This *r* value is an estimate of genetic similarity(GS). In the following way the software KIN was used to calculate the COP values.

(This methodology have been taken from following the methods of Tinker et al., 1993)

Step 1. An input text file was created where in each line had a cultivar with its two parents.

Step 2. A matrix was generated which contain the value r_{xy} for pairwise XY combinations. The value of r_{xy} was generated by using the formula: $r_{xy} = 1/2(r_{X.P} + r_{X.Q})$

(Here X and Y are the cultivars, and P and Q are the parents of Y)

The following assumptions were made at the time of calculating values for the COP data file:

1. There were no selection pressure, migration or drift in the population or at the segregating progeny.
2. Each parent contributed half of its alleles to its offspring.
3. All other ancestors were unrelated and all lines were completely inbred.
4. The genotypes, which were not related by its pedigree do not carry homologous DNA fragments.

The matrix generated from COP values were analyzed using the software NTSYS(SUNY, Stony Brook) to generate the tree diagram (Figure 6).

Calculation of AFLP Data Matrix

The software NTSYS (SUNY, Stony Brook) was used to generate the tree diagram for the AFLP (Figure 1, Appendix A). In the following ways, the data matrix generated from AFLP was calculated by using NTSYS.

Step 1: The initial data file(generated manually from AFLP) was analyzed using the program 'SIMQUAL'. The coefficient used for this calculation was Dice coefficient (Dice, 1945), which measures the amount of association among cultivars. This 'SIMQUAL' program calculates the similarity and dissimilarity coefficients among cultivars.

Step 2: The file generated from the 'SIMQUAL' program was used to form 'SAHN' clustering(Sneath and Sokal,(1973) referred 'SAHN' as 'Sequential, Agglomerative, Hierarchical, and Nested clustering methods'). At the time of forming the clustering, the 'UPGMA'(Unweighted pair-group method, arithmetic average) clustering method was used.

Step 3. The 'SAHN' clustering file was used to generate the tree diagram (Figure 1, Appendix A), which gives us a visual representation of genetic similarity (GS) among cultivars.

CHAPTER 5

RESULTS AND DISCUSSIONS

AFLP Analysis

The pairwise AFLP genetic similarity (GS) value appeared to be normally distributed(Figure 2, Appendix A) with a mean value of similarity of 0.5838. Genetic similarity(GS) was calculated using NTSYS. The input and output winter wheats were observed to form clusters at different positions in the dendogram. From the tree diagram (Figure 1, Appendix A), generated from AFLP data, the following observations (Table 4) were made.

Table 4: Genetic similarity (GS) among different winter and spring wheat input and output groups

Cluster 1 (GS 0.698)
a)Winter wheat input
Cree, Minturki, Centurk, Comanche and Yogo
b)Winter wheat output
MT91192
Cluster 2 (GS 0.68)
a)Winter wheat input
Norin10
b) Winter wheat output
Nuwest, McGuire and Tiber
c) Spring wheat
MT 9433
Cluster 3 (GS 0.662)
a)Winter wheat input
Cheyenne and Turkey Selection
b) Winter wheat output
Erhardt
c)Spring wheat
Hi-Line, Red River68
Cluster 4 (GS 0.662)
a)Winter wheat input
Winalta, Marquis and Neeley
b) Spring wheat
Thatcher and Hard Red Calcutta
Cluster 5 (GS0.68)
a)Winter wheat output
Rampart, Norwin and Vanguard

We observed few clusters in the dendogram (Figure 1, Appendix A). A cluster was formed with the basis of the following assumptions:

1. There must be at least 3 entries to form a cluster

2. The GS value must range above 0.662

From the above clusters and from the tree diagram (Figure 1, Appendix A) we can see that winter wheats and

spring wheats are not clustered separately.

Another analysis was done to check whether the AFLP data was consistent with prior expectations. Marquis, a pedigree of Thatcher, is found to be in the same cluster. Hard Red Calcutta, a pedigree of Marquis, found to be in the same cluster with Thatcher and Marquis. Cree, which is a back cross derivative of Cheyenne, was found to be widely separated from Cheyenne based on AFLP. Glenman, which is a solid stem variety, was not genetically similar to other solid stem varieties in the data set. Marquis and Hard Red Calcutta (which is the pedigree of Marquis) were found to be in the same cluster. The two Turkey selections (Turkey Selection, CI 12137) were found to be widely separated from each other. Closely related varieties like Vanguard and Rampart are found to be fairly close to each other in the tree diagram (Figure 1, Appendix A). From Table 5, we can see the relationship of related varieties based on AFLP analysis.

Table 5: The relationship of related varieties based on AFLP analysis(also see Table 12,Appendix B)

Comparison	Mean GS
All wheats	0.58
Solid stems	0.48
Turkeys	0.73
Winter wheats	0.66
Cree vs. Cheyenne	0.71
winter wheat input group	0.594
winter wheat output group	0.6048

Comparison Among the Solid Stem Varieties and Rest of the Wheat Populations

The solid stem varieties in my data set are- Glenman, Vanguard, Rampart, Rescue, Fortuna and MT9433.

A t-test(Table 6)was performed to check whether the mean genetic similarity of the solid stem varieties differed from the rest of the wheat genotypes as follows.

Table 6: t-test to compare the solid stem varieties with the rest of the wheat population

	Variance
Solid stem varieties	0.010257
Rest of the wheat populations	0.08265
solid stem varieties)	
The calculated t value= 1.61 with 34 degrees of freedom The table value of $t_{0.01,\ 34df} = 2.724$	

There is no significant difference in mean GS between the solid stem varieties and the remaining wheats. The solid stem varieties are found to be dispersed throughout the tree diagram(Figure 1, Appendix A) because of their low mean GS.

Comparison Between Turkey Selections

There are two Turkey selections in my data set. Those are- Turkey Selection and CItr No. 12137. The pairwise GS between these two are 0.733 (Table 10, Appendix B). This signifies that the mean GS between the Turkey Selections are fairly high but compared to all wheat genotypes in the data set they are widely separated from each other.

Comparison Between Winter Wheat Input and Output Group

Winter wheat input and output group matrices were generated from NTSYS and a t-test (Table 5) was performed to check, whether there was significant difference between winter wheat input and output groups.

There is no significant difference in mean GS between the winter wheat input and output group. In my opinion, since the beginning of the breeding history the winter wheat input and output groups have come closer to each other. As a result of which they lost their diversity and so we did not detect any difference in mean GS between the winter wheat input and output group.

Comparison Between Winter Wheats With Overall Wheat Population

All winter wheats in this data set were compared(by t-test, Table 5) with the entire wheat data set to check whether the winter wheats are different from the entire data set.

There is no significant difference in mean GS between the winter wheat population and the rest of the wheat data set.

COP Data Analysis

A tree diagram (Figure 4, Appendix A)was generated from the COP data using NTSYS. The COP data matrix was used to do the following tests. From the tree diagram, generated from the COP data matrix, the following observations were made. We found out two distinct clusters of spring and winter wheats. Here we defined cluster as follows:

1. There must be at least 10 entries
2. The GS value must be above 0.192 to 0.288

The following table (Table 7) may help to understand the cluster pattern.

Table 7:Genetic similarity among different winter and spring wheat groups from the COP data
Cluster 1 (Mean GS 0.192) **Winter wheats:** Cree, Cheyenne, Neeley, CItr 12137, Turkey Selection, Comanche, Nuwest, Rampart, Vanguard, Yogo, Tiber, Minturki, Winalta, Erhardt, Norwin, MT 91192, Centurk, Norin 10, McGuire
Cluster 2 (Mean GS 0.288) **Spring wheats:** Red River 68, Glenman, MT 9433, Fortuna, Rescue, Ceres, Pilot, Hiline, Marquillo, Thatcher, Marquis

Comparison Between the Solid Stem Varieties With Rest of the Wheat Population

A t-test was performed to check whether the mean GS of the solid stem varieties differed from the rest of the wheat population (Table 8) from the COP data set.

Table 8: t-test to compare the solid stem varieties with the rest of the wheat population

	Mean GS
Solid stem varieties	0.25
Rest of the wheat populations (excluding solid stem varieties)	0.20
	Variance
Solid stem varieties	0.033
Rest of the wheat populations (excluding solid stem varieties)	0.030

The calculated t value= 0.62 with 34 degrees of freedom (the table value of $t_{0.01,\ 34df} = 2.724$). From the above table

we can conclude that there is no significant difference in mean GS between the solid stem varieties and the remaining wheats from the COP data.

Comparison Among Closely Related Varieties From COP Data Set

An analysis was done to check whether the COP data was consistent with prior expectations. The solid stem varieties in the COP data set are found to have a very low mean GS value and as a result of which they have found to be separated in the tree diagram generated from the COP values. The Turkey Selections have been found to be placed closely with a mean GS of 0.75. Cree, which is a backcross derivative of Cheyenne have been found to be closely placed in the tree diagram generated from COP. From the following table (Table 9), we can see the relationship of related varieties based of COP analysis.

Table 9: Genetic similarity of related varieties based on COP analysis(also see Table 12,Appendix B)

Comparison	Mean GS
Solid stems	0.25
Turkeys	0.75
All wheats	0.20
Cree vs. Cheyenne	0.97
Vanguard vs. Rampart	0.76

Comparison Between the AFLP and Pedigree Data Matrices

When the data matrix from AFLP analysis and the data matrix files generated from the pedigree analysis are plotted by using NTSYS, a scatter diagram (Figure 7, Appendix A) was obtained. There was little relationship between these two data matrices. The matrix correlation was only 0.13(not significant). Burkhamer et al.(1998) worked on parental divergence in spring wheat. They found significant correlation between the pedigree (COP) with the AFLP data (r= 0.55).

When the frequency distribution of the data matrices from the COP data and the other from the AFLP data were plotted (Figure 2 and 3, Appendix A) we found that the AFLP data appear to be normally distributed (Figure 2) whereas the COP data matrix were skewed towards the low values (towards origin)(Figure 3).

From the above result we see that the mean of COP based GS matrix is 0.20 (skewed), whereas the mean of AFLP marker based GS matrix is 0.5838 (somewhat normally distributed).

The COP and AFLP data matrices were compared with cophenetic(COPH) values(Rohlf and Sokal, 1981). The 'COPH' values are used to test goodness of fit of a clustering to a set of data(NTSYS, 1997). The COPH matrix was generated from the 'SAHN' clustering from NTSYS. Then the COPH matrix

was plotted against the COP or AFLP data matrix(Figure 5 and 6). Very good goodness of fit (r=0.96)was obtained from COP data and moderate fit(r=0.67) was obtained from the AFLP data.

Conclusion

Spring and winter wheat genotypes were not found to be clustered separately based on the AFLP data. The genepool of 'inputs' and 'outputs' of winter wheats are found to be mixed with each other in most cases. Although, a tree diagram generated from COP data showed separation spring and winter wheats(Figure 4, Appendix A).

The mean COP based GS values was 0.20(skewed), whereas the mean of AFLP marker based GS was 0.5838(normally distributed). The two frequency distribution graphs(Figure 2 and 3, Appendix A) of AFLP and pedigree based data analysis tell us about the distribution of values in those two cases and also about the center of these distributions. The COP based analysis measures alleles which are identical by descent. The AFLP based analysis measures both the alleles identical by descent and also alleles identical in state. According to Barrett et al.(1998), because of the skewness in the data matrix of pedigree values, genetic resolution among different varieties or cultivars was not achieved. Barrett et al. (1998) made a similar observation

when they compared AFLP and pedigree based genetic diversity assessment with wheat cultivars from Pacific Northwest regions of U.S.

The mean GS of solid stem varieties based on AFLP was found to be less than the rest of wheat data set. This indicates that the rest of the wheat population is much more similar to each other than the solid stem varieties. As a result the solid stem varieties are dispersed, as expected, in the tree diagram generated from AFLP (Figure 1).

When the winter wheat input and output data matrices are compared, we found no significant difference between these two groups. But it must be emphasized that more the primer combinations used in AFLP, more accurate will be the data analysis. So, instead of 10 primer combinations, if 50 primer combinations were used, it is likely that we would get more accurate result. This fact must be kept in mind when anyone wants to interpret any result from molecular marker methods.

BIBLIOGRAPHY

Barbosa-Neto, J.F., M.E Sorrells and G. Cisar. 1996. Prediction of heterosis in wheat using coefficient of parentage and RFLp-based estimates of genetic relationship. Genome. 39:1142-1149.

Barrett, B.A., and K.K. Kidwell. 1998. AFLP-based Genetic Diversity Assessment Among Wheat Cultivars from the Pacific Northwest. Crop Science(In Press).

Beuningen, L.T. and R.H. Busch. Genetic diversity among North American spring wheat cultivars: I. Analysis of the coefficient of parentage matrix. Crop Science. 37:570-579.

Beuningen, L.T. and R.H. Busch. Genetic diversity among North American spring wheat cultivars: II. Ancestor contributions to gene pools of different eras and regions. 1997. Crop Science. 37:580-585.

Beuningen, L.T. and R.H. Busch. Genetic diversity among North American spring wheat cultivars: III.Cluster analysis based on quantitative morphological traits. 1997. 37:981-988.

Bos, I., and P. Caligari. Selection Methods in Plant Breeding. 1995. Chapman and Hall, NY.

Burkhamer, R.L., S.P. Lanning, R.J. Martins, J.M. Martin and L.E. Talbert. Predicting progeny variance from parental divergence in hard red spring wheat. 1998. Crop Science. 38:243-248.

Chen, H.B., J.M. Martin, M. Lavin, L.E. Talbert. Genetic diversity in hard red spring wheat based on PCR markers. 1994. Crop Science.34:1628-1632.

Cox, T.S., G.L. Lookhart, D.E. Walker, L.G. Harrell, L.D. Albers and D.M. Rodgers. Genetic relationships among hard red winter wheat cultivars as evaluated by pedigree analysis and gliadin polyacrylamide gel electrophoretic patterns. 1985. Crop Science. 25: 1058-1063.

Cox, T.S., Y.T. Kiang, M.B. Gorman and D.M. Rodgers. Relationship between coefficient of parentage and genetic similarity indices in the soybean. 1985.

Crop Science. 25:529-532.

Damania, A.B. (Editor). Biodiversity and Wheat Improvement. 1993. John Wiley and Sons, NY.

Dice, L.R. Measures of the amount of ecological association between species. 1945. Ecology. 26:297-302.

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

Eaton, D.L., R.H. Busch and V.L. Youngs. Introgression of unadapted germplasm into adapted spring wheat. Crop Science. 1986. 26:473-478.

Falconer, D.S. Introduction to Quantitative Genetics. 1989. Longman Group UK Ltd, Essex, England.

Feyerherm, A.M., G.M. Paulsen and J.L. Sebaugh. Contribution of genetic improvement to recent wheat yield increases in the USA. 1984. Agronomy Journal. 76:985-990.

Finney, D.J. Experimental Design and Its Statistical Basis. 1955. The University of Chicago Press, IL.

Hahn, V., K. Blankenhorn, M. Schwall and A.E. Melchinger. Relationships among early European maize inbreds:III. Genetic diversity revealed with RAPD markers and comparison with RFLP and pedigree data. 1995. Maydica. 40:299-310.

Kisha, T.J., C.H. Sneller, and B.W.Diers. 1997. Relationship between Genetic Distance among Parents and Genetic Variance in Populations of Soybean. Crop Science. 37:1317-1325.

Lewin, B. Genes. 1987. John Wiley and Sons, NY.

Martin, J.M., L.E. talbert, S.P. Lanning and N.K. Blake. Hybrid performance in wheat as related to parental diversity. 1995. Crop Science. 35:104-108.

Melchinger, A.E., M.M. Messmer, M. Lee, W.L. Woodman, and K.R. Lamkey. 1991. Diversity and Relationships among U.S. Maize Inbreds Revealed by Restriction Fragment Length Polymorphisms. Crop Science. 31:669-678.

Mercado, L.A., E. Souza and K.D. Kephart. Origin and diversity of North American hard spring wheats. 1996. Theoretical and Applied Genetics. 93:593-599.

Poulsen, G.B., G. Kahl, and K. Weising. 1993. Abundance and polymorphism of simple repetitive DNA sequences in Brassica napus L. Theor. Appl. Gennet. 85:994-1000.

Poehlman, J.M., and D.A. Sleper. Breeding Field Crops. 1995. Iowa State University Press, IA.

Rosner, B. Fundamentals of Biostatistics. 1982. Duxbury Press, MA.

Schut, J.W., X. Qi, P. Stam. 1997. Association between relationship measures based on AFLP markers, pedigree data and morphological traits in barley. Theor. Appl. Genet. 95:1161-1168.

Siedler, H., M.M. Messmer, G.M. Schachermayr. 1994. Genetic diversity in European wheat and spelt breeding material based on RFLP data. Theoretical and Applied Genetics. 88:994-1003.

Smith, E.L. (Editor). Genetic Improvement in Yield of Wheat. 1987. CSSA Special Publication. CSSA, WI.

Smith, J.S.C., E.C.L. Chin, H. Shu, O.S. Smith, S.J. Wall, M.L. Senior, S.E. Mitchell, S. Kresovich and J. Ziegle. An eavaluation of the utility of SSR loci as molecular markers in maize (*Zea mays*): Comparisons with data from RFLPs and pedigree. 1997. Theoretical and Applied Genetics. 95:163-173.

Sneath, P.H.A., and R.R. Sokal. Numerical Taxonomy: The Principles and Practice of Numerical Classification. 1973. W.H. Freemnan and Company, CA.

Snedecor, G.W. Statistical Methods: Applied to Experiments in Agriculture and Biology. 1938. Collegiate Press, Inc., IA.

Sokal, R.R., and F.J. Rohlf. Biometry. 1981. W.H. Freeman and Company, CA.

Srivastava, J.P. and A.B. Damania. Wheat Genetic Resources: Meeting Diverse Needs. 1990. John Wiley and Sons. NY.

Suzuki, D.T., and A.J.F. Griffiths, Jeffrey H. Miller, R.C. Lewontin. 1986. An Introduction to Genetic Analysis. W.H. Freeman and Company, NY.

Talbert, L.E., N.K. Blake, P.W. Chee, T.K. Blake and G.M. Magyar. Evaluation of "sequence-tagged-site" PCR products as molecular markers in wheat. 1994. Theoretical and Applied Genetics. 87:789-794.

Talbert, L.E., P.L. Bruckner, L.Y. Smith, R. Sears and T.J. martin. 1996. Theoretical and Applied Genetics. 93:463-467.

Tinker, N.A. and D.E. Mather. KIN: Software for computing kinship coefficients. The Journal of Heredity. 1993. 84(3):238.

Tsegaye, S., T. Tesemma, and G. Belay. Relationships among tetraploid wheat (*Triticum turgidumi*) landrace populations revealed by isozyme markers and agronomic traits. 1996. Theoretical and Applied Genetics. 93:600-605.

Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman, M. Kuiper and M. Zabeau. 1995. Nucleic Acids Research. 23:4407-4414.

Watson, J.D., J. Tooze and D.T. Kurtz. Recombinant DNA: A Short Course. 1983. W.H. Freeman and Company, NY.

Weising K., H. Nybom, K. Wolff and Wheat. Meyer. DNA Fingerprinting in Plants and Fungi. 1995. CRC Press, MI.

Watson, J.D., N.H. Hopkins, J.W. Roberts, J.A. Steitz and A.M. Weiner. 1987. Molecular Biology of the Gene. (Voulme 1).

Whitaker, J.W. (Editor). Farming in the Midwest 1840-1900. 1974. The Agricultural History Society, WA.

Yu, K. and K.P. Pauls. Rapid estimation of genetic relatedness among heterogeneous populations of alfalfa by random amplification of bulked genomic DNA samples. 1993. Theoretical and Applied Genetics. 86:788-794.

Appendix A

Figures

Fig. 1: Tree diagram generated from AFLP data using NTSYS