

90969

**A NOVEL METHOD FOR DEVELOPING MICROSATELLITE
MARKERS AND SEQUENCE ANALYSIS OF WHEAT
DISEASE RESISTANCE GENE ANALOGS**

**A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
THE MIDDLE EAST TECHNICAL UNIVERSITY**

BY

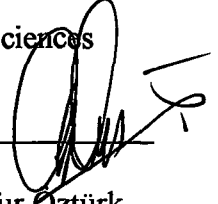
ERDOĞAN EŞREF HAKKI

**IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
IN
THE DEPARTMENT OF BIOTECHNOLOGY**

NOVEMBER 2000

**T.C. YÜKSEKÖĞRETİM KURULU
DOKÜMANTASYON MERKEZİ**

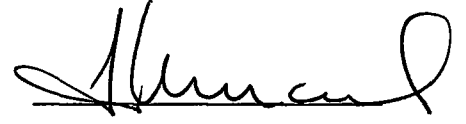
Approval of the Graduate School of Natural and Applied Sciences



Prof. Dr. Tayfur Öztürk

Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Doctor of Philosophy.



Prof. Dr. Suzan Kincal

Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Doctor of Philosophy.



Prof. Dr. Aykut Kence

Co-supervisor



Assoc. Prof. Dr. Mahinur S. Akkaya

Supervisor

Examining Committee Members

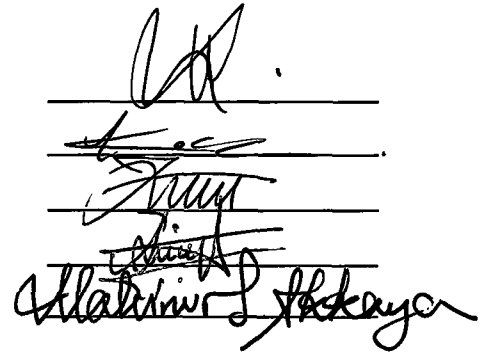
Prof. Dr. Ufuk Gündüz

Prof. Dr. Semra Kocabıyık

Assoc. Prof. Dr. Zümrüt Ögel

Assist. Prof. Dr. Ahmet Yıldırım

Assoc. Prof. Dr. Mahinur S. Akkaya



ABSTRACT

A NOVEL METHOD FOR DEVELOPING MICROSATELLITE MARKERS AND SEQUENCE ANALYSIS OF WHEAT DISEASE RESISTANCE GENE ANALOGS

Hakkı, Erdoğan Eşref

Ph.D. Department of Biotechnology

Supervisor: Assoc. Prof. Dr. Mahinur S. Akkaya

Co-supervisor: Prof. Dr. Aykut Kence

November 2000, 128 pages

Wheat is an important crop for the entire world and especially for Turkey, who is within the first 10 producers. The wheat consumption in Turkey is the largest per capita in the world. Wheat agriculture was originated in Anatolia, thus the rich gene pool of this geography should be characterized with advanced molecular genetics tools aiding the germplasm conservation and the characterization of the local varieties in terms of yield and resistance to co-evolved pathogens.

A novel approach in isolating microsatellites from wheat, equally applicable to any other eukaryotic genome, by utilizing AFLP system is developed in this thesis study.

Turkish durum wheat doubled haploid (DH) lines 'Çakmak-79', 'Berkmen-469' and 'Kunduru-1149' were compared at the DNA level, using highly polymorphic 10 microsatellite markers (two showing two loci), out of 15 markers studied. The genetic relationship among the lines was determined. Data produced also combined the progenitors of the DH lines as well as the other local varieties, in order to evaluate the genetic relationship in larger scale.

Yellow rust is an important wheat disease caused by the fungus *Puccinia striiformis*. It is considered the most important production limiting disease in Turkey, with major recent epidemics occurring in six of the past 25 years. The varieties 'Pastor', with one major and two minor, 'Avocet', with two minor resistance genes to yellow rust, the completely susceptible 'Chinese spring' variety and the promising wild types *T. monococcum*, *T. dicoccoides* and *A. speltoides* have been utilized in resistance gene analog (RGA) isolations *via* homology-based PCR approach. Genetic relationships of RGAs with other resistance genes (in GenBank) were determined.

Key Words: Genetic Relationship, Microsatellites or SSRs, Molecular Markers, Novel SSR Isolation, *Puccinia striiformis*, Resistance Gene Analogs (RGA), Wheat, Yellow Rust.



ÖZ

YENİ BİR METODLA MİKROSATELİT BELİRLEYİCİLERİNİN GELİŞTİRİLMESİ VE BUĞDAYDA HASTALIĞA DİREÇİLİLİK GENİ ANALOGLARININ SEKANS ANALİZİ

Hakkı, Erdoğan Eşref

Doktora, Biyoteknoloji Bölümü

Tez Yöneticisi: Doç. Dr. Mahinur S. Akkaya

Ortak Tez Yöneticisi: Prof. Dr. Aykut Kence

Kasım 2000, 128 sayfa

Buğday tüm dünya için, özellikle de en büyük 10 üretici ve dünyanın kişi başına tüketimi en fazla olan ülkesi, Türkiye, için önemli bir tahıldır. Buğday tarımı Anadolu orijinelidir, dolayısıyla bu zengin gen kaynaklarının korunması ve gerek verim, gerekse parojenlere karşı dirençlilik özellikleri gözetilerek, Anadolu gen havuzundaki yerel çeşitlerin ileri moleküler teknikler ile karakterizasyonu gerçekleştirilmelidir. Bu tezde, AFLP sistemi kullanılarak, yeni bir yaklaşımla buğdaydan mikrosatelitler geliştirilmiştir. Bu yeni yöntem, tüm ökaryotik genomlara eşit oranda uygulanabilir niteliktedir.

Bu çalışmada kullanılan 15 belirleyiciden, son derece yüksek polimorfizm gösteren 10 mikrosatelit belirleyicisi (ikisi iki ayrı lokusta amplifikasyon yapmakta) kullanılarak, çift haploid (ÇH) hatlar DNA düzeyinde karşılaştırılmış ve hatlar ('Çakmak-79', 'Berkmen-469' ve 'Kunduru-1149') arasındaki genetik ilişki belirlenmiştir. Türk durum buğdayları arasında daha iyi bir ilişkinin elde edilebilmesi için, ÇH örnekleri ile onların ebeveynleri ve diğer yerel varyetelere ait veriler biraraya getirilmiştir.

Sarı pas, *Puccinia striiformis*'in sebep olduđu önemli bir buğday hastalığıdır. Bu hastalık, son 25 yıl içinde önemli üretim kayıplarına sebep olan altı önemli epidemin görüldüğü Türkiye için daha da önemlidir. 'Pastor' varyetesi sarı pasa dirençlilik genlerinden bir major ve iki minor gene sahiptir. 'Avocet' varyetesi sarı pasa karşı iki minor gene sahipken 'Chinese spring' bu pasa tamamen hassastır. *T. monococcum*, *T. dicoccoides* ve *A. speltoides* yabani buğday türleri ise paslara karşı dirençlilik potansiyeli olan bitkilerdir. Homolojiye dayalı PCR yöntemi ile direnç geni analoglarının (DGA) izolasyonunda bu bitkiler kullanılmıştır. Genbankalarında mevcut olan direnç genleri ile DGA'lar arasındaki genetik ilişkiler belirlenmiştir.

Anahtar kelimeler: Genetik İlişki, Mikrosatelitler veya SSRlar, Moleküler Belirleyiciler, Yeni SSR İzolasyonu, *Puccinia striiformis*, Direnç Geni Analogları (RGA), Buğday, Sarı Pas.





To Our Baby

ACKNOWLEDGEMENTS

I would like to thank to my supervisor *Dr. Mahinur S. Akkaya* for all her helps, endless support and continuous encouragement during the research as well as out of the lab and for allowing me to undertake the research for this degree under her guidance.

I gratefully acknowledge the aids and supports of *Dr. Engin Akkaya* and would like to thank him for the proof reading of our papers and for tolerating my somewhat unexpected intrusions into their lives. I offer sincere thanks to all my *friends and colleagues in the lab*, including the recently graduated ones, and the students of Dr. Engin Akkaya who have made my working time more pleasant, for they shared all the advantages and/or disadvantages of a crowded lab. Thanks go to Prof. *Dr. Aykut Kence* for his positive contributions with his suggestions and for sharing a genetic analysis software with us. I am grateful to *Dr. Çiğdem Savaşkan* for supplying the DH lines. I express sincere appreciation to Dr. Steven S. Clark from the University of Michigan, Biology Department, for inviting me to his lab and allowing me to carry out sequencing analysis. I am also grateful to everybody in the Clark lab and especially to Lita and Amy without whose technical support everything would have been a great deal harder.

I acknowledge the financial supports of the Middle East Technical University (AFP-2000-07-02-00-11) and that of Tübitak (TBAG-1720), and especially for the project related to my research activities in Ann Arbor (Tübitak-NATO A2). Thanks also go to the company GAMA for support in the form of a scholarship.

I am highly indebted to my parents who keep on believing in me up until now and for their endless support.

TABLE OF CONTENTS

ABSTRACT.....	iii
ÖZ.....	v
ACKNOWLEDGEMENTS.....	viii
TABLE OF CONTENTS.....	ix
LIST OF TABLES.....	xiv
LIST OF FIGURES.....	xv
LIST OF ABBREVIATIONS.....	xvii
CHAPTER.....	
1. INTRODUCTION.....	1
1.1. WHEAT.....	1
1.1.1. Cultivation of Wheat.....	4
1.1.2. Wheat Genome.....	4
1.1.3. Geographical Centers of Origin and Centers of Diversity.....	7
1.1.4. Wheat Domestication and Human Civilization.....	8
1.1.5. An Historical Profile of Germplasm Flows.....	9
1.2. MOLECULAR MARKERS.....	10
1.2.1. Restriction Fragment Length Polymorphisms (RFLPs).....	10
1.2.2. Variable Number of Tandem Repeats (VNTR)- Hypervariable minisatellites.....	10
1.2.3. Randomly Amplified Polymorphic DNA (RAPD) Markers.....	11
1.2.4. Microsatellites.....	11
1.2.5. Amplified Fragment Length Polymorphism (AFLP).....	14

1.2.6. Comparisons of DNA-based Molecular Markers....	16
1.3. OUTLINE OF THE THESIS.....	17
1.4. A NOVEL METHOD DEVELOPED FOR MICROSATELLITE ISOLATION.....	18
1.5. GENOTYPING OF ANATOLIAN DOUBLED- HAPLOID DURUM LINES WITH SSR MARKERS. ..	21
1.6. ANALYSIS OF DISEASE RESISTANCE GENE ANALOGS..... ..	22
1.6.1. Resistance Breeding.....	22
1.6.2. Isolation of Resistance Genes.....	23
1.6.3. Resistance Gene Analogs.....	24
2. MATERIALS AND METHODS.....	29
2.1. MICROSATELLITE ISOLATIONS.....	29
2.1.1. Triple Affinity Capturing (TAC) Enrichment Method.....	29
2.1.1.1. Plant Materials.....	30
2.1.1.2. Isolation and Digestion of Wheat Genomic DNA..... ..	30
2.1.1.3. (CT) ₉ Enrichment of 'Kundurur-1149' DNA.....	31
2.1.1.4. Digestion of pKSII KS ^{+/} - DNA.....	32
2.1.1.5. Construction of Genomic DNA Library and Maintaining of the Clones with Insert DNA.....	32
2.1.2. Double Affinity Capturing (DAC) Enrichment Method.....	32
2.1.2.1. Sequencing of Enriched Microsatellite Candidates.....	32
2.1.3. Isolation of AFLP-based Microsatellites:A Novel Approach.....	33
2.1.3.1. Plant Materials.....	33
2.1.3.2. DNA Isolation	33
2.1.3.3. AFLP Sample Preparation.....	34

2.1.3.4.	Restriction Enzyme Digestion of Genomic DNA.....	34
2.1.3.5.	Double Stranded Adapter Preparations.....	34
2.1.3.6.	Ligation of Adapters.....	34
2.1.3.7.	Pre-selective Amplification.....	35
2.1.3.8.	Selective Amplification.....	35
2.1.3.9.	Enrichment of Selectively Amplified PCR Products.....	35
2.1.3.10.	Radiolabeled Reaction Conditions.....	36
2.1.3.11.	[γ - ³³ P]-ATP Labeling of ϕ x174 DNA/ <i>Hinf</i> I Molecular Size Marker.....	37
2.1.3.12.	Seperation of Selective Amplification Bands.....	37
2.1.3.13.	Extraction of DNA from the AFLP Gel.....	38
2.1.3.14.	Re-amplification of Band-isolated AFLP Fragments.....	38
2.1.3.15.	Selection of the Fragments According to their Restriction Sites.....	39
2.1.3.16.	Sequencing of Fragments.....	40
2.2.	GENOTYPING OF ANATOLIAN DOUBLED-HAPLOID DURUM LINES WITH SSR MARKERS. ..	41
2.2.1.	Plant Materials.....	41
2.2.2.	DNA Isolation and PCRs.....	41
2.2.3.	PCR Analysis.....	42
2.2.4.	Cluster Analysis.....	42
2.3.	ISOLATION OF WHEAT RESISTANCE GENE ANALOGS.....	43
2.3.1.	Plant Materials.....	43
2.3.2.	DNA Isolations.....	43
2.3.3.	Design of the PCR Primers for RGA Amplifications.....	44
2.3.4.	PCR Conditions in RGA Amplifications.....	44
2.3.5.	Cloning of PCR Products.....	45
2.3.6.	Sequencing of Cloned Amplification Products.....	46
2.3.7.	Sequence and Homology Analysis.....	47

3. RESULTS.....	48
3.1. ISOLATION OF NEW MICROSATELLITE MARKERS.....	48
3.1.1. Triple Affinity Capturing (TAC) Enrichment Method.....	48
3.1.2. Double Affinity Capturing (DAC) Enrichment Method.....	48
3.1.3. Isolation of AFLP-Based Microsatellites: A Novel Approach.....	50
3.2. GENOTYPING OF ANATOLIAN DOUBLED-HAPLOID DURUM LINES WITH SSR MARKERS...	58
3.3. ISOLATION AND SEQUENCING OF WHEAT RESISTANCE GENE ANALOGS.....	67
3.3.1. PCR Products of RGAs.....	68
3.3.2. Alignments of RGA Sequences.....	72
3.3.3. Homologies Obtained with the GenBank Sequences.....	73
3.3.4. Phylogenetic Analysis of Wheat RGAs.....	81
3.3.5. Alignment and Phylogenetic Analysis of Uninterrupted RGA Sequences Based on Amino Acid and Nucleotide Sequences.....	84
4. CONCLUSIONS.....	92
REFERENCES.....	97
APPENDICES.....	108
A. <i>T. dicoccoides</i> (Acc. No. TR02771) Resistance Gene Analog Sequences.....	108
B. <i>T. monococcum</i> (Acc. No. TR36938) Resistance Gene Analog Sequences.....	109
C. Resistance Gene Analog Sequences from 'Avocet' Variety.....	110
D. Resistance Gene Analog Sequences from 'Chinese spring' variety.....	114
E. Resistance Gene Analog Sequences from 'Pastor' Variety.....	116

F. CLUSTAL X (1.81) multiple sequence alignment of all wheat RGA sequences.....	120
CURRICULUM VITAE.....	126



LIST OF TABLES

TABLE

1.1.	Comparison of DNA-based molecular marker systems.....	16
2.1.	Degenerate primers, targeting NBS-LRR regions, used in resistance gene analog amplifications.....	44
2.2.	PCR conditions of RGAs generated by degenerate primer sets.....	45
2.3.	The total number of maintained RGA clones and the sequenced ones.....	46
3.1.	Bands cut off and isolated from AFLP gel, labeled as in Figure 3.3.....	53
3.2.	DNA sequences obtained by direct sequencing of the fragments containing best repeat regions.....	56
3.3.	Allele sizes (bp) of DH lines using 12 microsatellite loci.....	62
3.4.	The genetic analysis between the populations: "Rho" values, $(\Delta\mu)^2$ and 'Nei72' genetic distances (10000 permutations and bootstraps).....	64
3.5.	The clones selected for sequencing and their translation status.....	71
3.6.	Homologous sequences found in GenBank, using Blast Algorithm....	74

LIST OF FIGURES

FIGURE

1.1.	AFLP procedure.....	15
1.2.	Overview of the isolation steps.....	19
1.3.	Isolation and purification of enriched fragments from the sequencing gel and sequencing of the fragments.....	20
1.4.	The responses of the plant host cell to the pathogen elicitors.....	23
1.5.	A. Chromosome walking-jumping. B. Chromosome walking interruptions in repeating sequences.....	24
1.6.	The motives shared by plants, insects and mammals related to defense and developmental genes.....	28
2.1.	Triple affinity capturing enrichment procedure.....	29
2.2.	Double affinity capturing enrichment procedure.....	36
2.3.	Band extraction from the dried polyacrylamide gel.....	38
2.4.	Band isolation from the agarose gel.....	39
2.5.	Determination of restriction sites of the amplified fragments.....	40
2.6.	Direct sequencing of fragments with different 3' and 5' ends.....	41
3.1.	Sequences of the new microsatellites generated by DAC enrichment method.....	49
3.2.	A schematic representation of the SSR isolation steps, exemplified with Eco+ATT and Mse+ATA primers combination.....	51
3.3.	Autoradiograph of AFLP fragments	52
3.4.	PCR re-amplified band-isolated AFLP fragments from DNA denaturing polyacrylamide gel.....	54
3.5.	The sequences obtained by novel SSR isolation method.....	55
3.6.	Autoradiograph of microsatellite alleles at each locus:	59
3.7.	UPGMA clustering of DH individual lines based on simple matching coefficient similarity matrix of microsatellite data set.....	63

3.8.	UPGMA dendrogram generated from NTSYS genetic analysis program by using $(\delta\mu)^2$ genetic length matrix.....	64
3.9.	UPGMA generated dendrogram from NTSYS software determining the genetic relation of the DH lines with their progenitors and other Turkish cultivars.....	65
3.10.	PCR product obtained using degenerate primers	70
3.11.	Genetic similarity of all 84 RGAs based on neighbor-joining method of ClustalX 1.81.....	73
3.12.	Genetic distances of the homologs of groups.....	82
3.13.	Alignment of amino acid sequences of the RGAs without any stop codon.....	84
3.14.	Genetic relation of the RGA amino acid sequences that are uninterrupted	85
3.15.	Genetic relationship of the RGA nucleotide sequences that do not have any stop codons	87
3.16.	Amino acid sequence alignment of uninterrupted RGAs combined with available resistance genes in GenBank.....	89
3.17.	Genetic relationship of RGAs in amino acid level without any stop codons with available resistance genes from GenBank.....	90

LIST OF ABBREVIATIONS

$[\alpha\text{-}^{32}\text{P}]\text{-dATP}$	: Alpha ^{32}P deoxy-adenosine triphosphate
$[\gamma\text{-}^{33}\text{P}]\text{-ATP}$	: Gama ^{33}P adenosine triphosphate
μg	: Microgram
μL	: Microliter
A	: Adenine
APS	: Ammoniumpersulfate
ATP	: Adenosine triphosphate
bp	: Base pair
C	: Cytosine
CTAB	: Hexadecylmethylammonium bromide
DAC	: Double Affinity Capturing
dATP	: Deoxyadenosine triphosphate
DNA	: Deoxyribonucleic acid
EDTA	: Ethylenediaminetetraacetic acid
G	: Guanine
LRR	: Leucine rich repeat
M	: Molar
mg	: Miligram
min	: Minute
mL	: Mililiter
mM	: Milimolar
NBS	: Nucleotide binding site

Nei72 :

$$d_{ij} = -\ln \left[\frac{\sum_k |X_{ki} X_{kj}|}{\sqrt{\sum_k X_{ki}^2 X_{kj}^2}} \right]$$

ng	: Nanogram
NJ	: Neighbor-joining
pmol	: Pico mole
RGA	: Resistance gene analog
RNA	: Ribonucleic acid
rpm	: rotation per minute
RT	: Room temperature
SAHN	: Sequential, Agglomerative, Hierarchial, and Nested Clustering Methods
SDS	:Sodium dodecyl sulfate
sec	: Second
simgend	: Similarity for genetic data
simqual	: Similarity for qualitative data
SMC	: simple matching coefficient
SSC	: Sodium chloride sodium citrate buffer
T	: Thymine
TAC	: Triple Affinity Capturing
TEMED	: N, N, N', N'-Tetramethylethylene diamine
Tris-Cl	: tris-hydroxymethyl-aminomethane-HCl
UPGMA	: Unweighted pair-group method, arithmetic average
v/v	: volume/volume
v/w	: volume/weight

CHAPTER I

INTRODUCTION

1.1 WHEAT

The domesticated grasses, wheat, barley, rice and maize, from the Poaceae or Gramineae family constitute the main source of human and animal nutrition by making up about half of the total world food production. Gramineae is one of the largest and most diverse families of flowering plants, containing about 10,000 species. Major cereals present in this family are rice, wheat, maize, barley, sorghum, millet as well as sugar cane. The oldest known grass fossils were found in Paleocene-eocene deposits that are about 60 million years old but morphological and molecular clock data suggests that the grasses had a monophyletic origin over 70 million years ago. Not only grasses are very diverse in morphology and growth habit but also they are morphologically distinct from other plant families. These key grasses also differ considerably in chromosome number and genome size (Bevan and Murphy, 1999; Bennetzen and Freeling, 1997). The large-scale genome projects of these grasses is thus important as the Human Genome Project, since these efforts to engineer plants are expected to improve grain yield and drought, pests, salt resistances, and other problems for the future of the human being with a population that is expected to double by the year 2050 (Cohen, 1997).

Among the products of agricultural research are modern varieties developed in corn, rice, wheat and other crops. Wheat is the most widely grown crop in the world and its economical significance for the humankind is matched only by rice. Global production of wheat is now approaching 600 million t, with international trade approximately 100 million t annually (<http://www.fao.org/WAICENT/FAOINFO/ECONOMIC/economic.htm>). Up until recently, both global production and international trade in wheat have broadly kept pace with growth in the human population. However, in recent years there has been a slowdown in the growth of global wheat production relative to the growth in human population. As a consequence, global reserves of wheat have fallen to disturbingly low levels.

Given the critical role of wheat as a food crop, because of its wide adaptability and flexibility in end uses, it is important that the balance between global production growth and human population growth be restored. World wheat demand is conservatively estimated to increase by 40% by 2020. This will have to come predominantly from productivity increases rather than area expansion (Johnson, 1997; Anderson *et al.*, 1997; Bender, 1997).

Turkey is among the biggest wheat producers worldwide. The first producer is China, followed by India, USA, Russia, France, Germany and Turkey. Wheat production in Turkey varies between 16 and 21 million tons depending on the meteorological conditions and the epidemics countered in this geography quite often. Durum wheat makes 4-6 million of the total production, remaining being bread wheat. Wheat is sown on 9-9.5 million ha of which around 1 million ha are irrigated. There is another 5 million ha occupied by fallow. Winter and facultative wheat account for about 6.5 million ha (60-70% of the production). Bread wheat is mainly consumed as leavened bread in cities and flat bread in villages, pasta and bulgur are the major processed durum wheat products (Braun, pers. comm.). Due to the extremely divergent agro-ecological conditions, the grain yield varies from 1 t ha⁻¹ in the winter wheat areas of Eastern Turkey to 3 t ha⁻¹ in South and European part (Thrace) of Turkey. The wheat consumption in Turkey is among the highest in the world (about 200 kg per capita per year). Considering the world average yield is 2.7 t ha⁻¹ and major European producers having more than 7 t ha⁻¹, Turkey's 2.08 t ha⁻¹ average yield in the year 2000 is very low (<http://apps.fao.org/page/collections?subset=agriculture>).

Wheat belongs to the tribe *Triticeae* comprising some 300 species classified into 22 genera including several other important crops (barley, rye, and triticale) and a number of important forage species. Of the different species recognized as members of the *Triticum* genus that wheat belongs, only *T. aestivum* (bread wheat) and durum (pasta or macaroni wheat) are used commercially.

Wheat species underwent a natural evolutionary differentiation into three types: Winter, Spring and Facultative types of wheat. Winter wheat's total cycle span is 9-11 months. These varieties are planted during the fall, where the seeds emerge and produce vegetative growth. During winter, wheat undergoes an inactive stage. In the spring the grown stems rapidly initiate vegetative growth and produce during the summer. Spring wheats have a life cycle of 3 to 6 months. Unlike winter wheats, spring wheats cannot survive low temperatures. Facultative wheats are intermediate in their tolerance to cold. Each type of wheat has a different degree of resistance to diseases and environmental stress conditions.

In Turkey, winter and facultative wheats are primarily grown in eastern and central Anatolia and in Thrace. Until the early 1970's durum wheat predominated in Thrace but then replaced rapidly by the high yielding bread wheat cultivar Bezostaja. In the transitional zone, facultative wheats, especially winter durum cultivars are widely grown. Two thirds of Turkey's wheat producing area is located in the transitional zones and on the central Anatolian Plateau (CAP). Through time many varieties were selected out of the three types of wheat. Durum wheat production makes up of 20-30% of the total spring type production, major production areas being the low lands of southeastern Anatolia. Under favorable ecological conditions, the average yield of spring-habit wheat is much higher than the other wheat, 3 t ha⁻¹ with a potential of 10 t ha⁻¹, however epidemics of yellow rust occasionally occur and thus the high production limits are not achieved.

Traditionally all the varieties (TV) grown in the world were tall, which ranged in height between 125 and 150 centimeters. TVs did not require as much weed and pest control because of the low population planting density and the density of their plant canopy. Tall varieties normally produced a low grain yield but possessed high straw output. Because of their low yields, researchers tried to improve yields by experimenting with higher dosages of nitrogen fertilization. With heavier doses of nitrogen fertilization tall varieties increased the production of stem and leaves. In some cases their propensity to fall over (lodge) caused failure to increase yields. Researchers failed to improve stem strength, but did manage to improve rust resistance in tall varieties.

Modern varieties are shorter in stature, more responsive to fertilizers and yield more grain than traditional or "tall" varieties. As a consequence of the adoption of modern varieties and chemical fertilizers, world-wide wheat production has increased significantly.

In Turkey, wheat breeding started in 1925 at Eskişehir and in 1931 the first wheat variety, which is a selection of local populations, was released and the first one derived from crosses released in 1944. The "Turkish Green Revolution" was achieved from the National Wheat Research and Training Project established in 1967, and wheat production doubled within a decade. The main contributors of the revolution were the development of improved Turkish winter wheat cultivars, the introduction of spring wheat from CIMMYT and the winter wheat from Russia and USA and the improvements achieved in agronomic practices by the joint Turkey / Rockefeller Foundation / Oregon State University / USAID / CIMMYT (Braun, pers. comm.).

In many cases the release of newer varieties has been a response to newer strains of rusts and/or mildews which tend to break down rust resistance a few years after a variety is released. A major breakthrough was made the varieties descending from crosses of winter and spring type wheat (so called Veery types) during the 1980's. Veery type varieties dramatically improved yields, and at the same time maintained disease resistance. In 1982-1983 the proportion of the wheat area cultivated with Modern Varieties (MV) varies from 30.6% in the Near East and Communist Asia to a 79.2% in Asia.

1.1.1 Cultivation of Wheat

Some scientists speculate that events leading to the formation of today's wheat may have occurred many times in nature, rather than as a single hybridization event. Bread wheat is the most widely grown wheat species and one of the few crops for which no wild forms have been identified although its primary gene pool contains species (e.g., einkorn, emmer) that have wild forms and that continue to share a natural gene flow with wild grasses. If, as many suppose, bread wheat has been cultivated since its hybridization, then it has long been relatively isolated from other species and may have a much lower potential for genetic variation than its wild progenitors and other relatives. On the other hand, its early geographic dispersal may have contributed to its wide adaptability. According to some scientists, landraces maintained in collections offer only limited possibilities for diversifying the gene pool of domesticated bread wheat. By some standards, durums, emmers, and other species and genera have contributed at least as much to bread wheat breeding programs as have bread wheat landraces. Thus, wild germplasm is a valuable resource for improving bread wheat productivity and durability. As human and animal populations expand, however, the marginal lands on which these wild relatives thrive may need protection. Wide crosses and other biotechnological innovations have the potential for breaking the crossing barriers encountered in conventional breeding. So do transformation techniques that use genes extracted from other crops and organisms as well.

1.1.2 Wheat Genome

Wheat genomes, like all genomes in the tribe Triticeae, are large in comparison to the current plant models, arabidopsis and rice. *A. thaliana* has a genome size of 15×10^7 bases and was estimated to have ~25,000 genes. On the other hand, wheat has about 30,000 genes with a genome size as large as 16×10^9 (same number of genes with rice of genome size 43×10^7). Much of the difference in genome size is attributable to differences in amounts of repetitive

DNA; larger genomes like wheat or barley are composed of >75% repetitive DNA whereas smaller genomes like rice contain <50% repetitive DNA. It was revealed that in some large genome grasses retrotransposons inserted between genes account for the majority of the repetitive DNA (Benetzen and Freeling, 1997; Cohen, 1997). Transposable elements are an integral part of most genomes and probably the majority of interspersed repetitive DNA is composed of transposons with a crucial role in the evolution of developmental processes. This is dramatically seen by the wide presence of mutant phenotypes generated by transposon insertions. It was recently shown that, like in mammalian genes, wild type plant genes also harbor mobile elements. Retrotransposons of the *copia* class are ubiquitous in plant genomes flanking wild type plant genes. Short interspersed nuclear elements (SINEs) and long interspersed nuclear elements (LINEs) are similarly present in plant genomes but in lesser frequencies near plant genes. In contrary, the unusual inverted-repeat elements, *Tourist* and *Stowaway*, have been identified in association with over a hundred plant gene sequences (Bureau *et al.*, 1996).

Genome duplications have contributed to the expansion of genome size in cereals as in the case of allopolyploid bread-wheat genome that is composed of three ancestral diploid barley-sized genomes or that of maize with an ancient tetraploidization (2,400 Mb haploid genome size) (Bevan and Murphy, 1999).

While wheat genomes have probably similar numbers of gene motifs as the small-genome plants, the large genome size makes it unrealistic to anticipate complete sequencing any of the *Triticeae* genomes in the near future. Therefore, large-scale discovery, isolation and deciphering gene function in wheat (when the high percentage of repeating sequences, ~80%, also considered) and its relatives must rely on other, less direct methods. The current approach of map-based cloning, transposon mutagenesis, differential display techniques, and other strategies, are either not suitable for large genomes or are tedious and expensive because they deal with a single gene at a time (Bevan and Murphy, 1999). Therefore, it is necessary to devise alternative strategies to access genomes of wheat and its relatives on a large scale and thus ensure continued advances in the biology of this immensely important plant. There is such a strategy that focuses solely on the expressed component of the wheat genomes and takes a maximum advantage of unique strengths of the wheat genetic system. Moreover, it constructs a bridge between the wheat genome and other important grass genomes for fluid movement of information on gene location and function from other grass genomes to *Triticeae* and vice versa and facilitates exploitation of the high conservation of gene order among grass genomes (Benetzen and Freeling, 1997; Cohen, 1997; Briggs and Helentjaris, 1997; Devos and Gale, 1997) for comparative functional genomics. The observation of conserved gene content and

order in the genomes of grasses gave rise to the development of the model that individual grass species could be viewed best as the manifestations of a single grass genome and studies in each grass genome could generate data that can be utilized by other grass genome work. The most obvious use of this model is in understanding the nature of grass genome evolution. With the creative genius of evolution, minor changes in gene content, function, or regulation give rise to major new organismal properties. This can now be studied with much more accuracy across a range of species and times and we are much closer to the answer of the question of not only what the outcome is but also what it does and why it is like that. This approach can be utilized efficiently in the isolation and cloning of genes in large genome grasses by using the information from the small genome grasses. As a success story, the chromosome walking in rice has resulted in the isolation of a resistance gene in barley. This strategy is termed as microsynteny (Bennetzen and Freeling, 1997). Thus, the common gene content and similar physiological processes, information and genes can be applied across species boundaries. The comparative approach however relates only to genes. Intergenic regions are likely to have accumulated species-specific sequences and this avoids the prediction of physical distance between homologous genes in related species. Inversions, translocations and duplications are the main causes of large rearrangements that differentiate the grass genomes. But 60 million years of evolution has yet left large chromosomal segments untouched in the grasses (Devos and Gale, 1997). It is obvious that gene content and gene orders are highly conserved within the grass family while the amount and organization of repetitive sequences has diverged considerably. Isopycnic gradient centrifugation experiments of the cereal genomes, suggests gene-rich islands in the sea of poorly characterized repetitive DNA sequences. That would make new strategies developed to that which involve sequencing DNA fractions enriched for genes (Bevan and Murphy, 1999).

What exactly the genome of the ancestral grass was like is yet unclear though an awareness of wheat's ancestry is important for understanding both the range of genetic diversity in its primary and secondary gene pools and the potential for incorporating useful genes or gene complexes (such as those that confer disease resistance or stress tolerance) into cultivars. Cultivated bread and durum wheats descend from hybridized wild grasses. Durum wheat (*Triticum turgidum*), is a hybrid that occurred spontaneously between an as yet unidentified wild grass and einkorn, a primitive diploid wheat. Hexaploid bread wheat (*T. aestivum*), is an allopolyploid with genome constitution AABBDD, the product of a later spontaneous hybridization between a tetraploid and an unknown D genome diploid wheat (Devos and Gale, 1997).

1.1.3 Geographical Centers of Origin and Centers of Diversity

The concept of conservation has a long history in plant genetics and started with the Russian scientist N. I. Vavilov who is a pioneer in the classification of crop plants. He defined geographical "centers of diversity" for major crops based on patterns of observable variation in collections made from each species (1926, 1951). In recent years, studies based on DNA have shown that there is remarkable conservation in the sequences of the genes with similar function in different species (Devos and Gale, 1997). Vavilov's central idea was that the place of origin for a cultivated plant species is the area where it exhibits the greatest variation. The underlying assumption is that selective forces in the environment remain relatively constant throughout a species' evolutionary history. Vavilov recognized that his original hypothesis was a simplification. As scientists have now shown, selective forces are not constant, and as a consequence, within-species variation builds up at different paces in different places. Topography and geographical isolation play a role in how rapidly plant populations diverge genetically from one another. In addition, variation patterns are unique to the history and distribution of each crop and its wild progenitors. Some crops were domesticated several times, for example, others only once. Some spread early and developed secondary centers; others spread recently and can be traced to their origins by historical data. Vavilov's (1949-50) equation of centers of diversity of cultivated plants with places of origin, proved to be incorrect and is no longer used.

Early in this century, Vavilov and his colleagues found that the number of bread wheats in Europe and Siberia was already rather restricted. Anatolia, Syria, Palestine, and Transcaucasia contained other forms and types of wheat. Vavilov also defined several other centers of variation, including the Ethiopian plateau and the Mediterranean basin for durum wheats and Afghanistan for bread wheats (Harlan, 1975; Harlan, 1995).

More recently several disciplines, such as archaeology, botany, genetics, chemistry, anthropology, linguistics and recently molecular biology, have contributed to the study of the origin and spread of cultivated plants. However, different sources of evidence vary considerably in reliability. Thus, the modern synthesis is based on two major sources:

- Examination of plant remains coming from archaeological excavations and identifying what crop species they belong to. Radiocarbon dating and dendrochronology helps in understanding the age of the remnants in an acceptable accuracy.
- Analysis of living plants, in particular the wild ancestors of the crops. A major contribution is the identification of wild progenitors from which the various cultivated plants were derived. Since plant domestication is a relatively recent event, one can expect that most

wild ancestors are still alive and most probably include forms similar to those that existed in pre-agricultural times. Thus, the wild progenitors of most of the crop plants have been identified. The second step is the comparison of the cultivars with their progenitors in order to determine the main morphological, physiological, chemical, and genetic changes that took place under domestication. Assessment of changes in adaptation can also be evaluated. The classical taxonomic studies and cytogenetic analysis are the frequently used techniques before the developments in molecular genetics that now provide more sensitive and reliable new tools for determination of genetic differences within and between species (Zohary and Hopf, 1988). The most powerful molecular marker systems used in evaluating genetic variation in crops and solving problems in crop evolution are Simple Sequence Repeat DNA Polymorphisms (SSRs) or microsatellites and AFLPs.

1.1.4 Wheat Domestication and Human Civilization

Scientists generally conclude that the diploid grass progenitors of durum and bread wheats originated in the Fertile Crescent. New arrays of locally adapted cultivars evolved as wheat dispersed over vast geographic areas. Bread wheat is a relative latecomer among cereals, and its domination in cultivated area (replacing emmer and other wheat forms) in Europe and northern Africa did not occur until historical times. Although there were different ideas of origin of wheat domestication, until recently, due to archeological data and the diversity of wild varieties, generally the fertile crescent was accepted as the center of origin of wheat. The importance of this comes from the point that the place of wheat domestication reflects the site of the first settling of human populations, that is the transition from hunter-gatherer villages harvesting wild plants to farming villages planting fully domesticated plants (Harlan, 1975; Harlan, 1995). After only this prerequisite is met, the growth of large, dense, sedentary human populations, in which the food production of part of the population resulted in the storable food surpluses to feed the non-food producing parts of the population, namely kings, bureaucrats, professional soldiers, metal workers and other full-time craftspeople. Literacy, metallurgy, stratified societies, advanced weapons, and empires were developed starting first with food production. Thus a long straight line runs through world history, from those first domesticates in the Fertile Crescent, to the “guns, germs and steel” upon which the western civilization relies in expense of so many destroyed native societies of other continents by European colonists (Diamond, 1999). Thus location of the domestication of wheat means almost everything in human history due to the head start that is gained and this location was found to be the Karacadağ mountains of the Southeastern Turkey. Botanists already knew about the ancestor of the cultivated einkorn, a very similar wild cereal still growing in the natural habitats of the

crescent, but the location was not clear. In a recent study, 68 lines of cultivated einkorn (diploid *T. monococcum*, genomic designation AA), plus 261 wild einkorn lines from all over the crescent and outside this region were sampled and utilized in AFLP analysis. The result clearly indicated that the wild samples genetically most similar to the cultivated einkorn are the 11 samples collected from the Karacadağ mountains (Heun *et al.*, 1997; Diamond, 1997; Diamond, 1999).

1.1.5 An Historical Profile of Germplasm Flows

An important early period in the history of wheat germplasm, the Neolithic Dispersal (roughly 6000-3000 B.C.) was associated with the development of a stable food-producing economy in the Near East. This "unprecedented evolutionary expansion" resulted in a broad ecological differentiation as wheats adapted to different latitudes, altitudes, soil moisture regimes, and cultural practices. Traces of wheat in archaeological sites suggest that it spread into southern Europe between 5000-6000 B.C., into Egypt by 4500 B.C. (and probably earlier), the Netherlands by 4000 B.C., England and Scandinavia by 3000 B.C., the Indus cultures by at least 3000 B.C. (and evidence from Pakistan and Baluchistan suggests much earlier), and China by at least the second millennium (Zohary and Hopf, 1988; Harlan, 1995; Diamond, 1997; Diamond, 1999). After 3000 B.C., wheat populations continued to evolve and spread, in concert with changing human settlements and cultivation practices; however, the crop probably remained confined to the Afro-Eurasian landmass until about 1500 A.D.

Migrating farmers and governments seeking trade opportunities also spurred germplasm exchange. Transmitted as food in cargo holds and seed in settlers' sacks, wheat probably spread to almost all of the current production areas during the colonial period. Secondary flows occurred within colonial empires and commercial spheres for example, from Cape Town in South Africa to India and Australia; between South and North America; and among North Africa, East Africa, and other territories. At the close of the nineteenth century, scientific breeding programs began to develop throughout the wheat-producing world.

1.2 MOLECULAR MARKERS

Recent developments in the field of molecular genetics (molecular markers, marker mapping etc.) has effected all the areas of modern biology, from ecology to systematics, from transgenics to developmental biology, from breeding to genomics and physiology. Now that geneticists have the capacity to isolate and clone genes, the next step would be the vital cooperation that would be build with the physiologists to address the connections of the genes and the characters that they are coding. The first step of finding the genes of interest is on its way in numerous laboratories in all around the world. While occasionally a gene is identified in its function, we cannot usually do anything with the gene until we map it or locate it in a linkage group. That is exactly where the molecular markers come into scene (Jones *et al.*, 1997). The most important DNA-based molecular markers include:

1.2.1 Restriction Fragment Length Polymorphisms (RFLPs)

Unlike the classical morphological markers, the variations in the molecular ones are not detected in the phenotype. These are usually single nucleotide changes in a repeating sequence or in a gene. Their main advantage is that they are numerous and they do not effect the physiology of the organism. For example, restriction fragment length polymorphisms(RFLPs) are generated with the constant mutations in single nucleotides that change the restriction sites and then the RFLP pattern of the hybridized probe (Botstein *et al.*, 1980). The large number of markers are the main advantage over the classical morphological markers and isoenzymes (Jones *et al.*, 1997).

1.2.2 Variable Number Tandem Repeat (VNTR)-Hypervariable Minisatellites

A labeled probe is needed in order to detect polymorphism with RFLPs. Another well-known probe-based marker is the variable number tandem repeat (VNTR), or minisatellite which was discovered in human for the first time by Alec Jeffreys as early as 1980s and was used as a revolutionary tool in forensic DNA profiling (Jeffreys *et al.*, 1985). Tandem arrays of short repeated sequences in VNTR are highly dispersed throughout the genome at different loci. These are embedded in unique flanking sequences, and the loci are hypervariable in repeat units. A 'polycore' probe, prepared for a consensus sequence in different loci, can be used to detect up to 30 loci simultaneously and a detailed DNA fingerprint can be generated. Minisatellites are mainly useful in vertebrates but plant species can also be used as template to

produce low resolution fingerprinting in cultivar identification (Jones *et al.*, 1997). The minisatellite size changes during meiosis have been characterized in sequence level and it was suggested that degradation of the “recipient” DNA occurs during gene conversion, which is not easily detected with the systems of recombination investigations. Recombination events exist that do not lead to any change in size, which is again difficult to detect in mammalian systems. Thus, it was suggested that the mammalian minisatellites are even more active than previously believed. In the studies with minisatellite variants generated in yeast meiosis involving DNA removal during gene conversion, demonstrate a clear association between minisatellite size changes and recombination events like crossovers and gene conversions, in the flanking DNA region (Bishop *et al.*, 2000).

1.2.3 Randomly Amplified Polymorphic DNA (RAPD) Markers

From the PCR-based markers, the randomly amplified polymorphic DNAs (RAPDs) are used with arbitrarily chosen 10mers as primers, and several different regions of a genome are amplified. Statistically, priming occurs once in every million base pairs. Generally, 5-10 fragments are amplified from an average-sized genome and polymorphism is due to the sequence variation that changes the sites of primer binding. Thus, RAPDs are dominant markers (presence/absence) and segregate according to Mendelian genetics from a heterozygous diploid. The advantages of RAPDs include their low application cost (no need for probe labeling or prior sequence information), their simplicity, due to the different primer possibilities the infinite number of RAPDs possible. Application of RAPD markers involve fewer stages and is less time-consuming than RFLP analysis. RAPDs can be used for mapping but basically as species-specific markers (due to short primers used and inconsistent results not suitable to use between species) for diversity and phylogenetic studies. The disadvantages of RAPDs include their poor reliability and reproducibility and their extreme sensitivity to the experimental conditions (Williams *et al.*, 1990; Welsh and McClelland, 1990; Jones *et al.*, 1997).

1.2.4 Microsatellites

Microsatellite markers, also referred as Simple Sequence Repeat (SSR) DNA polymorphism, are among the most informative molecular markers for eukaryotic organisms. They are locus and genome specific, co-dominantly segregating, randomly and abundantly present (five to ten times more frequent than random sequences), highly polymorphic and

reproducible markers (Weber and May, 1989; Litt and Luty, 1989; Tautz, 1989; Kashi *et al.*, 1997). They were found near genes, in some introns of genes, within variants of repetitive Alu-elements and satellite sequences, as well as other nonfunctional genomic regions. That is why it was suggested that they arise by common mechanisms called slippage replication and unequal crossover and that they might have no function in gene expression (Tautz and Renz, 1984). However, the abundant (CA)_n repeat blocks in human genome (50,000–100,000 interspersed (CA)_n blocks with the 'n' ranging between 15–30) was suggested to serve as hot spots for recombination (Weber and May, 1989). Direct evidence on their function as transcriptional activating elements were found in the studies where upstream promoter regions containing SSRs enhance transcriptional activity in expression vectors, while constructs with deletions in the same promoter regions, lacking the SSR sequences, have their activity reduced considerably. There is a high rate of spontaneous mutation in the repeat number of SSRs, the variation of which is expressed in the high frequency of appearance of new alleles. Since they are highly polymorphic within species and populations, they become valuable tools of genome maps and kinship studies. Their repeat-number variation can cause quantitative changes in gene expression and function. Studies on human degenerative disorders have shown that functional variation due to SSR-length difference have phenotypic consequences, specifically, there is a strong correlation between trinucleotide-repeat length and age of onset and in some cases severity of disease (Kashi *et al.*, 1997). Requirement of sequence information to produce locus specific primer sets for PCR amplifications is a major limitation for their application in many organisms and for most none or insufficient numbers of markers are available. Also, for most organisms, a large number of microsatellite markers have already been reported in the literature, but in order to obtain highly condensed maps or to find close linkages to the important characters (genes of interests), more markers need to be isolated.

SSRs or microsatellites are tandemly repeating core (≤ 5 bases) sequences that are utilized as valuable DNA-based markers in numerous eukaryotic genomes (Wang *et al.*, 1994). (CT)_n and (GT)_n type of dinucleotide repeats are the most abundant ones in mammals (Hamada *et al.*, 1982) while (A)_n, (AT)_n and (GA)_n are most common in plants (n ranges from 10 to 80). SSRs are widely interspersed in human genomic DNA and show high frequency of variation in the number of repeats in different individuals or accessions, that is most probably due to slippage during the replication of DNA. Tautz and Renz, 1984; Tautz *et al.*, 1986; and Akkaya *et al.*, 1992, have named the repeated sequences as simple sequence repeat (SSR), or short tandem repeat (STS) and Litt and Luty, 1989, as microsatellite.

Co-dominantly inherited and multiallelic SSR markers are valuable for evolutionary studies (Schloetterer *et al.*, 1991; Buchanan *et al.*, 1994), regarding genetic relationship and differentiation of closely related species.

The tandemly repeated di-, tri-, tetra- or penta-nucleotide multiallelic sequences are abundant and widely distributed in the genomes of several plant species so far studied. While (AT)_n repeats ('n' is the number of repeating units within the simple sequence repeat (SSR) or microsatellite locus) are the most abundant ones, they may also be dinucleotides (AC)_n and (AG)_n; trinucleotides (TCT)_n, (TTG)_n; tetranucleotides (TATG)_n and so on.

The PCR amplified (by locus specific primers at the each conserved end of repeating units) SSR products are separated on the sequencing gels and visualized by either radioactive or fluorescent labeling, and by silver staining as well. The length variation of PCR products is a function of the number of SSR core units, which is the result in polymorphism of a particular locus. The high level of SSR polymorphism was first demonstrated in human genome (Litt and Luty, 1989; Weber and May, 1989; Tautz, 1989).

The high level of microsatellite polymorphism in plants and the fact that "AT" repeats are the most abundant di-nucleotide repeats in plants were first demonstrated in soybean by Akkaya *et al.* in 1992.

In recent years microsatellites have been developed and used successfully in numerous molecular genetics laboratories all around the world in rapidly increasing number of plant species. The following species are among them; soybean (Akkaya *et al.*, 1992; Morgante and Olivieri, 1993; Rongwen *et al.*, 1995; Diwan and Cregan, 1997), wheat (Roder *et al.*, 1995; Devos *et al.*, 1995; Bryan *et al.*, 1997; Korzun *et al.*, 1997; Peil *et al.*, 1998; Stephenson *et al.*, 1998; Shan *et al.*, 1999; Dograr *et al.*, 2000; Prasad *et al.*, 2000; Li *et al.*, 2000), wild wheat (Pestsova *et al.*, 2000), barley (Liu *et al.*, 1996), rice (Wu and Tanksley, 1993; Zhao and Kochert, 1993; Yang *et al.*, 1994), maize (Senior and Heun, 1993; Taramino and Tingey, 1996), brassica (Lagercrantz *et al.*, 1993; Szewc-McFadden *et al.*, 1996), arabidopsis (Bell and Ecker, 1994; Innan, 1997), potato (Provan *et al.*, 1996), beans (Yu *et al.*, 1999), grape (Thomas and Scott, 1993), sunflower (Brunel, 1994), avocado (Lavi *et al.*, 1995), tomato (Bredemeijer *et al.*, 1998), cabbage (Raybould *et al.*, 1999).

By providing high level of polymorphism information, microsatellites are ideal markers for wheat species (i.e. mapping agronomically important genes). They are abundant and distributed evenly over the genome and small amounts of genomic DNA is adequate for genetic

analysis. The analysis of microsatellites can easily be automated. The main drawback for microsatellites is the prior sequence knowledge requirement in order to design locus specific primers. In this respect, due to the large and complex genome (AA, BB and DD) of these plants, the microsatellite development for wheat is a very time consuming and expensive procedure (Röder *et al.*, 1995; Plaschke *et al.*, 1995; Bryan *et al.*, 1997; Röder *et al.*, 1998). Another disadvantage is the low percentage of functional SSRs (~30% of all primer pairs developed). Nonfunctional primer pairs on the other hand result in either a smear (due to the large number of fragments), no product or fragments of the incorrect size (usually monomorphic). Optimization of PCR conditions or designing new primers do not significantly improve the results for the unsuitable ones (Röder *et al.*, 1995; Bryan *et al.*, 1997).

The level of polymorphism is assessed by the use of the difference between the number of repeats in a PCR amplified specific locus. Since the marker information is valid for the two alleles the marker is co-dominant. Though the molecular mechanisms of new allele generation is unclear, replication slippage, sister chromatid exchange, unequal crossing over, and gene conversion are the suggested causes of variation (Tautz and Renz, 1984). The application of the marker is quite simple (PCR-based) and the marker is highly informative (due to the number and frequency of alleles detected and to their ability to distinguish between closely-related individuals).

Microsatellites can be utilized as markers for mapping, determination of hybridity, cultivar identification, protecting germplasm, analysis of gene-pool variation and as diagnosing traits of economic importance (Goodfellow, 1993). As their disadvantages, the establishment of SSRs are expensive, they have long development time and they need unique primers. In this study, however, we developed a new strategy that will effect the time and cost expense of these markers (see below).

1.2.5 Amplified Fragment Length Polymorphism (AFLP)

AFLP is based on selective amplification of digested genomic DNA by a series of extended primers. The genomic DNA is digested with two enzymes, a frequent and a rare cutter, e.g. *MseI* and *EcoRI*. Adaptors, with a core sequence and the appropriate overhangs for the enzymes used are ligated to the ends of the fragments. Pre-amplification of these fragments by PCR is achieved using primers with a core sequence, an enzyme specific sequence and a random-extension nucleotide. This generates a pool of fragments which are re-amplified by two primers of core sequence, enzyme specific sequence and this time further random-

extension nucleotides. The primer with the rare cutting enzyme sequence is end-labelled before the reaction so that those fragments only will be detected by electrophoresis and autoradiography (Vos *et al.*, 1995). Fragments generated by *Mse*I digestion will be small enough to amplify by PCR, while those with one *Mse*I and one *Eco*RI end will be much rarer but will also easily amplify. During the second amplification, further selection of fragments occurs using the extension nucleotides and finally only those fragments generated by the *Eco*RI-labelled primer will be resolved visually by autoradiography following polyacrylamide gel electrophoresis (Figure 1.1). The use of two enzymes also enables the AFLP reaction to be fine-tuned for the production of the maximum number of bands. This technique was originally developed by plant geneticists but can easily be transferred to the analysis of other genomes. The technique was applied to soybean (Maughan *et al.*, 1996; Keim *et al.*, 1997), poplar (Cervera *et al.*, 1996), sugar beet (Hansen *et al.*, 1999), barley (Qi and Lindhout, 1997; Powell *et al.*, 1997; Schwarz *et al.*, 1999), wild barley (Pakniyat *et al.*, 1997), wheat (Shan *et al.*, 1999), rice (Maheswaran *et al.*, 1997), *Oryza* species (Aggarwal *et al.*, 1999), many more different plants. The development of many polymorphic bands for each reaction should enable large numbers of marker loci to be mapped for a relatively small amount of work and input DNA. However, all these markers will be anonymous, and to clone one requires the extraction of DNA from a gel matrix.

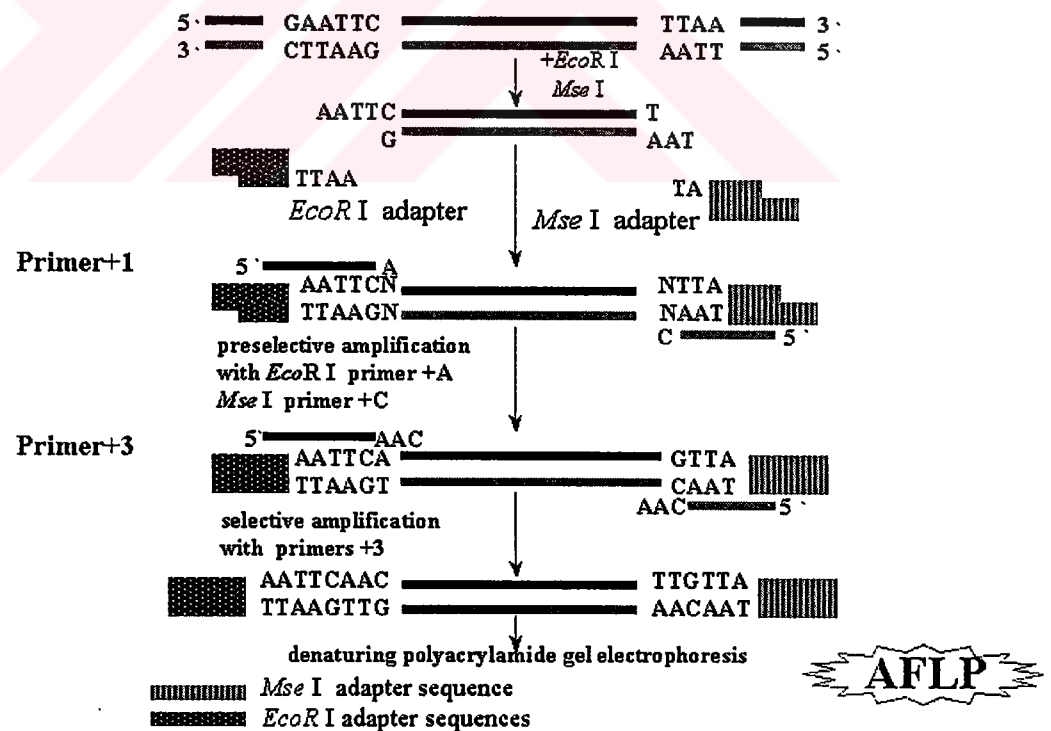


Figure 1.1 AFLP procedure

1.2.6 Comparisons of DNA-based Molecular Markers

The performances of the major molecular markers are compared in several plant species. The four types of marker assays in Table 1.1 have different properties: While SSRs have the highest expected heterozygosity, AFLP have characteristic high multiplex ratio. RAPDs are intermediate in both heterozygosity and multiplex ratio. RFLPs are moderate in heterozygosity but are appropriate for syntenic studies. In the comparison of AFLP, RAPD and RFLP assays for measuring genetic diversity in melon, AFLP was found the most suitable for cultivar fingerprinting in cucurbit species (Garcia-Mas *et al.*, 2000). AFLP markers were also found appropriate for cultivar identification in barley (Schut *et al.*, 1997) and for genetic diversity studies in wheat (Bohn *et al.*, 1999). Pejic *et al.* (1998) analyzed the genetic similarity among maize inbred lines using RFLP, RAPD, SSR and AFLP markers and concluded that both the SSR and the AFLP technologies can replace RFLPs. Ease of genotyping and the level of informativeness are the most important criteria in the choice of the marker system. The specifications and advantages/disadvantages of the most widely used DNA-based molecular markers are compared in Table 1.1. SSR and AFLP are the best applicable and highly informative and reliable markers. SSR have dominated the polymorphism research in mammalian genomes and they are likely to have major impact on plant genetics as well since they provide a highly informative, co-dominant PCR based assay which is compatible with the requirements of plant breeding and population genetic programs (Powell *et al.*, 1996).

Table 1.1 Comparison of DNA-based molecular marker systems

Specifications	Marker System			
	AFLP	RFLP	RAPD	SSR
Required DNA content (µg)	0.5–1.0	10	0.02	0.05–0.10
PCR-based	Yes	No	Yes	Yes ^a
Polymorphism (Hav ^b)	0.11	0.28	N/A ^c	N/A
	0.29–0.64	N/A	0.33–0.34	0.47–0.76
	0.32	0.41	0.31	0.6
	Medium	High	Medium	Very High
Effective multiplex ratio ^d	50–100	N/A	20–50	1.0
	19.2	0.25	1.56	1.0
	24.7	3.0	N/A	N/A
	6.14	0.1	0.48	0.6
Marker index (MI ^e)	6.14	0.1	0.48	0.6
Reproducibility	Very High	Very High	Medium ^f	Very High
Cost	N/A	N/A	2x AFLP	N/A
	\$105 ^g	178	N/A	N/A
	Medium	N/A	Low	High ^h
Application	Difficult in first use	Difficult	Easy	Easy ⁱ

^aSequence information required. ^fBetween laboratories

^bHav, average heterozygosity. ^gGain of time comparing RFLP

^cN/A, data not available. ^hFirst sequencing expense is high

^dNumber of polymorphic loci in tested samples per experiment. ⁱBy practice/time it becomes easy

^eAverage expected heterozygosity and effective multiplex ratio ^jSome problems in silver staining

1.3 OUTLINE OF THE THESIS

Considering the flora and fauna richness of Anatolia, with special emphasis on wheat (gene center of which is Turkey) due to its major role in the nutrition of Turkish people, advanced technologies should be applied and improved in characterization and effective utilization of these natural offerings. The outline and the content of this thesis research are summarized in the followings.

Within the context of the thesis, we applied a novel strategy for the development of microsatellites that will probably reduce the development cost and time consumption when the current applied standard methods are considered. The details of the method are explained in 1.4.

Microsatellite marker systems is one of the most efficient and reliable marker systems currently available and utilized in most of the relevant advanced laboratories around the world. In the section 1.5, application of SSR in the characterization of doubled haploid Turkish durum wheat varieties are presented.

In the section 1.6, information on disease resistance genes and gene analogs from the plant species are presented. Also, the isolation of disease resistance gene analogs from wheat varieties are explained.

All the materials and methods applied in the thesis are presented in Chapter II, in the same order as it is presented in the Introduction.

Chapter III includes the results of the studies on the novel AFLP-based microsatellite isolation first, followed by the application of microsatellites in doubled haploid Anatolian durum wheat varieties. Finally, results on the isolation, sequencing and genetic analysis of disease resistance gene analogs from wheat are included.

A general conclusion on the thesis is presented in the final part, Chapter IV.

In addition, all the RGA sequences obtained and their alignments with each other are presented in Appendices I to VI.

1.4 A NOVEL METHOD DEVELOPED FOR MICROSATELLITE ISOLATION

Database searches and sequence analysis was the primary source of finding microsatellite sequences (Weber and May, 1989; Akkaya *et al.*, 1992, 1995; Yang *et al.*, 1994; Devos *et al.*, 1995; Russell *et al.*, 1997). A common yet a time consuming microsatellite isolation approach has been the construction of libraries and screening *via* the colony hybridization using probes of short oligonucleotides of tandem repeat units, followed by the isolation of plasmid (Akkaya *et al.*, 1992, 1995; Cregan *et al.*, 1994; Sosinski *et al.*, 2000) or phage DNA (Röder *et al.*, 1995; Chen *et al.*, 1997; Pfeiffer *et al.*, 1997; Struss and Plieske, 1998) from the clones putatively possessing the repeat regions for sequencing. The positive clones are then sequenced and specific primers that flank the core repeat are designed for the locus of interest.

There were several attempts on our part to improve the microsatellite isolation procedures, since screening plasmid or phage libraries by colony hybridization with radioactively labeled oligonucleotide probes are time consuming and non-cost-effective. The disadvantage of this approach is because of the repetitive genome of wheat, we could only expect to obtain 1 microsatellite locus in every 100 kb or even more. Thus, it was decided that the genome should be enriched in terms of interest of di- or tri-nucleotide repeats. Enrichment of libraries in terms of targeted microsatellite core sequences were performed in several laboratories (Becher *et al.*, 2000; Prasad *et al.*, 2000; Fischer and Bachmann, 2000, 1998; Edwards *et al.*, 1996). There has been considerable effort in this study for the isolation of wheat microsatellites with various approaches. In the first procedure applied, which is referred as triple affinity capturing (TAC), target DNA was captured with a biotin labeled (GA)_n oligonucleotide, by utilizing the triple helix formation ability of DNA under mild acidic conditions. Streptavidin-coated magnetic beads were used to eliminate the unrelated DNA while keeping the target triplex apart. Enriched DNA was recovered in double stranded form by elution with a mild alkaline buffer destabilizing the triplex (Ito *et al.*, 1992). In our trials of constructing genomic DNA libraries, we have enriched the enzyme (*Hind*III) digested wheat genomic DNA, site selected at ~500 bp region, by carrying out the TAC methods.

A colleague from Bulgaria, Dr. Karagyozov, had developed a small-insert wheat genomic library after double affinity capturing enrichment (Figure 2.2) in terms of “GA” and “CA” repeats as a collaborative effort. We have sequenced screened and selected clones putatively containing the tandem repeats, generated in his lab. The contribution of this thesis work has been the sequencing of the microsatellite loci and the designing of the appropriate primers.

During the study, we have come up with a very new approach for the microsatellite marker development. It is an alternative method for isolation of microsatellite loci from any organism and eliminates labor-intensive library constructions, and the colony hybridization for selection of the putative positive clones (containing microsatellite repeats) and the maintenance of the colonies. The procedure makes use of AFLP marker generation (Zabeau and Vos, 1993; Vos *et al.*, 1995) and an enrichment step *via* double heteroduplex complex formation. Considering many research groups using microsatellite markers are most likely also be applying AFLP markers in genetic studies, this approach can be a practical one for microsatellite isolations with already available technical infrastructure. AFLP markers were obtained with various Eco+3 and Mse+3 selective amplification primer sets as in performing AFPL with only exception that the amplification is non-radioactive (Figure 1.1). The products then were enriched for each selective primer combination using a biotinylated target repeat oligonucleotide on the streptavidin coated magnetic beads. The enriched fragments were eluted separately as each AFLP selective primer combination product, and then were re-amplified in the presence of a radioactively labeled nucleotide, [α - ^{32} P]dATP, using the same set of AFLP selective primer combination (Figure 1.2).

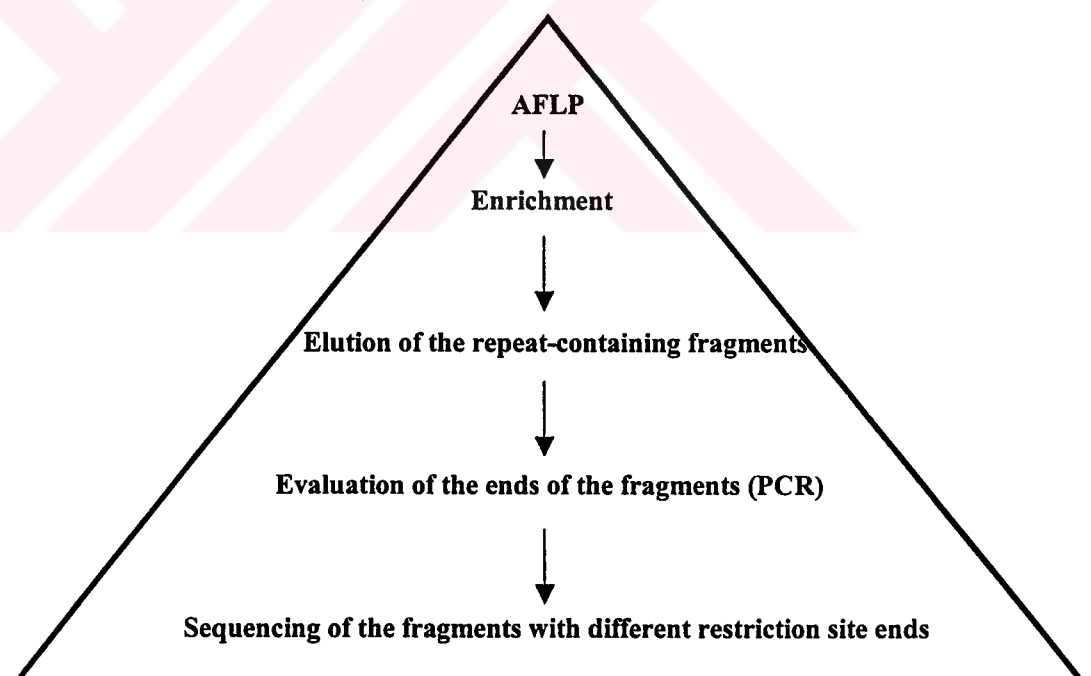


Figure 1.2 Overview of the isolation steps

1.5 GENOTYPING OF ANATOLIAN DOUBLED-HAPLOID DURUM LINES WITH SSR MARKERS

The doubled haploid (DH) lines generated from the durum wheat varieties, that are selections from Middle Anatolian landraces, 'Çakmak-79', 'Berkmen-469' and 'Kunduru-1149' (Savaşkan *et al.*, 1997), were analyzed using fifteen microsatellite primer sets (Röder *et al.*, 1995 and 1998) for genotyping and evaluation of genetic relationships between and within the doubled-haploid lines. However, only ten primer sets were polymorphic (two of them amplifying to different loci) in our samples and thus eligible in the analysis.

'Kunduru-1149' and 'Berkmen-469' are old genotypes of Central Anatolia, which were improved by selection from local populations based on yield increase. Turkish farmers are still using 'Kunduru-1149' for production due to adaptability and high yielding natures. 'Çakmak-79', is a cross of 'Üveyik-162' (also Central Anatolian landrace selection) and line '61-130' (US origin, semi-dwarf durum wheat). All three genotypes are facultative. The genetic relationship of doubled-haploid individuals of three-durum wheat cultivars, 'Çakmak-79', 'Berkmen-469' and 'Kunduru-1149' was investigated.

Characterizing the genetic distances of DH lines within and between populations will be valuable for assessing the genetic diversity, thus, facilitating the use of these lines in breeding programs. Also, by adding more cultivars in the analysis, a more informative genetic relationship on the Turkish durum wheat varieties will be obtained. The progenitors of the DH lines as well as other Turkish varieties are incorporated in the final analysis.

1.6 ANALYSIS OF DISEASE RESISTANCE GENE ANALOGS

1.6.1 Resistance Breeding

Protection of crops against the attacks of detrimental pathogens is the major area of research in breeding. While insects and bacteria are also effective, fungal infections are the major disease causing agents during all the crop production history. The control of fungal pathogenesis today involves, the husbandry procedures (like the rotation of planted crops and preventing infected plant material and soil), utilizing resistant cultivars and application of agro-chemicals. Classical plant breeding procedures have generated resistant cultivars to major pathogens, but the reaction of the fungi by developing resistant races was followed. Development of resistance with conventional breeding strategies takes ten to fifteen years; however, it takes only few years for the fungi to develop new virulent races. The breeder then should start all over again to develop new resistances. As a second strategy applied, modern agricultural practices utilize extensively the fungicidal chemicals. However, there are artifacts of high cost and harmful effects of these chemicals on the environment. Moreover, the pathogens evolve rapidly to develop resistance, just like the antibiotic resistance of human pathogens. At this point the advanced molecular genetic techniques come into scene with powerful tools to engineer resistant cultivars in meaningfully short time periods even allowing plant scientists cross species barriers in introducing genes from unrelated organisms. That is isolating the necessary genes from whatever organism available and introducing them to the target crop with transformation techniques.

Enhanced lignification and cell wall cross-linking, production of antimicrobial compounds, like phytoalexins, controlled cell death at the infection site and production of reactive oxygen species are among the inducible defense mechanisms of plants. Systemic acquired resistance (SAR) that involves the enhanced resistance in yet uninfected tissue is also an effective strategy of the plants against viruses, bacteria and fungi. Pathogenesis-related (PR) gene expression is also accompanied with SAR response. It is speculated that PR peptides are very effective in the defense of plants against fungi. The strategy of antifungal compounds aims the expression of these proteins constitutively in transgenic plants. In this respect, the expression of chitinase in transgenic tobacco enhanced the resistance against *Rhizoctonia solani*. The resistance level was considerably enhanced when certain combinations of PR genes, like chitinase/glucanase, were used. Similarly, promising results were obtained from the first monocots transformed, namely the 'Aly AFP' and 'Rice Chi' transgenes developed resistance against *Fusarium sp.* in wheat and *Rhizoctonia solani* in rice, respectively (Melchers and Stuiver, 2000).

Plant defensins, thionins, hevein, lipid transfer proteins are among the other plant antifungal peptides that have the potential of being used in combination that may be effective against a broad range of fungi (Melchers and Stuiver, 2000). A summary of events that are detected during the host-pathogen recognition and the possible responses of the plant cells are presented in Figure 1.4.

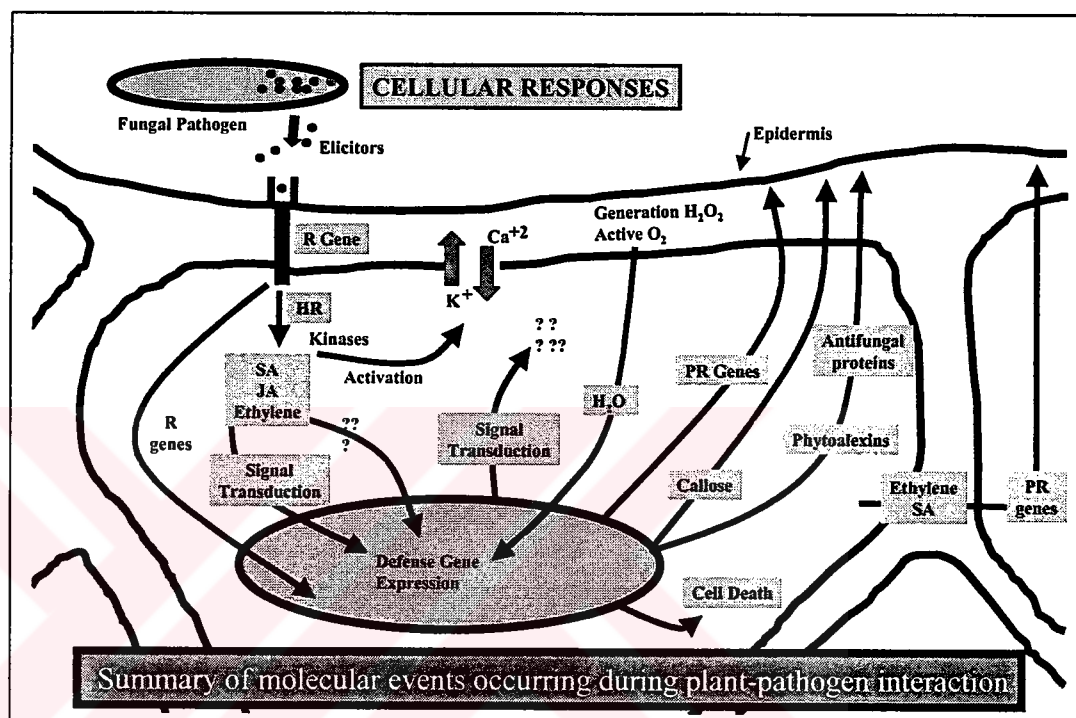


Figure 1.4 The responses of the plant host cell to the pathogen elicitors. HR: hypersensitive response; SA: salicylic acid; PR: pathogenesis related.

1.6.2 Isolation of Resistance Genes

During the decade of 1990's a number of disease resistance genes have been isolated from different representative plant species like tomato, arabidopsis etc. To date, from the monocot species only rice Xa21 disease resistance gene has been isolated, all other genes belonging to the dicotyledonous plants. The method of choice was map-based cloning. As it is seen in Figure 1.5, the walk begins with a clone containing *mkrB*. The ends of the clone (boxed) are used to probe a library. Clones from adjacent genome segments are thus identified and isolated. The distal ends of those clones are used to reprobe the library. These steps are continued until a clone contains either *mkrA* or *mkrC* sequences. Clones between *mkrB* and

mkrC must then be evaluated for the presence of *yfg* (<http://opbs.okstate.edu/~melcher/MG/MGW4/MG412.html>).

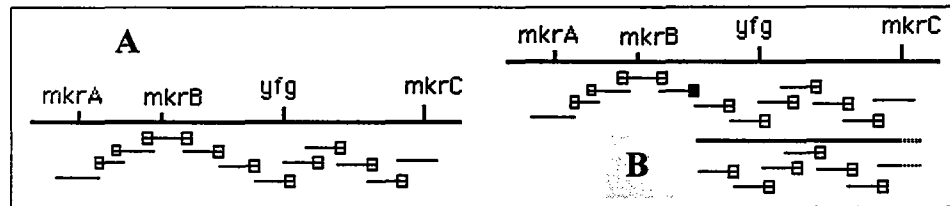


Figure 1.5 A. Chromosome walking-jumping. B. Chromosome walking interruptions in repeating sequences.

While the method is suitable for gene isolations from simple-genome plants (arabidopsis, tomato etc.), this is especially impossible for genomes with high level of repetitive elements like that of wheat which has >80% repeating DNA sequences. Chromosome walks can be detoured in many directions if one of the clone end probes (filled box) is a repeated sequence (Figure 1.5).

This drawback in gene isolations from complicated genomes is going to be overcome by alternative genomic approaches. A promising alternative is homology-based PCR amplification method. Resistance gene analogs are amplified from a couple of plant species to be used in probing the whole resistance genes.

1.6.3 Resistance Gene Analogs

In the new era of genomics new tools has started to be utilized by the plant molecular genetics community. Research on plant disease resistance is not an exception and new understanding on gene isolation and functional research has been targeted in numerous laboratories around the world. With the accumulation of considerable sequences related to the resistance genes (R-genes) (mainly through classical map-based cloning and transposon tagging strategies) a new approach, based on homologies, allowed the identification of candidate R genes. Acquiring and processing large amounts of biological data, in respect of comparative genomics, has been facilitated with the advances in bioinformatics (Baxevanis and Ouellette, 1998; Bisby, 2000; Edwards *et al.*, 2000), and all these serve on novel gene isolations by utilizing available information. Most of the generated data belong to few model organisms like

arabidopsis, rice, tomato, which are representative plant species. As the technologies improve and cost relatively decrease the aim of sequencing large portions of the genomes of important crop plants are on their way (Michelmore, 2000). This seems to be facilitated by the novel gene-rich island identifications within genomes. Sequence information from the dicot species *Arabidopsis thaliana*, a small genome model plant, has shown a gene density level of a gene per 4-5 kb. The genomes of the grasses on the other hand are variable in size and in repeating sequences (rice genome being 415 Mb and that of wheat being 16,000 Mb in size). Although there is such a huge genome size difference, the number of the genes are quite close to that of small genome plant species, the size difference reflecting repeating sequence accumulations (up to 80% of the whole genome in wheat). Sequence analysis of large genomic fragments from maize, rice, and sorghum genomes have suggested that gene density changes in accordance with genome size but this information is based on limited data available. In wheat, for instance, deletion mapping studies on the distal end of the chromosome group 1 have demonstrated the gene-islands. In another study, 23 kb genomic fragments around the receptor-like kinase gene *Lrk10* (mapped to the short arm of wheat chromosome 1A, located in a very distal region on this chromosome arm with a distance of 8 cM to the (Glu-3)-1A (Feuillet *et al.*, 1997)) were isolated and compared in the genomes of wheat, barley, maize, and rice. The presence of regions rich in genes (different receptor-like kinase gene families) in different grass genomes were demonstrated. In the chromosome group 1 of Triticeae, four receptor-like kinase genes (*Lrk/Tak*) and six NBS and LRR region genes were identified. In the species used, the relative positions of *Lrk* and *Tak* (*Triticum aestivum* kinase) to each other were identical but their orientations were opposite to each other. This organizational conservation makes us think of the presence of a selection pressure during evolution (Feuillet and Keller, 1999). Resistance gene clusters most likely evolve more rapidly than other regions of the genome and the motifs that are related to the signaling and recognition events (such as leucine-rich repeat (LRR)) are subject to adaptive selection. The function and the evolution of such domains have recently been analyzed by nucleotide substitution characterizations. The ratio of nucleotide substitutions that result in amino acid replacements (dn) and substitutions that do not alter the amino acids (ds) are specifically informative. In consistence with the function of the proteins (with amino acid replacements) the ds/dn ratio is usually <1.00 in most protein coding genes while a ratio that is significantly greater than unity indicates the divergence between genes via adaptive selection. The available data on adaptive selection is limited but it is most probable that they are more common in regions functioning in host/pathogen recognition. It can be speculated that the ligand binding regions will be subject to stronger adaptive selection than the regions related to structure. A comparison of nucleotide substitutions in the LRR coding regions of Xa21 and Xa21D have shown that there is a 99.1% sequence identity but substitutions that result in amino acid replacements are significantly greater than the

substitutions that do not lead to amino acid replacements in the LRR regions. These observations indicate that the region has a functional importance in ligand binding. The dn/ds ratio is greater than 1.00 which indicates the role of LRR in race-specific pathogen recognition and this is subject to adaptive evolution. The high level of sequence variation observed among resistance gene family members suggest that gene conversion and/or unequal crossing over is sufficiently low as to prevent homogenization of the family members. While most models predict that resistance genes possessing novel ligand-binding capabilities arise from recombination events, there is also a predicted novel ligand specificity acquisition *via* single amino acid substitutions in the solvent-exposed surface of the LRR (Richter and Ronald, 2000).

In disease resistance loci of plants, it is already well-established that there is a clustering of different genes working on the same signal transduction pathway (for instance the tight linkage between the *Pto* and *Pfr* genes in tomato). The distance between these genes, in wheat and barley, were varied from 450 bp to 7 kb. In rice genome the maximum distance in this region was determined to be 2 kb. This information of genome size independent gene-rich islands is important in using model species in cross-genome map-based cloning. There are limitations however, due to the rearrangements in collinear regions. It was suggested that, using a model species which is the most close to the genome of interest, that is using barley as a model instead of rice while looking into wheat genome for instance, is the most logical approach.

Receptor-like kinase genes were found together with NBS-LRR protein encoding genes in disease resistance loci leading us to think about the effective resistance expression in the presence of both type of genes in a region. It is already known that the *Pto* kinase and *Pfr* are NBS-LRR genes in tomato. They are effective against *Pseudomonas syringae* when expressed together (Feuillet and Keller, 1999).

The majority of the disease resistance genes so far cloned from different organisms have the nucleotide binding site (NBS) at the N-terminal of the leucine-rich region. The two groups distinguished in NBS are: those with toll/interleukine-1 receptor homology N-terminal to the NBS (namely the toll-interleukine receptor or TIR group) and those of non-TIR group. While resistance genes *N* from tomato, *M* and *L6* from flax, and *RPP5* from *Arabidopsis*, are belong to the TIR class, the non-TIR one includes the products of resistance genes *RPS2* and *RPM1* from *Arabidopsis*, and *I2*, *Mi* and *Prf* from tomato and *Dm3* from lettuce (http://www.ncgr.org/research/rgenes/rgenes_data1.html).

Evidence on the numbers and organization level of disease resistance genes are coming from the rice and *Arabidopsis* genome sequences. About 14% of the genes in 1.9 Mb *Arabidopsis* genome was suggested to be related to disease resistance related genes by signaling components or anti-microbial proteins. *Arabidopsis* genome is estimated to have about 200 NBS-encoding genes (at the present ~50% of its genome was analysed and 120 genes with similarity to NBS harboring plant R-genes were detected) ~150 of which are TIR class and remaining ~50 non-TIR class. This represents 1% of the whole *Arabidopsis* genome and the genes were found to be highly clustered. The rice genome on the other hand (data from 5% of the genome) was suggested to have about 750-1500 NBS encoding genes, all non-TIR class. None of the genomic or EST sequences from any grass species so far sequenced represent TIR-class which is very interesting in terms of evolutionary studies on the diversion of dicotyledonous plants and monocots.

From the clusters of resistance genes so far sequenced the *RPP5* (NBS-LRR type protein gene) from *Arabidopsis*, the *Cf-4/9* (LRR-transmembrane-type gene) and *Pto* (Protein kinase gene) in tomato, are sequenced completely and *Dm3* (NBS-LRR type gene) is partially sequenced. Homologs (~10) of *RPP5* are scattered over a region of ~90 kb, within pseudogenes and retrotransposons while the *Cf-4/9* cluster with five resistance genes spread over 36 kb. The *Pto* cluster with five homologs are spread over 60 kb region. The *Dm3* in lettuce is the largest cluster with 24 resistance gene homologs scattered over a 3.5 Mb region.

In spite of the fact that the most abundant group is NBS-LRR class there is limited information on their function. The motives they harbor give clues on their role in the beginning of signal transduction pathways. This suggests their functions (development and response to abiotic environment) other than disease resistances. The present data indicate their constitutive expression at low levels (Michelmore, 2000).

It was determined that different plant disease resistance genes have some common properties. When plants are repeatedly infected with a certain pathogen they first develop an hypersensitive response (HR) and thus avoid the development of the pathogen. This is performed through the specific receptor molecules produced by the resistance genes of the plant against the avirulence signal peptides produced by the pathogen specifically. It is thought that HR and other systemic acquired disease resistance due to necrotic reactions (Systemic Acquired Resistance (SAR)), is generated nonspecifically. R genes in receptor-ligand model in recognition and resistance can be an extracellular receptor like that of the *Cf-9* gene in tomato as well as an intracellular receptor like that of N resistance gene in tobacco. The ligand in this model can be a direct or indirect product of pathogen avirulence gene. It is thought that the

specific recognition event develops *via* the activation of a signal transduction pathway where resistance mechanism is generated with the role of protein kinases. A more general defense mechanism including mammals and the insects operates with a signal transduction pathway against fungal, bacterial and viral pathogens and R genes have common structural domains. The common characteristics of all five types of R genes are that they harbour a "leucine-rich" region (LRR) and a 'serine-threonine kinase' domain. These groups of R genes are: 1) LRR and Nucleotide Binding Site (NBS) containing cytoplasmic receptor-like proteins. As example of this group there are *Arabidopsis thaliana* RPS2, RPP5, RPM1, and tomato Prf, tobacco N genes. 2) Raf and IRAK genes in mammals, a "serine-threonine kinase" coding Pto gene in tomatoes that is homologous to *Drosophila* Pelle kinase gene, are examples to this group. 3) Cf-2 and Cf-9 genes in tomatoes and HS1^{pro-1} gene in sugar beet, that contain large extracytoplasmic R domains with "putative transmembrane" receptors. 4) The first and the only monocot resistance gene from rice, Xa21 gene, with an extracytoplasmic LRR domain and a cytoplasmic "serine-threonine kinase" region containing transmembrane receptor gene was cloned. 5) HM1 gene in maize is within this group, and encodes an "NADPH-dependent reductase" (Figure 1.6 and Baker *et al.*, 1997).

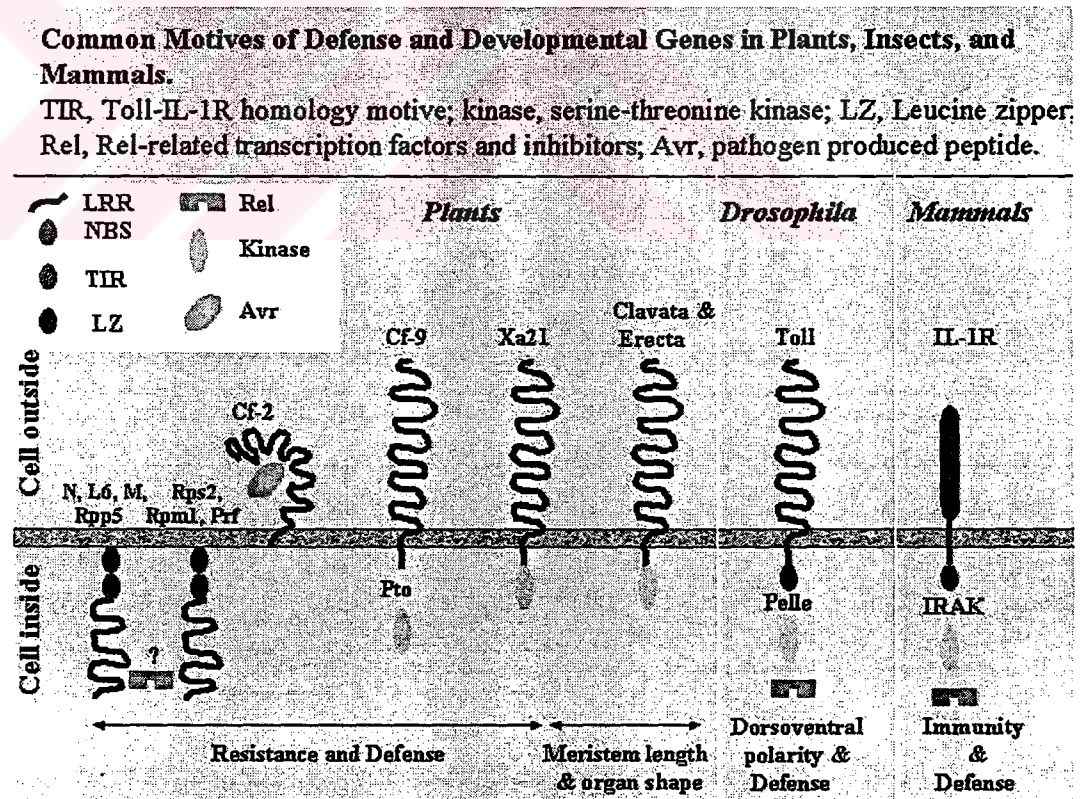


Figure 1.6 The motives shared by plants, insects and mammals related to defense and developmental genes

CHAPTER II

MATERIALS AND METHODS

2.1 MICROSATELLITE ISOLATIONS

2.1.1 Triple Affinity Capturing (TAC) Enrichment Method

The triple helix forming ability of DNA under mild acidic conditions was utilized. Target DNA (*Hind*III digested, ~500 bp region, enriched wheat genomic DNA) was captured with a biotin-labeled (GA)₉ oligonucleotide and streptavidin-coated magnetic beads were used in eliminating the remaining DNA. Enriched, double-stranded DNA was recovered with a mild alkaline buffer by destabilizing the triplex DNA. The genomic DNA library was constructed with the eluted repeat containing DNA fragments (Figure 2.1).

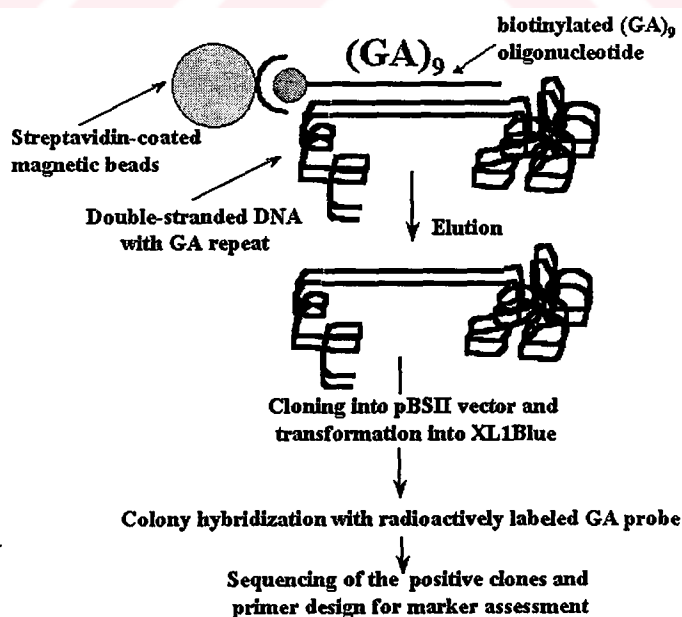


Figure 2.1 Triple affinity capturing enrichment procedure

2.1.1.1 Plant Materials

'Kundurur 1149', a landrace selected Turkish durum cultivar of Central Anatolia, was used for the enrichment of the genomic DNA in terms of repeat units.

2.1.1.2 Isolation and Digestion of Wheat Genomic DNA

Genomic DNA isolations of a durum wheat cultivar 'Kundurur 1149' was performed according to Saghai-MarooF and his co-workers (Saghai-MarooF *et al.*, 1984) with some modifications (Incirli and Akkaya, 2000).

Freeze-dried 200 mg of individual plant tissue was powdered in liquid nitrogen by flame-sealed 1 mL pipet tips quickly so that the crushed tissue was not thawed and dispersed in 4 mL extraction buffer (50 mM Tris-Cl, pH8.0, 10 mM EDTA, pH 8.0, 0.7 M NaCl, 1% CTAB v/w [Hexadecyltrimethylammoniumbromide] and 0.1% β -mercaptoethanol just before usage). The mixture was incubated in 60 °C for 1 hour by inverting the tubes gently several times. 4 mL of chloroform:octanol, 24:1 (v/v), was added and the solution was mixed by inversion to form an emulsion which was centrifuged at 4000 g for 25 minutes at RT to perform phase separation (Denley BS400 Bench Centrifuge, England). The aqueous phase was transferred into a fresh microfuge tube containing 2/3 volume of isopropanol. After two-to-four quick, gentle inversions the mixture was centrifuged at 15,000 rpm for 15 minutes at 4 °C. Aqueous phase was discarded, and the precipitated DNA pellet was washed with 70% ethanol. By centrifugation at 15,000 rpm for 15 minutes at 4 °C the ethanol solution was poured off and the samples were air-dried. The pellet was suspended in 600 μ L of 1xTE buffer. The DNA samples and the remaining plant tissue were stored at -85 °C.

Absorbance values of 100 times diluted DNA samples were measured at 230 nm, 260 nm and 280 nm with a double beam spectrophotometer (Shimadzu UV-1601). DNA absorbs UV light at 260 nm, whereas proteins absorb UV light at 280 nm and carbohydrates at 230 nm. The optimum ratio of A_{260}/A_{280} and A_{260}/A_{230} is indicated as 1.5-1.8 and 1.5-2.0, respectively.

Concentrations were determined according to the equation below. For 50 μ g/mL double stranded DNA (ds DNA), A_{260} is 1.00.

$$A_{260} \times \text{dilution factor} \times (50 \mu\text{g/mL}) \text{ ds DNA} / A_{260} = \dots \mu\text{g/mL DNA}$$

10 µg of 'Kunduru 1149' DNA was digested with 40 units of *Hind*III (New England Biolabs), in the presence of 3µL of NE 4 buffer (New England Biolabs) in 30µL of reaction volume at 37°C for 3 hours. Digested 'Kunduru 1149' DNA was separated on 1% agarose gel with a co-migrating molecular weight standard marker. The region corresponding to 500bp was cut and isolated from the gel using a DNA isolation kit (MBI Fermentas, Lithuania). *Hind*III digested and gel isolated 'Kunduru 1149' DNA was extracted with Phenol/chloroform/isoamyl alcohol, and after 3 M NaOAc (pH5.2) addition it was precipitated with ethyl alcohol (Sambrook *et al.*, 1989).

2.1.1.3 (CT)₉ Enrichment of 'Kunduru 1149' DNA

Streptavidin (Sigma, S-2415, Lot number: 76H0229) was homogenized by gentle mixing. An aliquot of 20 µL streptavidin-coated magnetic bead slurry was transferred into a 1.5 mL tube and 150 µL of 6 X SSC. Another 20 µL of 6 X SSC was added and the mixture was left on the rack at RT. In a separate tube, 600 pmol of biotinylated (CT)₉ was mixed with 30 µg of *Hind*III digested and size selected (~500bp region) 'Kunduru 1149' DNA, in 6 X SSC, within a final volume of 50 µL. The mixture of DNA and oligonucleotide was incubated at 94 °C, for 5 minutes and cooled at room temperature for 30 minutes allowing triple helix formation.

Triple helix complex of 50 µL was transferred on streptavidin-coated magnetic beads, washed as above. The biotin and streptavidin association was allowed to take place by incubation at room temperature for 15 minutes. The supernatant of the triple helix complex was discarded. The washing steps below were then applied: First, 500 µL of '2 X SSC + 0.1% SDS' was added, the mixture was incubated for 5 minutes at 25 °C and supernatant was discarded. This was repeated once more. Secondly, 500 µL of '1 X SSC' was added and incubated for 5 minutes. Again the step was repeated. In the third step, 500 µL of '1 X SSC' was added and the mixture was incubated for 1 hour at 50°C. The step was repeated with an incubation period of 15 minutes at room temperature. Enriched DNA was eluted (as double stranded) in 10 µL water after 15 minutes incubation at room temperature and 2 minutes on the magnet. Another 10 µL water was added on the beads and the eluted solutions were combined (20 µL final volume). The sample was kept at room temperature until usage.

2.1.1.4 Digestion of pKSII KS^{+/−} DNA

10 µg of pKSII KS^{+/−} DNA was digested with 40 units of *Hind*III (New England Biolabs), in 1 X NE 4 buffer (New England Biolabs) in 30µL of reaction volume at 37°C for 3 hours. Digested pKSII KS^{+/−} DNA was isolated from 1% agarose gel using a DNA isolation kit (MBI Fermentas, Lithuania).

2.1.1.5 Construction of Genomic DNA Library and Maintaining of the Clones with Insert DNA

(CT)_n enriched 'Kunduru 1149' DNA was cloned in *Hind*III digested pBSIIKS^{+/−} vector and recombinant vector was transformed in competent *E. coli* Dh5α cells (prepared freshly) according to standard protocols in "Current Protocols in Molecular Biology, Asubel, 1994, Current Protocols, John Wiley and Sons, Inc." and let grown on LB (Amp⁺ / X-Gal / IPTG) agar medium. The white colonies were picked up under sterile conditions and transferred to the microtiter plates and grown in 50 µL LB medium and stored in glycerol at −86 °C until further analysis.

2.1.2 Double Affinity Capturing (DAC) Enrichment Method

The group of Dr. Karagyozov at the Bulgarian Academy of Sciences, Bulgaria, have constructed small-insert wheat genomic libraries enriched in terms of (GT)_n and (GA)_n repeats. In their method, DNA fragments were generated after sonication and were ligated to adaptors followed by PCR-amplification. The PCR products were enriched by hybridisation to (CA)_n and (CT)_n oligonucleotides bound to a nylon membrane. The enrichment products were digested with *Eco*RI and electrophoresed on a 2% agarose gel. The 250 to 450 bp fragments were purified and cloned into dephosphorylated pBluescript SK+ (Stratagene). The clones were transferred to nitrocellulose and hybridised to ³²P-labeled (CA)_n and (CT)_n probes (Karagyozov *et al.*, 1993).

2.1.2.1 Sequencing of Enriched Microsatellite Candidates

The sequencing reactions of the 200 ng purified pBluescript SK+ (Stratagene) recombinant clones were performed using Big-dye fluorescent labelling sequencing reaction

mixture in 20 µL reaction volume with 25 ng T3 primer. The cycling conditions were as following: 25 cycles of 96 °C for 30 seconds, 56 °C for 15 seconds, and 60 °C for 4 minutes. The extension products were purified using Centri-Spin Columns (CS-900, Princeton Separations, NJ). The dried DNA was suspended in 15 µL of Perkin-Elmer Template Suppression Reagent (Perkin-Elmer, CA). The sequences were read on ABI Prism-310 Genetic Analyzer.

2.1.3 Isolation of AFLP-based Microsatellites: A Novel Approach

AFLP marker generation (Zabeau and Vos, 1993; Vos *et al.*, 1995; Figure 1.1) and an enrichment step *via* double heteroduplex complex formation (Figure 2.2) were utilized. AFLP markers were obtained with various Eco+3 and Mse+3 selective amplification primer sets as in performing AFPL (non-radioactively). The products then were enriched for each selective primer combination using a biotinylated target repeat oligonucleotide on the streptavidin-coated magnetic beads. The enriched fragments were eluted separately as each AFLP selective primer combination product, and then were re-amplified in the presence of a radioactively labeled nucleotide, [α -³²P]dATP, using the same set of AFLP selective primer combination (Figure 2.2).

2.1.3.1 Plant Materials

The durum wheat cultivar, 'Selçuklu-97', originally produced in Bahri Dağdaş Milletlerarası Kışlık Hububat Araştırma Merkezi (BD.MİKHAM), Konya, Turkey, was obtained from Assist. Prof. Dr. Ahmet Yıldırım, Gaziosmanpaşa University, Tokat.

2.1.3.2 DNA Isolation

Genomic DNA isolation and concentration determination of the cultivar 'Selçuklu-97' was performed as in Section 2.1.1.2.

2.1.3.3 AFLP Sample Preparation

The reaction steps, the adapters, and the primers of AFLP marker generation were essentially the same as Vos and Zabeau (Zabeau and Vos, 1993; Vos *et al.*, 1995) with optimization of reaction conditions (Incirli and Akkaya, 2000), by using common *EcoRI* and *MseI* adapters.

2.1.3.4 Restriction Enzyme Digestion of Genomic DNA

DNA (300 ng) was digested with *EcoRI* and *MseI* enzymes (3 units from each, MBI Fermentas, Inc. Lithuania) in a final reaction volume of 40 μ L containing, 1 X Universal Buffer (0.1 M KOAc, 25 mM Tris-Acetate, pH 7.6, 10 mM MgOAc, 0.5 mM β -Mercaptoethanol, 10 μ g/mL BSA (Stratagene, CA)) and 0.1 μ g Bovine Serum Albumin (BSA) (MBI Fermentas, Inc. Lithuania). The reactions were carried out for 3 hours incubation at 37 °C.

2.1.3.5 Double Stranded Adapter Preparations

Double stranded (DS) adapters, *EcoRI* DS adapter and *MseI* DS adapter, with respective *EcoRI* and *MseI* complementary ends were ligated to the genomic DNA fragments harboring *EcoRI* and *MseI* restriction sites. Common sequences of double stranded adapters were the same as used in Vos *et al.*, 1995. The single strand of each adapter were separately annealed together by denaturation at 95 °C for 5 minutes, followed by slow cooling at room temperature.

2.1.3.6 Ligation of Adapters

The ligation reaction mixture of 25 μ L volume contained 20 μ L of restriction digest aliquot, 7.5 pmol *EcoRI* DS adapter and 75 pmol *MseI* DS adapter, 1000 units of T4 DNA Ligase (New England Biolabs, MA, USA), 0.2 x T4 DNA ligase buffer (New England Biolabs, MA, USA), 0.5 μ g BSA, and 0.2 mM of ATP. The ligation reaction was carried out for 5 hours at 37 °C.

2.1.3.7 Pre-selective Amplification

Pre-selective amplification was performed using Eco+A and Mse+A primers (20 cycles of three steps as 94 °C for 30 seconds, 60 °C for 30 seconds, 72 °C for 1 minute). Both of the primers used have the extra base A at their 3' ends of the restriction sites. The mixture for pre-selective amplification reaction contained 5 µL ligation reaction products as template in PCR amplification, 75 ng of each of the pre-amplification primers, 0.2 mM dNTPs (MBI Fermentas, Inc. Lithuania), 1 X PCR buffer (10 mM Tris-Cl, pH 8.3, 50 mM KCl), 2.5 mM MgCl₂, 1.5 units of *Taq* DNA polymerase (MBI Fermentas, Inc. Lithuania). The reaction volume was brought to 50 µL by PCR quality water. PCR amplifications were performed in Techne Genius Thermal Cycler, UK.

2.1.3.8 Selective Amplification

Pre-selectively amplified fragments were selectively amplified without radioactive labeling and using nine primer combinations with 3'end extensions reported in Figure 2.2, (11 cycles of denaturation at 94 °C for 30 seconds, annealing at 65 °C (-0.7 °C/each cycle) for 30 seconds, extension at 72 °C for 1 minute and additional 24 cycles of denaturation, annealing and extension at 94 °C for 30 seconds, 56 °C for 30 seconds, and 72 °C for 1 minute, respectively).

2.1.3.9 Enrichment of Selectively Amplified PCR Products

Each selectively amplified PCR product was enriched according to Fisher and Bachmann, 1998, with the following changes. Streptavidin coated magnetic beads (Streptavidin S-2415; Sigma Corp., MO) of 20 µL suspension was washed with 150 µL of 6 X SSC in 1.5 mL microcentrifuge tube using magnetic separator (Sigma Corp., MO), and equilibrated in 20 µL 6 X SSC. In a separate tube, 30 pmol of 5'-end biotinylated (GA)_n oligonucleotide was incubated with 20 µL selective amplification PCR product in a final volume of 50 µL, containing 6 X SSC for 5 minutes at 94 °C. The mixture was cooled to room temperature slowly for 30 minutes. Hybridization complex was loaded on beads in 20 µL of 6 X SSC, allowing streptavidin and biotin association for 15 minutes at room temperature. The supernatant was discarded and the immobilized hetero duplex DNA complex was washed to remove the non-specifically associated fragments, twice in 500 µL 2 X SSC, 0.1% SDS, after incubation for 5 minutes at 25 °C. Washing was repeated two more times in 500 µL of 1 X

SSC at 25 °C as above, and once again in 500 µL of 1 X SSC at 47 °C for 2 minutes. The single stranded "GA" enriched AFLP fragments were eluted twice in 10 µL PCR quality ddH₂O after 15 minutes incubation at 25 °C, in two steps (Figure 2.2).

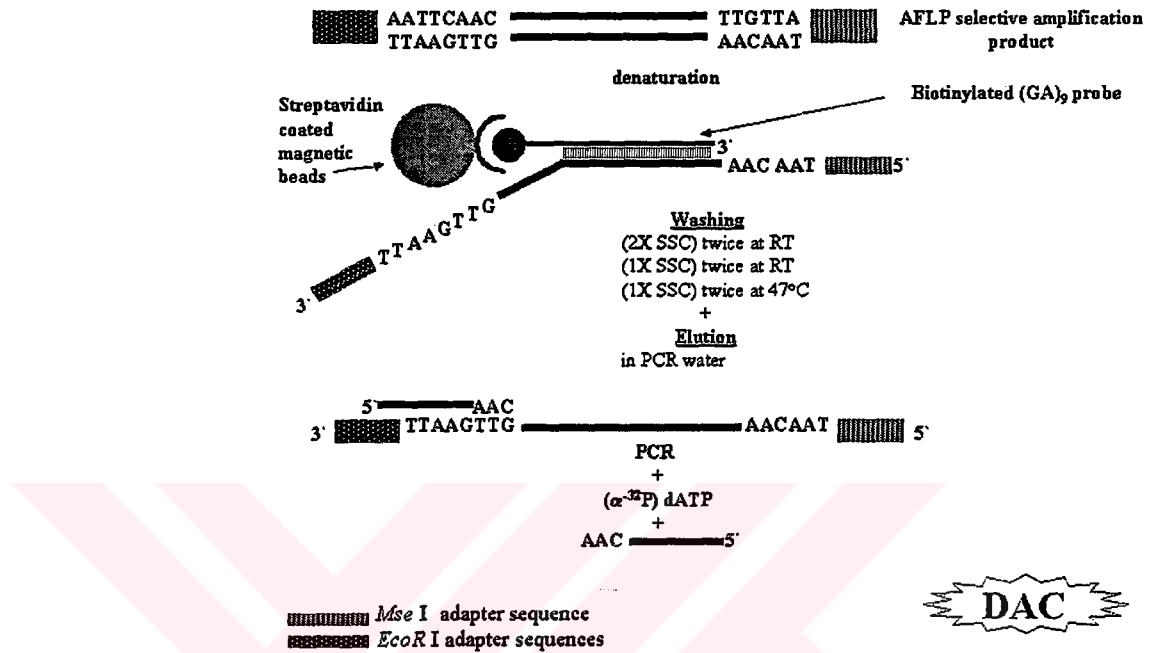


Figure 2.2 Double affinity capturing enrichment procedure

2.1.3.10 Radiolabeled AFLP Reaction Conditions

The radiolabeled selective PCR amplification of the enriched fragments was performed as above, but in the presence of 0.1 µL [α-³²P] dATP (3000 mCi/mmol) (Institute of Isotopes Co. Ltd., Hungary) in a final volume of 20 µL, containing 10 µL eluent as the template DNA (enriched in terms of "GA" repeat), 50 ng of each Eco+ANN and Mse+ANN, 3mM Mg²⁺ (MBI Fermentas, Inc. Lithuania), PCR buffer (75 mM Tris-HCl, pH 9.0, 20 mM (NH₄)₂SO₄, 0.01 % Tween-20) (MBI Fermentas, Inc. Lithuania) and 1 unit *Taq* DNA polymerase (MBI Fermentas, Inc. Lithuania) (Figure 2.2).

2.1.3.11 [γ -³³P]-ATP Labeling of ϕ x174/*Hinf*I Molecular Size Marker

5'-end labelling reaction of ϕ x174/*Hinf*I marker (0.5 μ g) was carried out in 20 μ L final volume (adjusted with PCR grade water) and contained the following reagents: 0.1 μ L of 3000 mCi/mmol [γ -³³P]-ATP (Institute of Isotopes Co. Ltd., Hungary), 0.2 units T4 DNA kinase (MBI Fermentas, Inc. Lithuania), 1 X T4 DNA kinase buffer (MBI Fermentas, Inc. Lithuania). The reaction was performed for 1 hour at 37 °C that was followed by the enzyme inactivation at 70 °C for 10 minutes. At the end of the reaction, 8 μ L of stop solution (95% formamide, 20mM EDTA, 0.25% bromophenol blue and 0.025% xylene cyanol) was added.

The enriched AFLP-fragments were co-electrophoresed with 2-3 μ L of 1:10 diluted (in 1 X TE-buffer) molecular weight standard marker (ϕ x174/*Hinf*I: 726, 713, 553, 500, 427, 417, 413, 311, 249, 200, 151, 140, 118, 100, 82, 66, 48 bp).

2.1.3.12 Separation of Selective Amplification Bands

The selective amplification reactions were stopped by the addition of 8 μ L of stop solution (95 % formamide, 20 mM EDTA, 0.25 % bromophenol blue and 0.025 % xylene cyanol) and denatured for 10 minutes at 94 °C prior to loading onto the 30 minute pre-electrophoresed 4.5 % of denaturing polyacrylamide gel (4.2 % acrylamide, 0.3 % N, N'-methylene-bis-acrylamide, 8 M urea, 1 X TBE (90 mM Tris base, 90 mM boric acid, 2 mM EDTA)). To 60 mL of gel solution, 650 μ L of 10 % APS (ammonium persulphate) and 25 μ L of TEMED (N, N, N', N'-Tetramethyl ethylene diamine) was added and gels were cast in GIBCO-BRL/Life Technologies sequencing gel plates. The electrophoresis was terminated at the point where the bromophenol blue dye reaches to the bottom of the gel, that takes about 2 hours of electrophoresis at 50-60 watts.

At the end of the electrophoresis the gel was transferred to a filter paper (Whatman-3MM) and dried for about 45 minutes at 78 °C on the gel dryer (Slab Gel Dryer SGD2000). Prior of exposing X-ray film to the gel, the filter paper was marked with radioactive nucleotide at its corners for correct orientation. X-ray (Bio-Max/MR) film was exposed to the dried gel for several days and then developed at the M.E.T.U. Health Center.

2.1.3.13 Extraction of DNA from the AFLP Gel

Developed film was aligned precisely; the corresponding bands were cut off together with the attached filter paper behind or scratched from the paper using a sterile blade. The cut-off gel pieces were incubated in 50 μ L PCR grade water at 4 °C overnight, following couple of freeze-thaw periods by holding in liquid nitrogen and at room temperature (Figure 2.3).

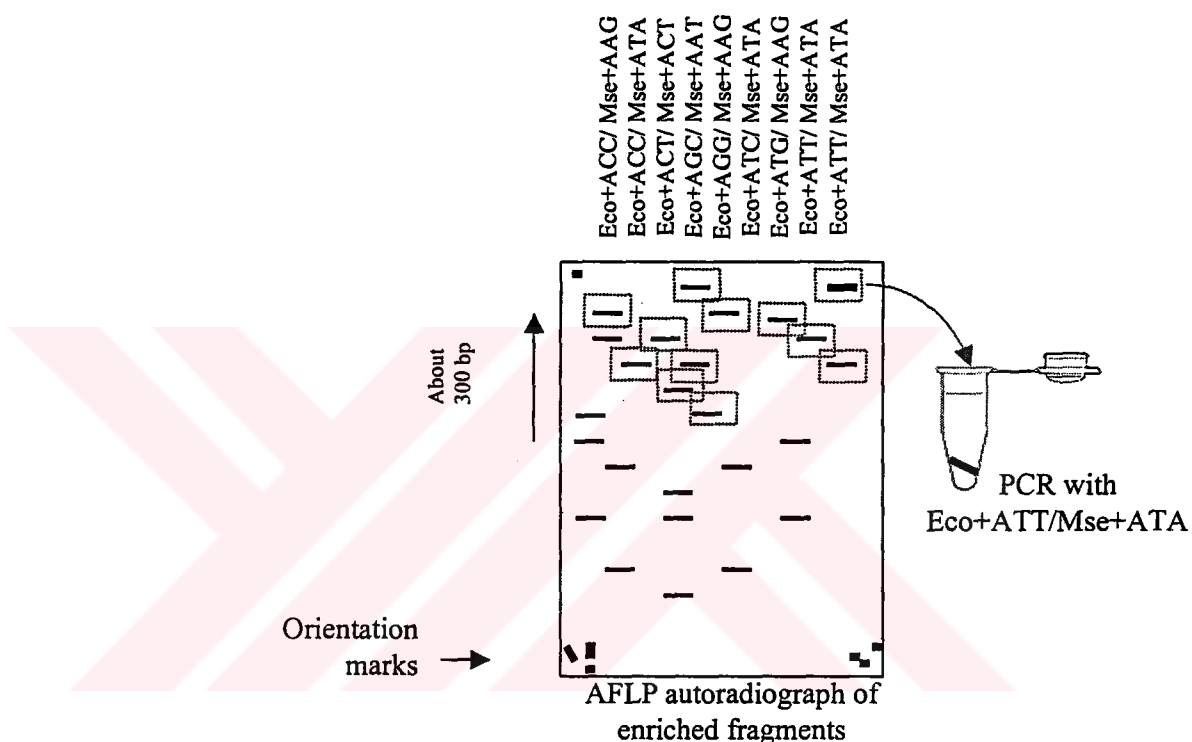
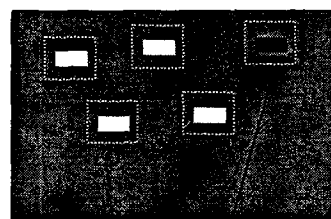


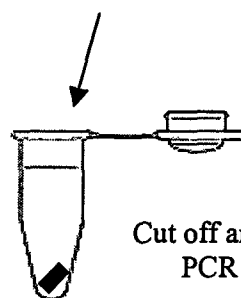
Figure 2.3 Band extraction from the dried polyacrylamide gel.

2.1.3.14 Re-amplification of Band-isolated AFLP Fragments

DNA from water suspended gel pieces (5 μ L) was used for re-amplification (30 cycles of denaturation, annealing and extension at 94 °C for 30 seconds, 56 °C for 30 seconds, and 72 °C for 1 minute, respectively) in a reaction containing 25 ng Eco+ANN and 75 ng of Mse+ANN selective primers, 3 mM Mg²⁺, PCR buffer (75 mM Tris-HCl, pH 9.0, 20 mM (NH₄)₂SO₄, 0.01 % Tween-20) and 1 unit *Taq* DNA polymerase in 50 μ L reaction volume. Individually amplified fragments were analyzed on 1.5 % agarose gels (Figure 2.4).



Separation on 1.5 %
agarose gel



Cut off and isolate the
PCR products

Figure 2.4 Band isolation from the agarose gel

2.1.3.15 Selection of the Fragments According to their Restriction Sites

The ends of the DNA fragments were checked for the presence of both of the *Mse*I and *Eco*RI site adapters, by re-amplifying with primers, Mse+ANN alone, Eco+ANN alone, and Mse+ANN/Eco+ANN together. Those only amplified with both site primers, but not with either of the single primer alone (Eco+ANN alone or Mse+ANN alone) were used for direct sequencing (Figure 2.5). The DNA fragments from the cut off agarose gel pieces were extracted by using QIAquick gel extraction kit (Qiagen, CA) in 30 μ L PCR grade water. The amount of DNA was assessed in comparison to 2 μ L of 1 kb DNA ladder (10 μ g/ μ L) on 0.8 % agarose gel.

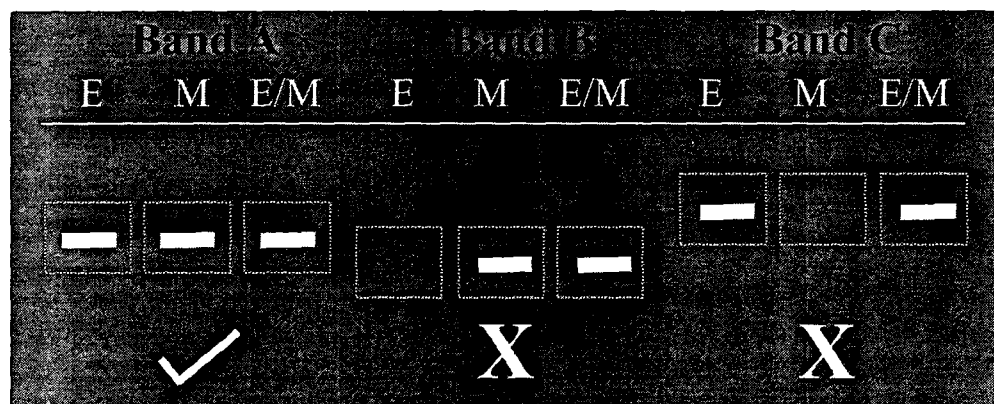


Figure 2.5 Determination of restriction sites of the amplified fragments in 1.5 % agarose gel. Band A is to be sequenced with either Eco+NNN or Mse+NNN primer. E: Eco+NNN, M: Mse+NNN, E/M: Eco+NNN/Mse+NNN

2.1.3.16 Sequencing of Fragments

The sequencing reactions of the 20-25 ng DNA of the isolated and purified fragments were performed using Big-dye fluorescent labelling sequencing reaction mixture in 20 μ L reaction volume with 25 ng Eco+ANN or 75 ng of Mse+ANN primer. The cycling conditions were as following: 25 cycles of 96 $^{\circ}$ C for 30 seconds, 56 $^{\circ}$ C for 15 seconds, and 60 $^{\circ}$ C for 4 minutes. The extension products were purified using Centri-Spin Columns (CS-900, Princeton Separations, NJ). The dried DNA was suspended in 15 μ L of Perkin-Elmer Template Suppression Reagent (Perkin-Elmer, CA). The sequences were read on ABI Prism-310 Genetic Analyzer (Figure 2.6).

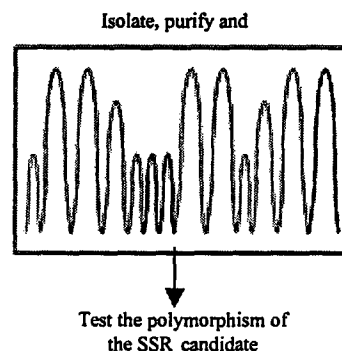


Figure 2.6 Direct sequencing of the fragments with different 3' and 5' ends

2.2 GENOTYPING OF ANATOLIAN DOUBLED-HAPLOID DURUM LINES WITH SSR MARKERS

2.2.1 Plant Materials

A total of 18 individual DH lines; 3 DH seeds from 'Çakmak-79' (designated as Ç1-3); 3 DH seeds from 'Berkmen-469' (B1, B3, B5) and 12 DH seeds from 'Kunduru-1149' (K1-K12) were investigated (Savaskan, 1997). The progenitors of DH lines, 'Berkmen-469', 'Çakmak-79' and 'Kunduru-1149' were also utilized in our analysis. From the local Anatolian varieties, 'Akbaşak-073/44', 'Ankara-04/95', 'Ankara-08/95', 'Kızıltan-91', 'Çeşit-1252', 'Ankara-05/95', 'Ankara-06', 'Ankara-02/94', 'Ankara-03/94', 'Gökgöl-79', 'Tunca-79' and 'Ankara-07/95' were also analyzed with seven of the polymorphic microsatellites used in the analysis of DH lines (Dograr, 2000 and 1999).

2.2.2 DNA Isolation and PCRs

DNA was isolated from half cut seeds according to a literature procedure (Plaschke *et al.*, 1995) with some modifications (Dograr *et al.*, 2000). DNA samples of genotypes were amplified with 15 different microsatellite primer sets; WMS2, WMS5, WMS6, WMS10, WMS11, WMS18, WMS30, WMS43, WMS46, WMS67, WMS88, WMS95, WMS120, WMS131 and WMS135, the sequences of which and the chromosome locations were reported earlier (Röder *et al.*, 1995 and 1998).

2.2.3 PCR Analysis

The PCR amplification conditions, electrophoresis and detection of microsatellites were performed as previously reported (Dograr *et al.*, 2000; Röder *et al.*, 1995 and 1998).

2.2.4 Cluster Analysis

Numerical Taxonomy and Multivariate Analysis System software (NTSYS, version 1.70) (Rohlf, 1992) was used to generate UPGMA dendrogram by using the simple matching coefficient similarity matrix of 18 DH individuals at 12 microsatellite loci. The genetic distances between the populations of doubled-haploid durum wheat lines, 'Çakmak-79', 'Berkmen-469' and 'Kunduru-1149', were standardized and processed using the software RST-22 (Goodman, 1997) which calculates $(\delta\mu)^2$ as defined by Goldstein (Goldstein *et al.*, 1995). The genetic distances were also calculated with 'Nei72' algorithm (Rohlf, 1992) using allele frequencies. Polymorphism information content (PIC) values (Table 3.3) were obtained according to the formula below (Russell *et al.*, 1997; Dograr, 1999).

$$PIC_i = 1 - \sum_{j=1}^n P_{ij}^2$$

2.3 ISOLATION OF WHEAT RESISTANCE GENE ANALOGS

2.3.1 Plant Materials

The cultivar 'Pastor' has one major and two minor resistance genes and 'Avocet' has two minor resistance genes to the fungal pathogen *Puccinia striiformis*. The seeds of 'Pastor' and 'Avocet' as well as 26 individual seeds from their F3 cross were supplied from the CIMMYT-Mexico. The complete-susceptible variety 'Chinese spring' was kindly provided by Dr. Ahmet Yıldırım from the Gaziosmanpaşa University, Faculty of Agriculture, Department of Crop Sciences, Tokat. The wild wheat varieties, *T. monococcum* (Acc. No. TR36938), *T. dicoccoides* (Acc. No. TR02771), *A. speltoides* are plant species (obtained from ICARDA, Aleppo, Syria) and used in other projects in the Middle East Technical University, Chemistry Department, Biochemistry Division, Akkaya lab.

2.3.2 DNA Isolations

Freeze-dried 200 mg of each individual plant tissue, grown for 15 days, was powdered in liquid nitrogen by home-made hand-held homogenizer before adding extraction buffer (2% CTAB w/v, 100 mM Tris, pH 8.0, 20 mM EDTA, pH 8.0, 1.4M NaCl, and 0.1% β -mercaptoethanol just before usage). The crushed tissue was never let to thaw. The samples were dispersed in 0.7 mL extraction buffer and incubated at 65 °C for 45 minutes by occasional gentle swirling. 700 μ L chloroform:isoamylalcohol, 24:1 (v/v) was added and the solution was centrifuged at 7,000 rpm for 5 minutes at RT. Then the aqueous phase was transferred into a fresh tube containing 700 μ L of chloroform:isoamylalcohol, 24:1 (v/v). The mixture was centrifuged at 7,000 rpm for 5 minutes. The supernatant was transferred to a fresh tube again and mixed with 0.6 volume of isopropanol. The pellet was washed with 70 % ethanol by centrifuging at 15,000 rpm for 5 minutes. Ethanol was poured off and the samples were air dried. The pellet was dissolved in 200 μ L 1 X TE buffer. The DNA samples and the remaining plant tissue samples were stored at -85 °C (Incirli and Akkaya, 2000).

Alternatively, isolations were carried out with the genomic DNA isolation kit *Nucleon phytopure* (Scotlab) according to the manufacturers' recommendations.

Concentrations of DNA samples are determined as in section 2.2.4.

2.3.3 Design of the PCR Primers for RGA Amplifications

By examining the conserved regions of dicotyledonous plant resistance genes, targeting nucleotide-binding site and leucine rich repeat regions, so far isolated and sequenced the following primers have been designed (Pan *et al.*, 2000; Leister *et al.*, 1998 and 1996; Yu *et al.*, 1996). We have used the same primers used by Leister *et al.*, 1998. The primer sequences are presented in Table 2.1.

Table 2.1 Degenerate primers, targeting NBS-LRR regions, used in resistance gene analog amplifications (Leister *et al.*, 1998).

• Forward Direction Primer “P-loop motif” (G G V/I G K T T): S1: 5'-GGI GGI (A,G)TI GGI AAI ACI AC-3' (T _A = 48 °C)	
• Reverse Direction Primers (G L P L A/T L): AS1: 5'-IAG IG(C,T) IAG IGG IAG ICC-3' (T _A = 48 °C) AS2: 5'-IAG IG(C,T) IAG IGG IAA ICC-3' (T _A = 46 °C) AS3: 5'-IAG IG(C,T) IAA IGG IAG ICC-3' (T _A = 46 °C) AS4: 5'-IAG IG(C,T) IAA IGG IAA ICC-3' (T _A = 46 °C) AS5: 5'-IAA IG(C,T) IAG IGG IAG ICC-3' (T _A = 46 °C) AS6: 5'-IAA IG(C,T) IAG IGG IAA ICC-3' (T _A = 44 °C) AS7: 5'-IAA IG(C,T) IAA IGG IAG ICC-3' (T _A = 44 °C) AS8: 5'-IAA IG(C,T) IAA IGG IAA ICC-3' (T _A = 42 °C)	

2.3.4 PCR Conditions in RGA Amplifications

All of the reactions were performed in standard conditions. PCRs contained: 1 X PCR Buffer (10 mM Tris-HCl (pH 8.8), 50 mM KCl, % 0.08 Nonidet P40), 1.5-2.0 mM MgCl₂, 0.8 mM total dNTP, 4 units of *Taq* DNA Polymerase and 50 ng plant genomic DNA. The reactions were carried out in 50 µL total volume. The concentration of both primers in the reactions were kept 25 pmol each. The reactions were performed on PTC-100 Programmable Thermal Controller (MJ Research, Inc.). Preheating of 2 min at 95 °C was applied prior to cycling steps. The annealing temperatures of the reactions were in between 41-47 degrees and

the cycling numbers were 35 times (1 min at 94 °C, 1 min at 72 °C and 1 min at the annealing temperature) (Table 2.2).

Table 2.2 PCR conditions of RGAs generated by degenerate primer sets

Plant DNA	Primer combination	Annealing temp (°C)	Cycles	Thermocycler
'Pastor'	S1 / AS1	47	35	PTC-100, MJ Research, Inc.
	S1 / AS2	45	35	PTC-100, MJ Research, Inc.
	S1 / AS2	45	35	Genius 2000, Techne, Inc.
	S1 / AS3	45	35	PTC-100, MJ Research, Inc.
	S1 / AS4	45	35	PTC-100, MJ Research, Inc.
	S1 / AS5	45	35	PTC-100, MJ Research, Inc.
	S1 / AS6	43	35	PTC-100, MJ Research, Inc.
	S1 / AS8	41	35	PTC-100, MJ Research, Inc.
'Avocet'	S1 / AS1	47	35	PTC-100, MJ Research, Inc.
	S1 / AS2	45	35	PTC-100, MJ Research, Inc.
	S1 / AS3	45	35	PTC-100, MJ Research, Inc.
	S1 / AS4	45	35	PTC-100, MJ Research, Inc.
	S1 / AS5	45	35	PTC-100, MJ Research, Inc.
	S1 / AS6	43	35	PTC-100, MJ Research, Inc.
	S1 / AS8	41	35	PTC-100, MJ Research, Inc.
'Chinese spring'	S1 / AS2	45	35	PTC-100, MJ Research, Inc.
	S1 / AS5	45	35	PTC-100, MJ Research, Inc.
	S1 / AS6	43	35	PTC-100, MJ Research, Inc.
	S1 / AS8	41	35	PTC-100, MJ Research, Inc.
<i>T. monococcum</i> (Acc. No. TR36938)	S1 / AS3	45	35	PTC-100, MJ Research, Inc.
<i>T. dicoccoides</i> (Acc. No. TR02771)	S1 / AS3	45	35	PTC-100, MJ Research, Inc.

2.3.5 Cloning of PCR Products

The PCR products of wild wheat varieties, 'Avocet', 'Pastor' and 'Chinese spring' at all the combinations were band isolated (Qiagen Gel Isolation Kit) from the expected sizes (~500 bp). These band isolated DNA fragments were used as a template for re-amplification. The reamplified PCR products were directly cloned into a pCR 2.1 (Invitrogen) vector according to the manufacturers' recommendations. In another experiment, the 500 bp band of PCR product from Pastor variety, amplified with S1-AS1 primer set, was isolated and purified from 1.5% agarose gel with Qiagen Gel Isolation Kit and cloned into a pGEM-T Easy (Promega) vector according to the manufacturers' recommendations (Table 2.3).

Table 2.3 The total number of maintained RGAs clones and the sequenced ones

Plant DNA	Primer combination	Vector	Total number of colonies	Number of colonies seq
'Pastor'	S1 / AS1	PCR 2.1 (Invitrogen)	18	5
	S1 / AS2	PCR 2.1 (Invitrogen)	20	5
	S1 / AS2	pGem T Easy (Promega)	12	9
	S1 / AS3	PCR 2.1 (Invitrogen)	20	6
	S1 / AS4	PCR 2.1 (Invitrogen)	20	5
	S1 / AS5	PCR 2.1 (Invitrogen)	16	2
	S1 / AS6	PCR 2.1 (Invitrogen)	37	3
	S1 / AS8	PCR 2.1 (Invitrogen)	45	5
'Avocet'	S1 / AS1	PCR 2.1 (Invitrogen)	30	5
	S1 / AS2	PCR 2.1 (Invitrogen)	20	5
	S1 / AS3	PCR 2.1 (Invitrogen)	22	6
	S1 / AS4	PCR 2.1 (Invitrogen)	20	6
	S1 / AS5	PCR 2.1 (Invitrogen)	25	2
	S1 / AS6	PCR 2.1 (Invitrogen)	36	4
	S1 / AS8	PCR 2.1 (Invitrogen)	44	5
'Chinese spring'	S1 / AS2	PCR 2.1 (Invitrogen)	20	5
	S1 / AS5	PCR 2.1 (Invitrogen)	25	2
	S1 / AS6	PCR 2.1 (Invitrogen)	36	2
	S1 / AS8	PCR 2.1 (Invitrogen)	40	1
<i>T. monococcum</i> (Acc. No. TR36938)	S1 / AS3	PCR 2.1 (Invitrogen)	100	3
<i>T. dicoccoides</i> (Acc. No. TR02771)	S1 / AS3	PCR 2.1 (Invitrogen)	80	5
Total number of colonies:			686	91

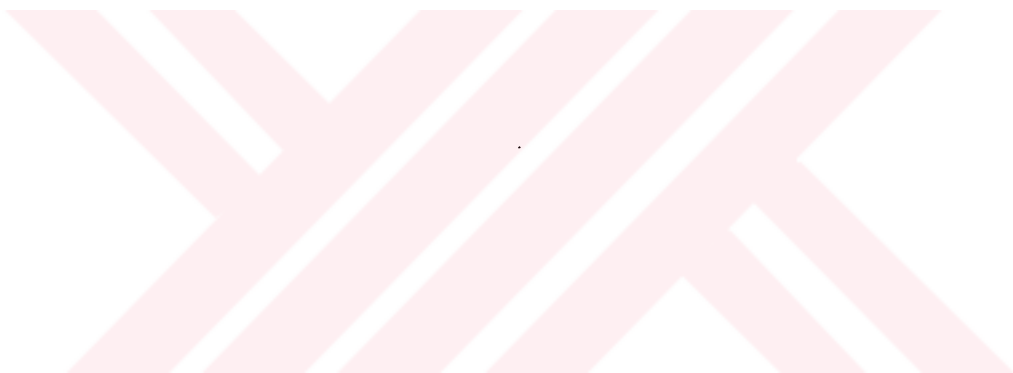
* Randomly selected

2.3.6 Sequencing of Cloned Amplification Products

At least a resistance gene analog from every group was sequenced from both directions using an automated DNA sequencer (ABI PRISM-310 Perkin Elmer). The clones to be sequenced were picked up with a sterile tip and passed into 50 μ L PCR water. The tubes were mixed, by vortex, and 45 μ L of the homogenates were removed. The remaining 5 μ L were used as template in 10 or 20 μ L sequencing reaction (ABI-PRISM Sequence Kit) and the reactions were carried out on PTC-100 Programmable Thermal Controller (MJ Research, Inc.). The products were purified using G-50 sephadex columns and dried in Speed-vac. Samples were suspended in 20 μ L ABI-PRISM Template Suppression Reagent and sequenced with an ABI PRISM-310 sequencer. The total number of sequenced colonies are presented in Table 2.3.

2.3.7 Sequence and Homology Analysis

NCBI databanks were searched using Blast Algorithm (Altschul *et al.*, 1990) and homologies of the sequences with available information from plant and other living systems were compared. Sequences in the GenBank were classified according to the best homologies. Sequence analysis was performed using Jellyfish 1.1 DNA analysis software (http://www.biowire.com/bw_jsp/j_welcome_top.jsp). The sequence alignments were obtained using ClustalX 1.81 (Jeanmougin *et al.*, 1998) and Boxshade (http://www.ch.embnet.org/software/BOX_form.html) computer programs. Phylogenetic relationships were obtained using neighbor-joining method (Saitou and Nei, 1987).



CHAPTER III

RESULTS

3.1 ISOLATION OF NEW MICROSATELLITE MARKERS

3.1.1 Triple Affinity Capturing (TAC) Enrichment Method

From the (CT)_n enriched genomic DNA library construction 3200 positive clones were selected and stored at -86 °C. These clones should have been screened with a probe and the positive clones should have been further assessed for the repetitive sequence. However, this is a tedious procedure and was not followed since our novel SSR isolation procedure worked with a high efficiency (see below).

3.1.2 Double Affinity Capturing (DAC) Enrichment Method

The group of Dr. Karagoyozov, Bulgarian Academy of Sciences, have constructed small-insert wheat genomic libraries enriched in terms of (GT)_n and (GA)_n repeats. In the method used by them, DNA fragments were generated after sonication and were ligated to adaptors followed by PCR-amplification. The PCR products were enriched by hybridisation to (CA)_n and (CT)_n oligonucleotides bound to a nylon membrane. The enrichment products were digested with *Eco*RI and electrophoresed on an agarose gel. The 250 to 450 bp fragments were purified and cloned into dephosphorylated pBluescript SK+ (Stratagene). The clones were transferred to nitrocellulose and hybridised to ³²P-labeled (CA)_n and (CT)_n probes (Karagoyozov *et al.*, 1993).

Clone #311

AGCTGGAGCTCCACCGCGGTGGCGGCCGCTAGATCTCCGATTGCTCCGTCCTCGCATGATCGCATGTAACACATGCATCAACACGC
ACATTTTGACACCCTCACCAACCCCTGCCCACCTCAAGGCACGCACAACCTCTTCGTGAATACCCATTCTCTCCTTTGGTTTCTTT
TCTCTCAATATCTG C C C C C C C T C C C C C C CCCTCCCCGAGC C C ATTCATCCCCATCCAAGGTTATA
GGGCGACCTCCTCTCTCTTGTGCTTCTTTGCACCCTAGCATGAACAACACCTCAATCTTCAAAGAGGGCATCGGAGATCTAGAACTAGT
GGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCCGGTACCCAATTGCGCCTATAGT
GAGGTCGTATTACAATTCACTGGCCGTCGTTTACAACGTCGTGACTGGGAAAACCTGGCGTTACCCAACCTAATCGCCTTGCAGCA
CATCCCCCTTTGCGCAG

Clone #312

CGCATGTCAACACATGCATCAACACGCACATTTTGACACCCTCACCAACCCCTGCCACCTCAAGGCACGCACAACCTCTCTTCGTGA
ATACCCATTCTCTCCTTTGGTTTCTTTTCTCTCAATATCTGAG C C C C C C C T C C C C C C CCCTCCCCGAG
CACACAATTCATCCCCATCCAAGGTTATAGGGCGACCTCCTCTCTCTTGTGCTTCTTTGCACCCTAGCATGAACAACACCTCAATCTTC
AAAGAGGGCATCGGAGATCTAGAACTAGTGGATCCCCCGGGCTGCA

Figure 3.1 (Continued).

The polymorphic status of these microsatellites are currently assessed in our lab as a part of another project.

3.1.3 Isolation of AFLP-based Microsatellites: A Novel Approach

Considering most of the laboratories using microsatellites also probably are applying AFLP, we developed a novel SSR isolation procedure that utilizes AFLP during the first steps. Thus, the only weakness of SSR marker system, which necessitates expensive sequence information generation via cloning and screening and a lot of time consumption, will probably be overcome to some extent by utilizing this shortcut procedure we developed in our lab. The procedure is equally applicable to any eukaryotic species. Part of the method is summarized in Figure 3.2. About half of the total number of bands that were labelled and marked with dots (Figure 3.3), each agarose gel isolated band was checked for the presence of both of the sites. That is very important since otherwise sequencing of the fragments is impossible. Out of 30 selected bands, 21 had both *EcoRI* and *MseI* site adaptors (Figure 3.2).

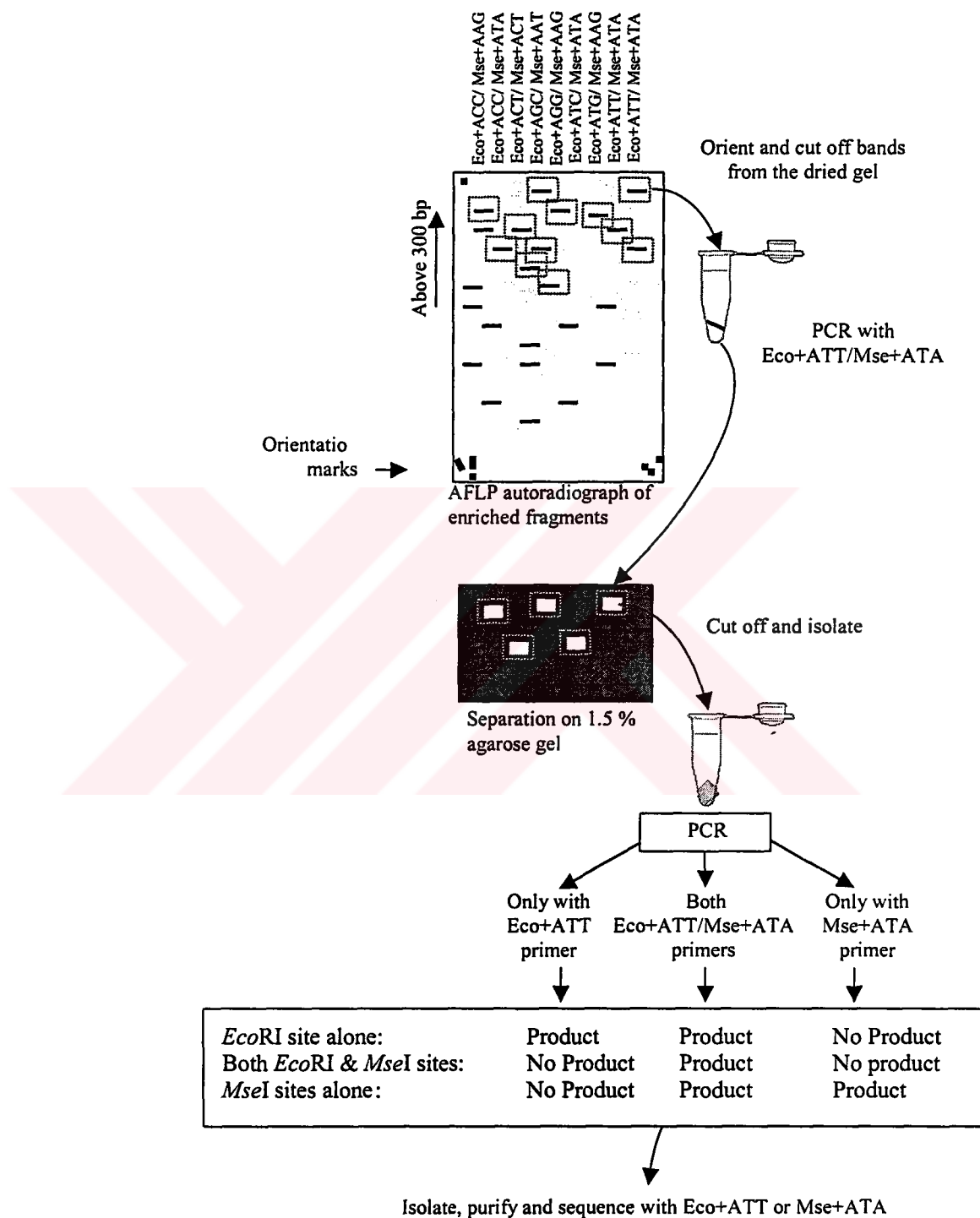


Figure 3.2 A schematic representation of the SSR isolation steps, exemplified with Eco+ATT and Mse+ATA primers combination.

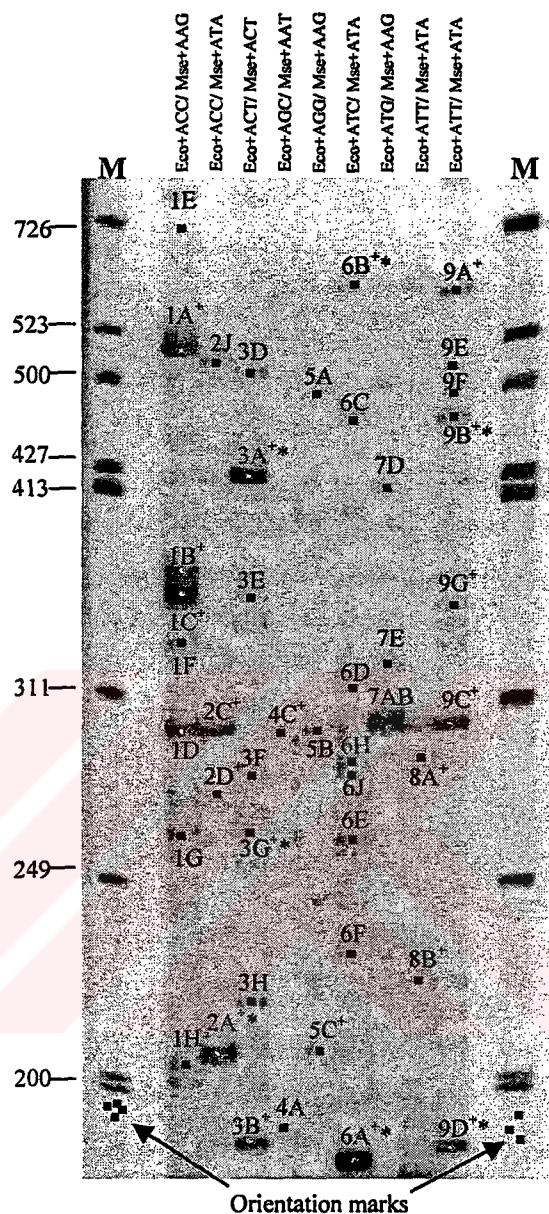


Figure 3.3 Autoradiograph of AFLP fragments: The selectively amplified 'Selçuklu-97' AFLP fragments generated after "GA" repeat enrichment using the set of primers indicated on top of each lane. The bands that are labeled and marked with dots (57 bands) were randomly selected for reamplification. From 30 bands checked, 21 marked with '+' had both *Eco*RI and *Mse*I site adapters. Among 21 purified fragments, 7 marked with '*' were sequenced. The radioactively labeled marker (M) is ϕ X 174/Hinf I.

Many well-separated bands were observed on the AFLP gel after selective amplification PCR of the "GA" enriched AFLP fragments (Figure 3.3), of which 57 bands were re-amplified and purified. All the re-amplified bands were observed as pure fragments on the agarose gels (Figure 3.4). Among these products, those with >300-400 bp in length were selected for designating the fragments with both *EcoRI* and *MseI* site adapters. Each agarose gel isolated band (30 bands) was checked for the presence of both of the sites, by amplifying the agarose gel-purified bands (Figure 3.4) once more in the presence of one type of primer (*Eco*+ANN) or the other (*Mse*+ANN) or in the presence of both of the primers. The fragments that were reamplified in the presence of both type of primers but not amplified with only one type of primer (21 bands) were selected for further analysis (Table 3.1; 70 %). In addition, a few short fragments with lengths about 200 bp, such as 6A and 9D were also isolated and purified in order to evaluate if such short fragments can be sequenced directly. We were able to obtain sequences of those as well.

Table 3.1 Bands cut off and isolated from AFLP gel, labeled as in Figure 3.3.

Bands								
1A+	2A*+	3A*+	4A	5A	6A*+	7AB	8A+	9A+
1B+	2C+	3B+	4B	5B	6B*+	7C	8B+	9B*+
1C+	2D+	3C	4C+	5C+	6C	7D		9C+
1D+	2E	3D		5F	6D	7E		9D*+
1E	2F	3E			6E	7F		9E
1F	3J	3F			6F			9F
1G		3G*+			6G			9G+
1H		3H			6H			
1J		3J			6J			
1K		3K						
1L								

Selected bands for isolation and purification (30/57 = 45 %)

+ Fragments with both *EcoRI* and *MseI* site adapters (21/30 = 70 %)

* and bold characters: Sequenced fragments (7 bands)

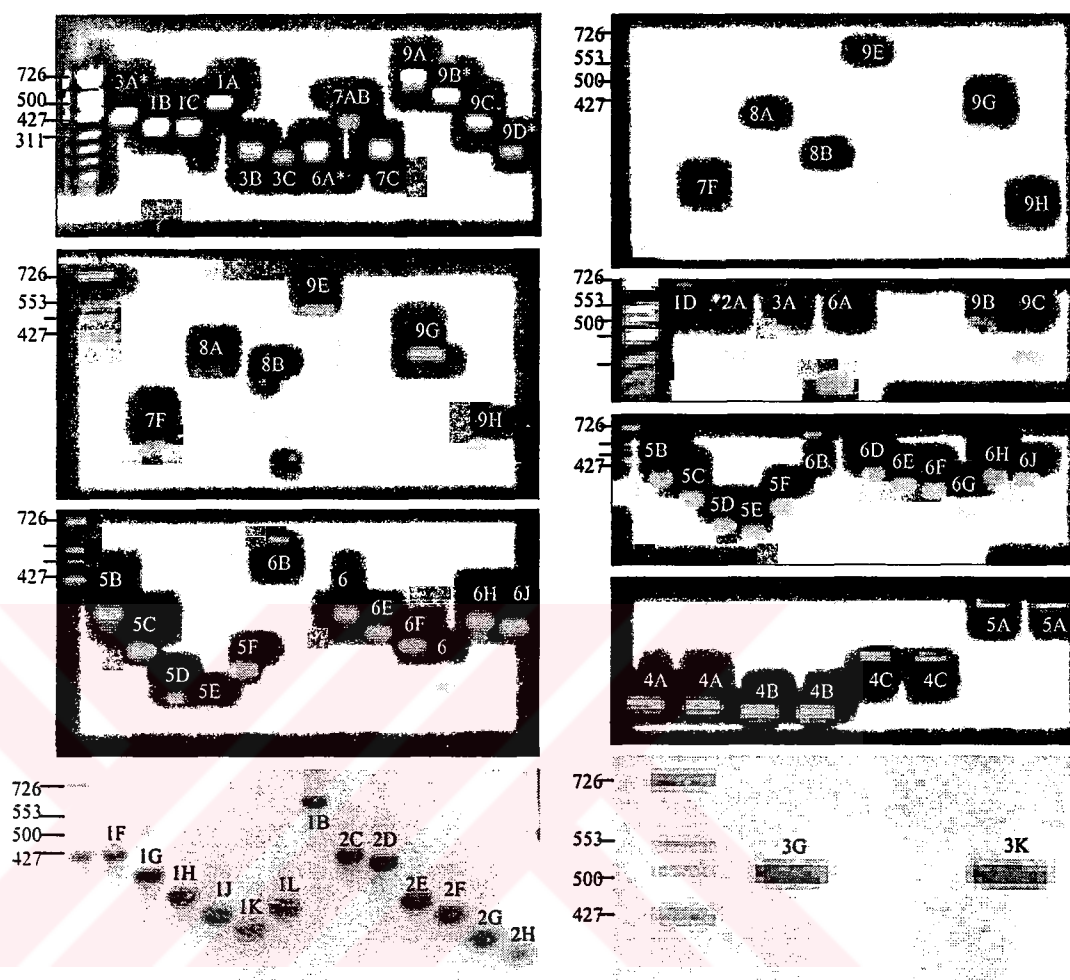


Figure 3.4 PCR re-amplified band-isolated AFLP fragments from DNA denaturing polyacrylamide gel separated on 1.5% agarose gel (The bottom gel pictures are the negative images of the photographs).

Seven of the positive fragments harbouring both Eco+ANN and Mse+ANN sites (Table 3.1) were selected for sequence analysis. Sequence reactions were performed as explained in the section 2.1.3.16 The sequence of the selected 7 bands are as given in Figure 3.5.

The microsatellite primers (Table 3.2) were designed for 3A, 6B and 9B fragments using Primer Detective software, version 1.01 (Clontech Corp., CA) and they were found to contain composite repeat types of (CT)₃(CCCCT)₃(CT)₈, (C)₁₂(CA)₅AA(GA)₆TG(AG)₃, and (TG)₄AT(TC)₁₂(GT)₂, respectively. We did not attempt to design primers for 1A, 2A, 6A and 3G since the repeats in these fragments were not consecutive (sequences other than the core repeat unit are scattered within the repeat regions). However, the region was still enriched in terms of the enriched repeat units (the uninterrupted (GA)_n repeats being ≤4) indicating the low stringency of our enrichment method (Figure 3.5).

[illegible]

The other fragments sequenced were all very rich in "GA" or "CT" repeats, however under the stringency conditions applied for enrichment, a minimum of 3 repeating units (9D) could be selected. In order to find loci with the tandem repeats of 10 or more, the stringency can be increased during enrichment, since the larger the repeat number, the higher the observed polymorphism levels.

This method, using different AFLP selective primer combinations, would be almost endless source for isolating sufficiently long fragments containing any target repeat units. In the methodology, further improvements can be done by taking some precautions such as, longer electrophoresis of the AFLP gel in order to separate the fragments further apart. However, in this trial, there were no contaminated fragments observed on the agarose gels. Also, no sequencing problems were encountered, as long as they possessed different adapters at the ends of each individual AFLP fragments. All the AFLP fragments sequenced were homogenous (not different loci with the same length). Secondly, we believe, washing temperature can be more stringent than the temperature we used, and the number of washing steps we had performed for "GA" enrichment. The same stringency conditions applied in "GA" enrichment would be applicable to the "GT" enrichment, too. However, for the enrichment in terms of "GC" and "AT" di-nucleotide repeats due to higher and lower melting temperatures, respectively, appropriate stringent washing conditions should be applied. There exists, of course, another problem for the "AT" enrichment, it is a self-annealing repeat, and this is also the most likely reason for the lower numbers of markers found containing "AT" repeats by colony hybridization screening. By this method however, it may be possible to prevent the self-hybridization more efficiently since the formation of heteroduplex DNA is in a small-volume and it takes place in a short time, 50 μ L and 15 minutes, versus 5-10 mL of hybridization solution and overnight incubation period are needed during colony hybridization screenings. In order to find multiple repeat types at once, the aliquots of the eluted DNA of each enrichment in terms of different repeat types can be pooled prior to the radiolabeled amplification per each AFLP selective primer combination. Increased stringency of the enrichment step would lower the number of bands that were observed on the AFLP gel, but the ratio of fragments containing repeats will be higher. Thus, combining high stringency with pooling enriched DNA in terms of various repeat types would increase the ratio of the sequences with longer repeats to the bands on the AFLP gel in folds.

We believe this short-cut isolation procedure for finding and sequencing of microsatellite loci is a viable alternative to current practice and may prove to be highly advantageous and productive for most laboratories.

3.2 GENOTYPING OF ANATOLIAN DOUBLED-HAPLOID DURUM LINES WITH SSR MARKERS

The markers selected are from the published sequences of Röder and co-workers (Röder *et al.*, 1995; 1998, Plaschke *et al.*, 1995). The 15 wheat SSR markers tested are WMS2, WMS5, WMS6, WMS10, WMS11, WMS18, WMS30, WMS43, WMS46, WMS67, WMS88, WMS95, WMS120, WMS131 and WMS135. Among them, 10 markers showed high level of polymorphism with PIC values ranging from 0.291 and 0.790. The allele sizes of polymorphic markers determined precisely are shown in Table 3.3. The PCR amplifications of WMS10, WMS43, WMS67, WMS88 and WMS95 markers resulted monomorphic lengths, therefore, they were not included in characterizations.

SSR markers of WMS2 and WMS6 amplified two alleles with each individual plant. The primer set of WMS2 marker was reported amplifying two independent loci at chromosomes 3A and 3D as, *Xgwm2-3A* and *Xgwm2-3D* (Röder *et al.*, 1998). All of the DH individual lines in this study too, resulted two PCR products (Figure 3.6). Thus this particular observation with WMS2 marker is in accordance with the earlier report.

All the samples of DH lines 'Berkmen-469' and 'Çakmak-79' amplified the same 171 bp allele while in 'Kundur-1149' samples, the alleles of 165 bp and 169 bp were observed with WMS5 primer set.

Although WMS6 marker primer set also amplified two PCR products (Figure 3.6 and Dograr *et al.*, 2000), there is only one loci reported for the primer set at chromosome 4B, as *Xgwm6-4B* (Röder *et al.*, 1998). Our observation shows that, WMS6 is indeed amplifying two independent loci, since the plant materials used here are DH lines. This suggests that the appearances of two amplification products are not due to heterozygous alleles at a single locus. In this study, the locus of the second PCR product has not been identified. The allele sizes ranging from 195 bp to 207 bp were considered as alleles of one locus and the ones ranging in 199 bp and 223 bp, as another locus.

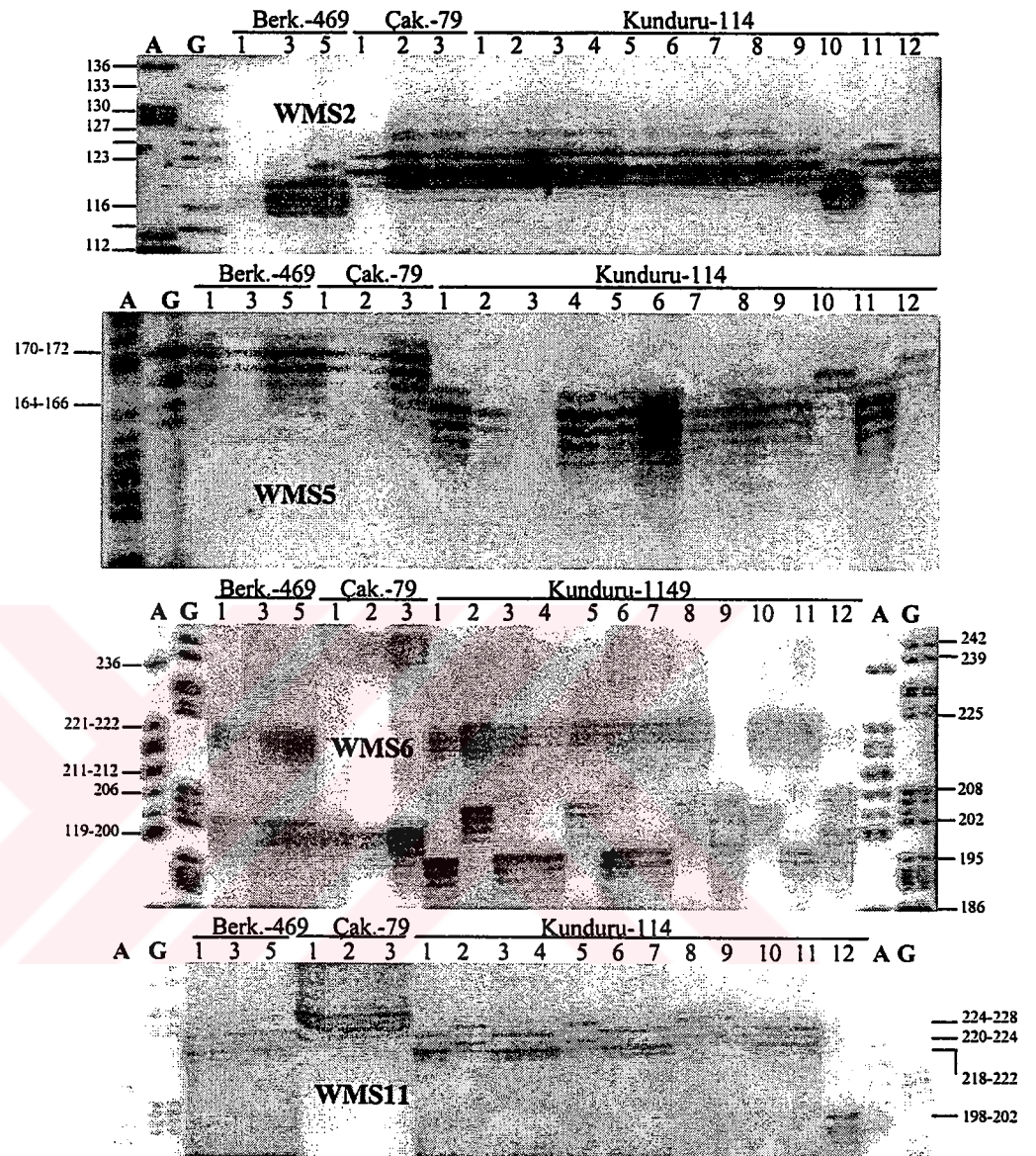


Figure 3.6 Autoradiograph of microsatellite alleles at each locus: PCR products amplified with the primer sets of different SSR markers, separated on a standard DNA sequencing gel. The numbers indicate the DH lines of each cultivar. A and G are the sequence products of M13mp18 ssDNA. The sizes are indicated as bp on the left or right site of the autoradiograph image.

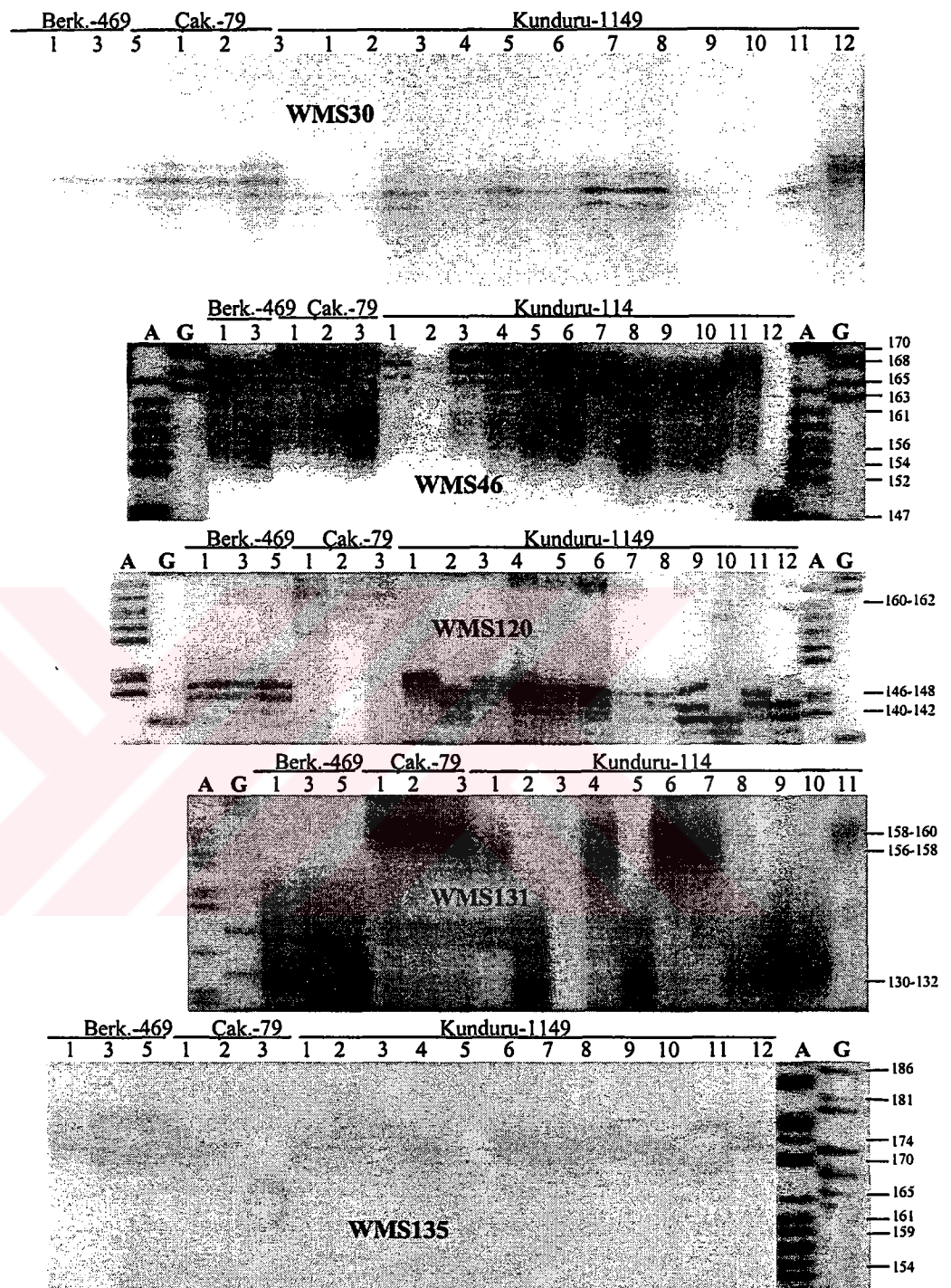


Figure 3.6 (Continued)

From the polymorphic microsatellites, WMS11 primer set generated a single allele from 'Berkmen-469' (224 bp long) and 'Çakmak-79' (228 bp long) samples while 'Kunduru-1149' population resulted in three alleles (201 bp, 222 bp and 224 bp) out of twelve individual plants. The results are indicated in Figure 3.6.

WMS18 amplified the allele of 181 bp in 'Berkmen-469', the allele of 183 bp in 'Çakmak-79' and the alleles 181 bp, 183 bp and 189 bp in 'Kunduru-1149' DH samples. The SSR marker WMS30 resulted in a single allele (210 bp sequence) in 'Berkmen-469' and 'Çakmak-79' DH samples while 'Kunduru-1149' DH lines showed two alleles of 208 bp and 210 bp in length. WMS46 had the amplification products of 167 bp from the 'Berkmen-469' and 169 bp from 'Çakmak-79' and the alleles 149 bp, 167 bp and 169 bp from 'Kunduru-1149' DH samples. WMS120 generated a single allele of size 145 bp and 162 bp in 'Berkmen-469' and in 'Çakmak-79' groups respectively and the alleles 143 bp, 146 bp and 148 bp from 'Kunduru-1149' population. The primer set of the marker WMS131 amplified a single band of 131 bp long in 'Berkmen-469'. 'Çakmak-79' group had the alleles of 157 bp and 159 bp. 'Kunduru-1149' population had the bands 131 bp and 155 bp amplified with the same primers. WMS135 set amplified a single band of 174 bp long in 'Berkmen-469'. 'Çakmak-79' group had the alleles of 165 bp and 174 bp while 'Kunduru-1149' samples amplified the bands of 172 bp and 174 bp with the same primers (Figure 3.6).

The data in Table 3.3 shows that three of the 'Berkmen-469' DH lines (B1, B3, B5) represent the same genotype. 'Çakmak-79' DH lines represent two different genotypes, the first two (Ç1, Ç2) are the same, Ç3 is a different genotype than the others. We had four times as many 'Kunduru-1149' DH lines than that of the other two varieties. 'Kunduru-1149' DH lines were shown to represent a total of 8 different genotypes. K1, K4, K6, K7, K11 gave identical alleles. All the others represent different genotypes. These results were summarized in a dendrogram (Figure 3.7).

Tablo 3.3 Allele sizes (bp) of DH lines using 12 microsatellite loci.

		Markers											
DHs	^a WMS2	^b WMS2	WMS5	WMS6	WMS6	WMS6	WMS11	WMS18	WMS30	WMS46	WMS120	WMS131	WMS135
B1	nd	Nd	171	203	223	223	224	181	210	167	145	131	174
B3	118	220	171	203	223	223	224	181	210	167	145	131	174
B5	118	220	171	203	223	223	224	181	210	nd	145	131	174
Ç1	122	224	171	201	247	247	228	183	210	169	162	159	174
Ç2	122	224	171	201	247	247	228	183	210	169	162	159	174
Ç3	122	224	171	201	247	247	228	183	nd	169	162	157	165
K1	122	224	165	195	223	223	222	183	nd	169	148	155	174
K2	122	224	165	207	223	223	224	181	208	167	146	131	174
K3	122	224	165	195	223	223	222	183	208	169	148	131	174
K4	122	224	165	195	223	223	222	183	208	169	148	155	174
K5	122	224	165	207	223	223	224	181	208	169	148	131	174
K6	122	224	165	195	223	223	222	183	208	169	148	155	174
K7	122	224	165	195	223	223	222	183	208	169	148	155	174
K8	122	224	165	207	223	223	224	181	208	167	148	131	172
K9	122	224	165	199	207	207	224	181	208	167	146	131	174
K10	118	220	169	205	223	223	222	183	nd	167	143	131	172
K11	122	224	165	195	223	223	222	183	208	169	148	155	174
K12	120	222	169	199	207	207	201	189	210	149	146	131	174
PIC	0.381	0.381	0.568	0.790	0.438	0.667	0.537	0.480	0.527	0.715	0.599	0.291	

(Ave: 0.531)

nd: not determined,

^a: locus *Xgwm2-3A* (Röder et al., 1998),

^b: locus *Xgwm2-3D* (Röder et al., 1998)

The highest variation within the cultivar was observed in 'Kunduru-1149'. Its DH lines show two sub-clusters with a phenon line of 0.65 (Figure 3.7). DH lines evaluated in this study are also being studied to exploit local varieties for improvement of germplasms in wheat breeding programs. 'Kunduru-1149' DH lines which have different morphological characters, such as dark color coleoptile or different spike and seed shapes, were confirmed as different genotypes according to the results obtained in this study.

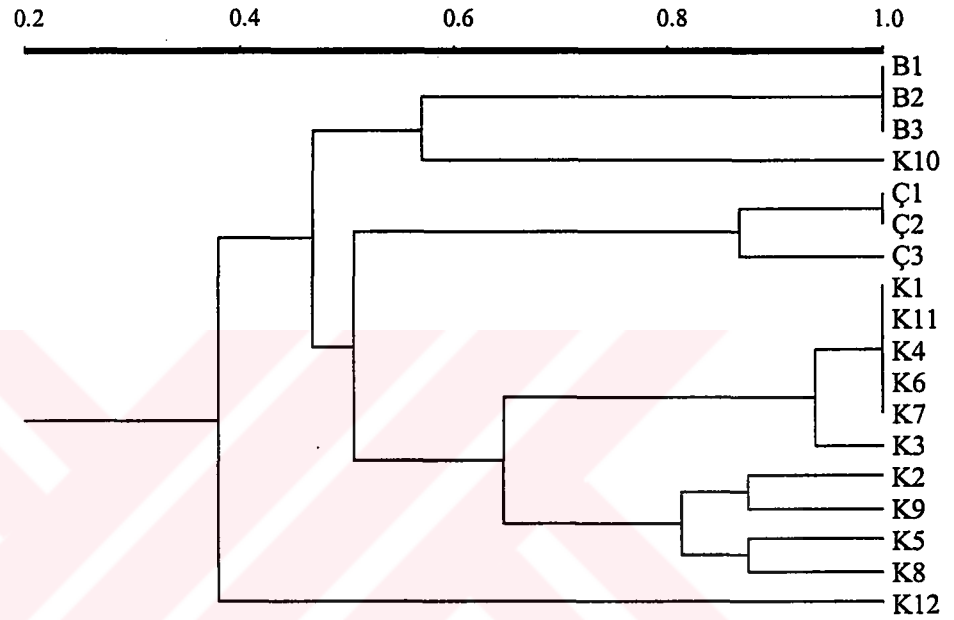


Figure 3.7 UPGMA clustering of DH individual lines based on simple matching coefficient similarity matrix of microsatellite data set. In the dendrogram, 'Çakmak-79' samples were represented by 'Ç', 'Berkmen-469' samples by 'B' and 'Kunduru-1149' samples by 'K'.

When assessing genetic relationships using microsatellite markers, especially for molecular ecology and evolution studies, using an appropriate distance or similarity algorithms specifically developed for the mutation models of microsatellite markers would be more proper. The algorithm used in this study (Goodman, 1997) is based on the stepwise mutation model applied to microsatellites (Slatkin, 1995) for "rho" calculations and genetic distance is denoted as $(\delta\mu)^2$ (Goldstein, et al., 1995). Thus, we believe a better estimation of genetic distance is obtained between the populations of DH lines in this study. 'Berkmen-469' x 'Çakmak-79' and 'Berkmen-469' x 'Kunduru-1149' pair-wise comparisons show a closer genetic relationship,

$(\delta\mu)^2$ of 0.84 and 0.85, respectively, than the other pair-wise comparisons. This might be due to a possible close relationship between 'Üveyik-162', the parent of 'Çakmak-79', and 'Berkmen-469', since they are both landrace selections. When K10 and K12, which are clustered very differently than the other 'Kunduru-1149' DH lines as can be seen in Figure 3.7, were omitted in the calculations of $(\delta\mu)^2$ based genetic distance, the 'Berkmen-469' and 'Kunduru-1149' became closer with $(\delta\mu)^2$ value of 0.65 (Figure 3.8).

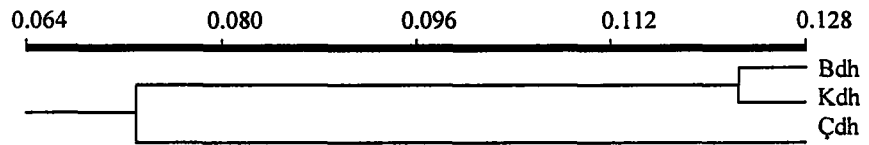


Figure 3.8 UPGMA dendrogram generated from NTSYS genetic analysis program by using $(\delta\mu)^2$ genetic length matrix.

This distance is probably meaningful, if K10 and K12 genotypes are represented with much lower frequencies in the real population than the one observed in this study. It is shown that the differentiation between the populations is the highest between 'Berkmen-469' and 'Çakmak-79' with "rho" values averaging over variance components and over loci were 0.996 and 0.767, respectively. On the other hand, 'Nei72' genetic distance calculations yielded different results, an opposite trend in relationships between populations were found (Table 3.4). The genetic relationship found between 'Berkmen-469' and 'Çakmak-79' using $(\delta\mu)^2$ is better reflecting the distances between these populations based on the accepted historical view and other studies (Zencirci *et al.*, 1994).

Table 3.4 The genetic analysis between the populations: "Rho" values, $(\Delta\mu)^2$ and 'Nei72' genetic distances (10000 permutations and bootstraps).

Populations	ρ^a	P^a	ρ^b	P^b	$(\delta\mu)^2$	'Nei72'
'Berkmen-469' x 'Çakmak-79'	0.99620	0.02430	0.76655	0.02100	0.83527	1.4655
'Berkmen-469' x 'Kunduru-1149'	0.42741	0.04790	0.34975	0.06570	0.85153	0.9175
'Çakmak-79' x 'Kunduru-1149'	0.55916	0.01180	0.44143	0.02950	1.41930	0.8872

^a: averaging over variance components, ^b: averaging over loci

When more samples are used in the analysis, it is obvious that more accurate genetic relation will be obtained. In this respect seven markers of our data have been combined with the data of our colleague Nurdan Dograr, who analysed the Turkish durum cultivars with these seven markers (Doğrar, 1999) and the results of NTSYS Analysis were found to be in agreement with our doubled haploid Turkish durum wheat variety analysis. Thus, it was possible to use the progenitors of the doubled haploid (DH) lines in harmony with the DH samples themselves. From these progenitor lines 'Berkmen-469' with 13 seeds, 'Çakmak-79' and 'Kunduru-1149' with 10 seeds were represented. Analysis was carried out in 7 microsatellite loci (WMS5, WMS6, WMS11, WMS37, WMS120, WMS131) and the dendrogram in Figure 3.9 was generated. It was found that, 'Çakmak-79' and 'Çakmak-79 DH' lines were clustered together while within the 'Berkmen-469' and 'Berkmen-469 DH' lines, samples belonging to 'Kunduru-1149' were scattered.

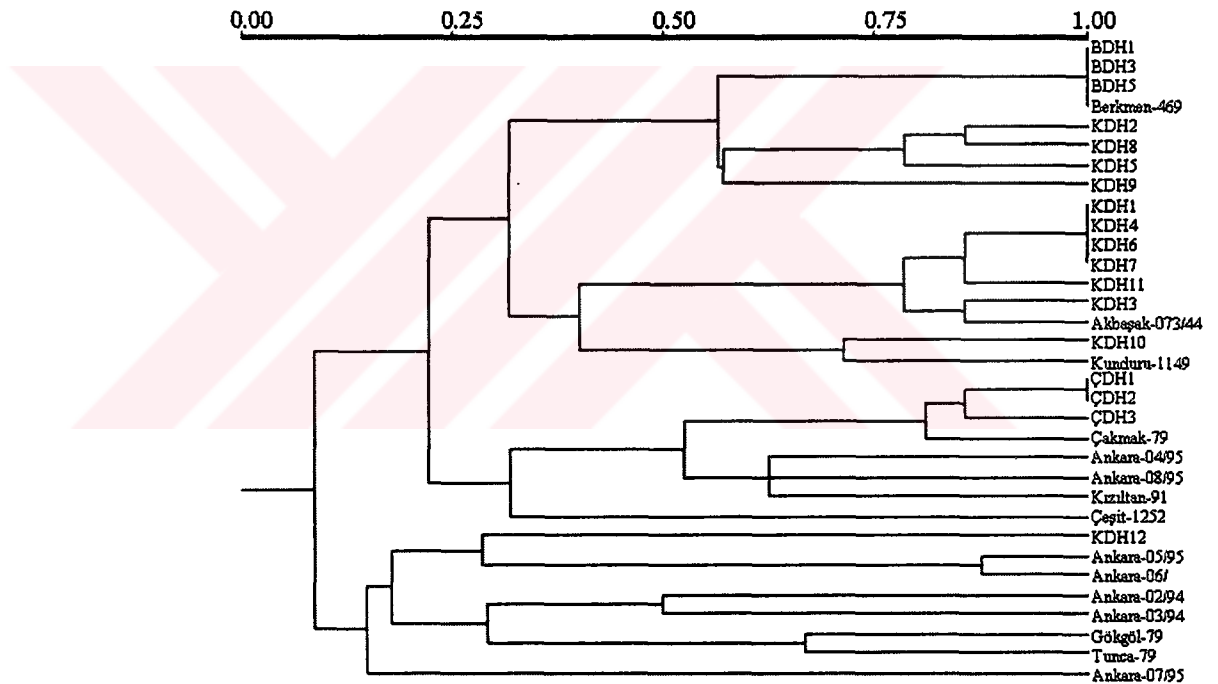


Figure 3.9 UPGMA generated dendrogram from NTSYS software determining the genetic relation of the DH lines with their progenitors and other Turkish cultivars.

This is the first molecular level analysis of Turkish durum wheat varieties and the genetic relation within and between these 3 cultivars ('Berkmen-469', 'Çakmak-79' and 'Kunduru-1149') in combination with their derivatives (DH lines). 'Çakmak-79' variety was shown to be relatively pure while 'Berkmen-469' and 'Kunduru-1149' were in a highly

genetically mixed form. The genetic relation of the three DH lines and their progenitors were also in accordance with the relation obtained from the DH samples alone in population level.

We believe the SSR marker system should be applied to all the cereal resources of Turkey which are the most critical food sources of today's Turkish population and the genetic characterization of these plants as well as the other important gene pools of the country should be completed as soon as possible.



3.3 ISOLATION AND SEQUENCING OF WHEAT RESISTANCE GENE ANALOGS

Sequences of the few functionally defined resistance genes and a much larger number of resistance gene homologues were utilised successfully, in a couple of laboratories around the world, in amplifying conserved motives in the NBS domain of the plant species soybean, potato, rice, barley, *Arabidopsis thaliana*, tomato and citrus (Leister *et al.*, 1996, 1998; Yu *et al.*, 1996; Deng *et al.*, 2000; Hayes *et al.*, 2000; Pan *et al.*, 2000). A considerable number of sequences were accumulated on specific plant species like that of *A. thaliana*, soybean and *oryza*.

There are efforts to generate RGAs from wheat and find linkages to important fungal pathogens, but up until now the information on RGA sequences from this immensely important crop are limited (http://www.ncgr.org/research/rgenes/rgenes_data8.html and Leister *et al.*, 1998) There is a report on a NBS-LRR sequence at the *Cre3* nematode resistance locus (Lagudah *et al.*, 1997). Spielmeyer *et al.*, 1998, targeted the leaf rust resistance locus (*Lr34* locus), but they were not successful. However, they generated a detailed RFLP map of the distal end of the short arm of chromosome 1D of *Aegilops tauschii*, which is the D genome donor of wheat (Spielmeyer *et al.*, 2000). In their recent publication, they have mapped disease resistance genes effective against leaf, stripe and stem rusts in close proximity to the prolamin loci which is a highly recombinogenic region. Lrk10, Lr21 as well as a yellow rust resistance gene analog probe (from the wheat cultivar 'Moro' carrying the *Yr10* gene for stripe rust resistance, mapped distal to the gamma-gliadins on the end of chromosome 1BS) were mapped in this distal end of chromosome 1DS by RFLP mapping in D genome donor *A. tauschii*. As a control, they utilized the homozygous susceptible wheat 'Avocet' as well as the near-isogenic line of 'Avocet' carrying *Yr10* gene (Spielmeyer *et al.*, 2000).

The recent findings above give strong support on our approach of looking for RGAs in the cultivars 'Pastor' with a major and two minor resistance genes and 'Avocet' with two minor resistance genes to the fungal pathogen *Puccinia striiformis*, the causal agent of stripe rust. We also utilised the complete-susceptible variety 'Chinese spring' and the wild wheat varieties, *T. monococcum* (Acc. No. TR36938), *T. dicoccoides* (Acc. No. TR02771), *A. speltooides*.

3.3.1. PCR Products of RGAs

The PCR products were obtained as explained in the Materials and Methods section. The conditions were successfully optimised that all the primer combinations resulted ample amounts of PCR products (Figure 3.10). However, the very same primers were not successful in isolating RGAs from the wheat genome, in the hands of the designers of the primers (Leister *et al.*, 1998). We have tried 7 primer combinations for PCR amplifications of 'Pastor' and 'Avocet', 4 primer combination for 'Chinese spring' and only one primer combination for the wild relatives of wheat. The sizes of PCR products were ranged from ~ 200 bp to ~1000 bp. The bands in the 500 bp regions were isolated and cloned as disease resistance gene candidates. From each PCR product many colonies obtained after cloning; 188 clones from the cultivar 'Pastor', 192 clones from 'Avocet', 121 clones from 'Chinese spring', 100 clones from *T. monococcum* and 80 clones from *T. dicoccoides* were saved for further analysis, summing up 686 clones. From each cloning experiment randomly selected few number of white clones were sequenced (Table 3.5). From the cultivar 'Pastor' 40 clones, from 'Avocet' 33 clones, from 'Chinese spring' 10 clones and 3 and 5 clones from the wild types *T. monococcum* and *T. dicoccoides*, respectively, were sequenced. The total number of RGA sequences obtained are 92. The sequences of all the RGAs are presented in the Appendices A to E.

All the ligation reactions were made with PCR 2.1 vector (Invitrogen) except that of the duplicate experiment with the primer set S1/AS2 where pGEM-T-Easy vector (Promega) was utilized. In this specific experiment the 9 clones sequenced (out of 12) were all generated the same sequence. Our degenerate primers were expected to result in different homolog genes or sequences of some non-functional pseudogenes from diverse regions of the genome. However, unlike all the other sequences revealed, in this particular case the same fragments were cloned, as was obvious with the sequences. The remaining 3 clones were not sequenced to further assess the situation. Our RGA, named P2-8, had a poly(A) sequence at the 161-168 bp region. From the GenBank sequence homology analysis we retrieved the following hits:

Observed Homology:

- Triticum aestivum resistance protein (RGA1) gene, partial cds.
ACCESSION AF087518
- Hordeum vulgare NBS-LRR type resistance protein (b2) gene, partial cds.
ACCESSION AF032680
- Triticum aestivum resistance protein (RGA3) gene, partial cds.
ACCESSION AF087520
- Hordeum vulgare NBS-LRR type resistance protein (b5) gene, partial cds.
ACCESSION AF032683
- Oryza sativa clone RGA7 resistance protein, partial cds.
ACCESSION AF074888
- Oryza sativa NBS-LRR type resistance protein (r4) gene, partial cds.
ACCESSION AF032691
- Oryza sativa NBS-LRR-like protein (YR19) gene, partial cds.
ACCESSION AF220741

- *Oryza sativa* subsp. *japonica* NBS-LRR-like protein (YR38) pseudogene, partial sequence.
ACCESSION AF227002
- *Oryza sativa* NBS-LRR-like protein (YR28) pseudogene, partial sequence.
ACCESSION AF220748
- *Oryza sativa* NBS-LRR-like protein (YR29) gene, complete cds.
ACCESSION AF220749
- *Oryza sativa* NBS-LRR-like protein (YR16) gene, partial cds.
ACCESSION AF220739
- *Oryza sativa* NBS-LRR-like protein (YR26) gene, partial cds.
ACCESSION AF220747
- *Oryza sativa* NBS-LRR-like protein (YR18) gene, partial cds.
ACCESSION AF220740
- *Oryza sativa* NBS-LRR-like protein (YR14) gene, partial cds.
ACCESSION AF220738
- *Oryza officinalis* NBS-LRR-like protein gene, partial cds.
ACCESSION AF226999
- *Triticum aestivum* resistance protein (RGA4) gene, partial cds.
ACCESSION AF087521

The sequence homology of P2-8 to the RGA1 sequence in the GenBank was very high, with *expect ratio* of 0.0, with an *identity rate* of 503/515 (97%). Interestingly however, there was a low *gap rate* of 1/515 (0%) at the Poly(A) region mentioned above, starting at 161 bp, and the point mutation at that position (single nucleotide deletion), has resulted the distortion of amino acid sequence then on. This situation was observed in all the nine clones sequenced. A better explanation of the presence of this mutation can be possible only after the function of the protein coded by the gene is determined.

Out of 92 RGA sequences 24 of them were uninterrupted with the stop codons, but the others possessed stop codons. The presence or the absence of stop codons were confirmed by sequencing all of them for the second time from the opposite directions. Thus the ones with no stop codons are considered as R genes, which need confirmation by linkage analysis studies. The interrupted RGAs are considered to be pseudogenes, genes which have lost their functions in the process of evolution.

There were three RGAs coding for amino acids obtained from 'Pastor', (P2-4 from the primer set S1/AS2; P6-4 and P6-5 from the S1/AS6 primer set). From the cultivar 'Avocet', 10 uninterrupted sequences were generated (A2-1, A2-2, A2-4, A3-1, A4-2, A4-6, A6-1, A6-2, A6-4 and A8-4). From the 3 uninterrupted sequences of 'Chinese spring' variety CS2-1 was from the primer set S1/AS2, CS6-4 from S1/AS6 and CS8-1 from AS/AS8. From the clones using S1/AS3 in *T. dicoccoides* and in *T. monococcum* 2 uninterrupted sequences from each wild type were generated, TD3-3; TD3-16 and TM3-2; TM3-10, respectively (Table 3.5).

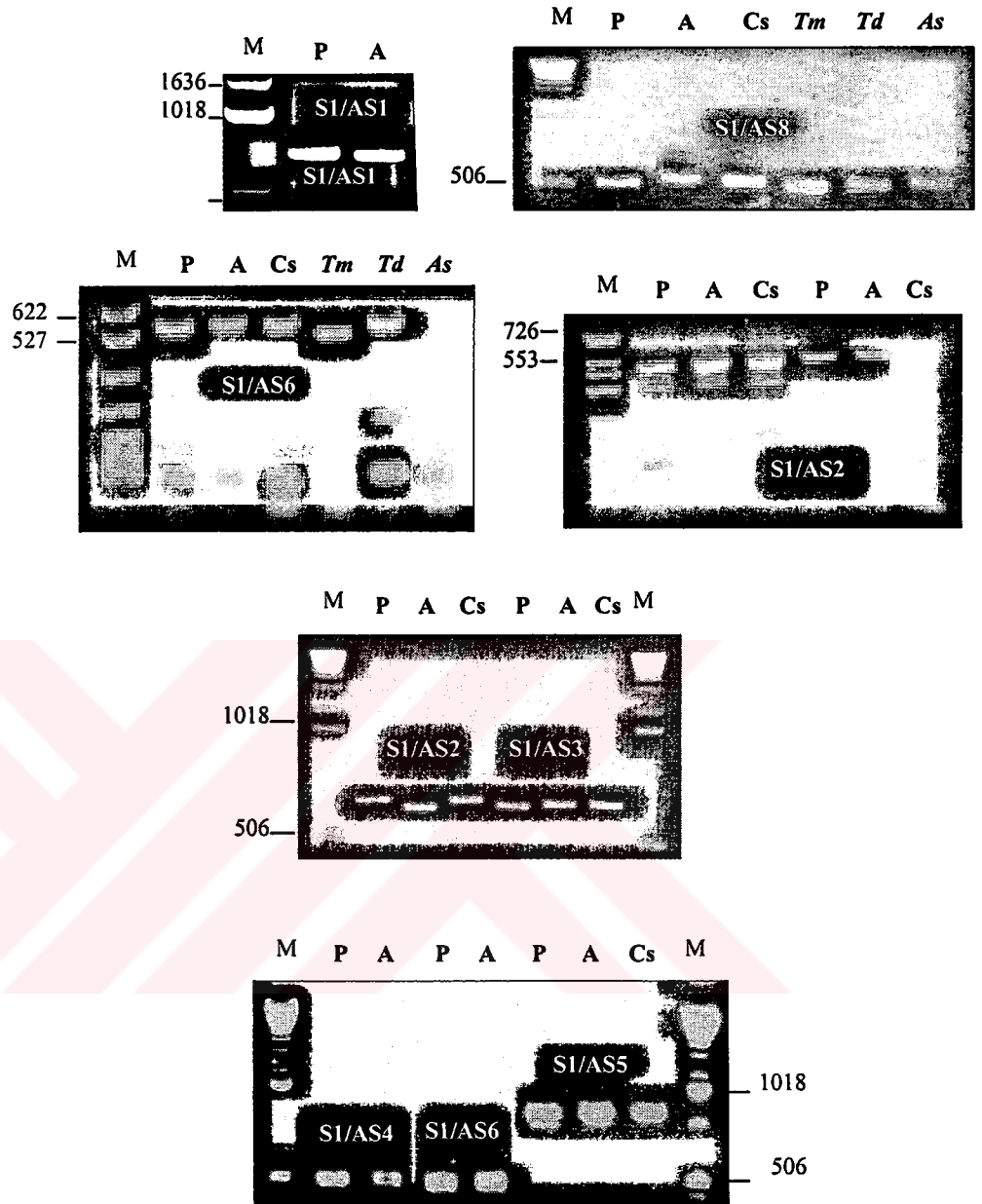


Figure 3.10 PCR products obtained using degenerate primers targeting the conserved NBS and LLR regions. The primer combinations are indicated on the photographs of 1.5 % agarose gels. Lanes are labeled as M: Marker, P: Pastor, A: Avocet, Cs: Chinese spring, *Tm* : *T. monococcum*, *Td*: *T. dicoccoides*, *As*: *A. speltoides*.

Table 3.5 The clones selected for sequencing and their translation status

PCR product	Sequenced clone	Appendix	Interrupted	Uninterrupted
Pastor/S1-AS1	P1-1	E	+	
	P1-2	E	+	
	P1-4	E	+	
	P1-5	E	+	
	P1-6	E	+	
Pastor/S1-AS2	P2-1	E	+	
	P2-2	E	+	
	P2-3	E	+	
	P2-4	E		+
	P2-5	E	+	
Pastor/S1-AS3	P2-8	E	+	
	P3-1	E	+	
	P3-2	E	+	
	P3-3	E	+	
	P3-4	E	+	
Pastor/S1-AS4	P3-5	E	+	
	P3-6	E	+	
	P4-2	E	+	
	P4-3	E	+	
	P4-4	E	+	
Pastor/S1-AS5	P4-5	E	+	
	P4-6	E	+	
	P5-1	E	+	
Pastor/S1-AS6	P5-2	E	+	
	P6-1	E	+	
	P6-4	E		+
Pastor/S1-AS8	P6-5	E		+
	P8-1	E	+	
	P8-2	E	+	
	P8-3	E	+	
	P8-4	E	+	
Avocet/S1-AS1	P8-5	E	+	
	A1-1	C	+	
	A1-2	C	+	
	A1-4	C	+	
	A1-5	C	+	
Avocet/S1-AS2	A1-6	C	+	
	A2-1	C		+
	A2-2	C		+
	A2-3	C	+	
	A2-4	C		+
Avocet/S1-AS3	A2-6	C	+	
	A3-1	C		+
	A3-2	C	+	
	A3-3	C	+	
	A3-4	C	+	
Avocet/S1-AS4	A3-5	C	+	
	A3-6	C	+	
	A4-1	C	+	
	A4-2	C		+
	A4-3	C	+	
Avocet/S1-AS5	A4-4	C	+	
	A4-5	C	+	
	A4-6	C		+
	A5-1	C	+	
	A5-2	C	+	
Avocet/S1-AS6	A6-1	C		+
	A6-2	C		+
	A6-4	C		+
	A6-6	C	+	
	A8-1	C	+	
Avocet/S1-AS8	A8-2	C	+	
	A8-3	C	+	
	A8-4	C		+
	A8-5	C	+	
	CS2-1	C		+
C. spring/S1-AS2	CS2-3	C	+	
	CS2-4	C	+	
	CS2-5	C	+	
	CS2-6	C	+	
	CS5-1	C	+	
C. spring/S1-AS5	CS5-2	C	+	
C. spring/S1-AS6	CS6-1	C	+	
C. spring/S1-AS6	CS6-2	C	+	
C. spring/S1-AS6	CS6-4	C		+
C. spring/S1-AS8	CS8-1	C		+
T. dicoccoides/S1-AS3	TD3-1	A	+	
	TD3-2	A	+	
	TD3-3	A		+
	TD3-12	A	+	
	TD3-16	A		+
T. monococcum/S1-AS3	TM3-1	B	+	
	TM3-2	B		+
	TM3-10	B		+

3.3.2 Alignments of RGA Sequences

All the RGA sequences obtained were aligned together by the aid of the software ClustalX (1.81) and the alignments were presented in Appendix F. The homology was specifically high at the 5'-end of the sequences, which is very meaningful, considering the same forward direction primer (S1), targeting the P-loop motive, was used in all the amplifications. Based on the sequence similarities, the RGAs were clustered as various groups, shown in Figure 3.11. The sequences presented in Appendices I-V were all analyzed according to different criteria. First, all the sequences were aligned in ClustalX 1.81 software and the clustering were done in nucleotide sequence level, without taken into account whether the sequences were interrupted with a stop codon or not (Appendix VI).

ClustalX 1.81 software was used to align all the 84 resistance gene analog sequences that were obtained using Pastor, Avocet, Chinese spring wheat varieties and the wild types *T. monococcum* and *T. dicoccoides*. The same software was utilized to generate their dendrogram reflecting their genetic similarities based on Neighbor-joining algorithm. The reliability of the tree was then established by conducting default (1000) bootstrap re-sampling steps of the program.

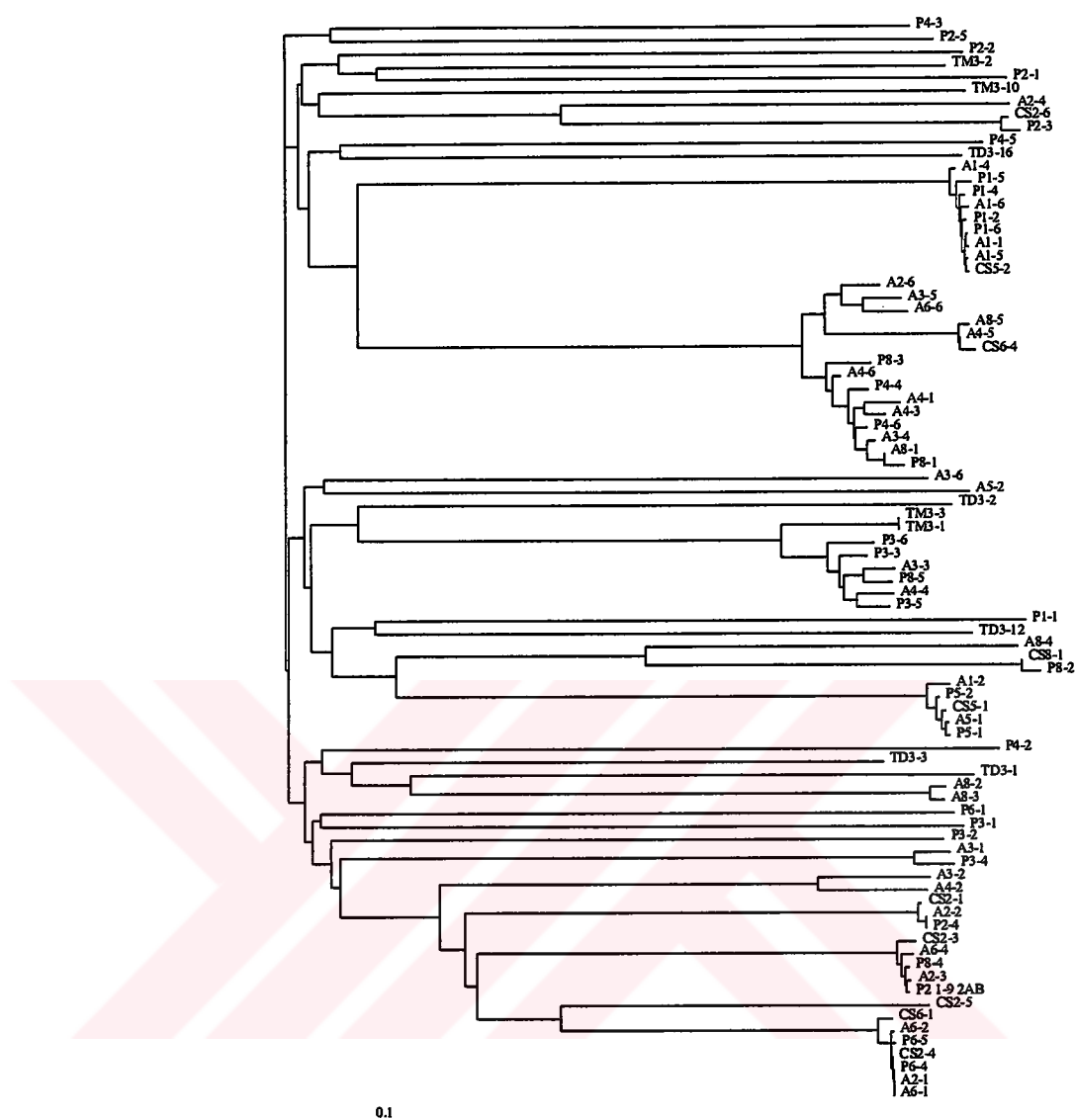


Figure 3.11 Genetic similarity of all 84 RGAs based on neighbor-joining method of ClustalX 1.81. Dendrogram was obtained from TreeView.

3.3.3 Homologies Obtained with the GenBank Sequences

Only 84 RGA sequences were used in the cluster analysis, since the Pastor/S1-AS2 PCR products cloned into pGEM-T Easy vector were resulted in 9 identical RGA sequences. Based on the clustering of the RGAs and their best blast hits, the sequences with high homologies were grouped accordingly.

Table 3.6 Homologous sequences found in GenBank, using Blast algorithm.

CLONE	BLAST HITS	ACCESSION NUMBER	SCORE (BITS)	E VALUE	IDENTITIES
P1-1	Avena sativa cultivar Sun II NBS-LRR type disease resistance protein O2 gene, partial cds	gb AF078874.1	101	3e-19	63/67 (94%)
	Avena sativa cultivar Sun II NBS-LRR type disease resistance protein O1 gene, partial cds	gb AF032679.1	75.8	2e-11	44/46 (95%)
P1-2	Hordeum vulgare NBS-LRR type desistance protein (b1) gene, partial cds	gb AF032679.1	309	7e-82	198/212 (93%)
	Avena sativa cultivar Sun II NBS-LRR type disease resistance protein O2 gene, partial cds	gb AF078874.1	303	4e-80	216/237 (91%)
	Avena sativa cultivar Sun II NBS-LRR type disease resistance protein O1 gene, partial cds	gb AF078873.1	303	4e-80	216/237 (91%)
	Avena sativa cultivar Sun II NBS-LRR type disease resistance protein O2 gene, partial cds	gb AF056157.1	305	1e-06	123/154 (79%)
P1-4	Zea mays disease resistance gene analog PIC17 (pic17) gene, partial cds	gb AF032679.1	317	3e-84	199/212 (93%)
	Hordeum vulgare NBS-LRR type resistance protein (b1) gene, partial cds	gb AF078874.1	311	2e-82	217/237 (91%)
	Avena sativa cultivar Sun II NBS-LRR type disease resistance protein O2 gene, partial cds	gb AF056157.1	540	6e-05	42/49 (89%)
P1-5	Zea mays disease resistance gene analog PIC17 (pic17) gene, partial cds	gb AF032679.1	309	7e-82	198/212 (93%)
	Hordeum vulgare NBS-LRR type resistance protein (b1) gene, partial cds	gb AF078874.1	303	4e-80	216/237 (91%)
	Avena sativa cultivar Sun II NBS-LRR type disease resistance protein O2 gene, partial cds	gb AF056157.1	61.9	3e-07	49/55 (89%)
P1-6	Zea mays disease resistance gene analog PIC17 (pic17) gene, partial cds	gb AF032679.1	317	3e-84	199/212 (93%)
	Hordeum vulgare NBS-LRR type resistance protein (b1) gene, partial cds	gb AF078874.1	311	2e-82	217/237 (91%)
	Avena sativa cultivar Sun II NBS-LRR type disease resistance protein O2 gene, partial cds	gb AF056157.1	60	1e-06	123/154 (79%)
P2-1	Zea mays disease resistance gene analog PIC17 (pic17) gene, partial cds				
P2-2	No Significant Hit				
P2-3	No Significant Hit				
	Triticum aestivum resistance protein (RGA1) gene, partial cds	gb AF087518.1	898	0.0	485/497 (97%)
	Hordeum vulgare NBS-LRR type resistance protein (b2) gene, partial cds	gb AF032680.1	605	e-171	439/485 (90%)
	Triticum aestivum resistance protein (RGA3) gene, partial cds	gb AF087520.1	232	1e-58	204/233 (87%)
P2-4	Triticum aestivum resistance protein (RGA2) gene, partial cds	gb AF087519.1	839	0.0	474/491 (96%)
	Hordeum vulgare NBS-LRR type resistance protein (b4) gene, partial cds	gb AF032682.1	747	0.0	458/485 (94%)
	Oryza sativa NBS-LRR type resistance protein (r1) gene, partial cds	gb AF032688.1	305	1e-80	259/294 (88%)
P2-5	No Significant Hit				
P2-8	Triticum aestivum resistance protein (RGA1) gene, partial cds	gb AF087518.1	934	0.0	503/515 (97%)
	Hordeum vulgare NBS-LRR type resistance protein (b2) gene, partial cds	gb AF032680.1	634	e-180	457/504 (90%)
	Triticum aestivum resistance protein (RGA3) gene, partial cds	gb AF087520.1	232	1e-58	204/233 (87%)
P3-1	No Significant Hit				
P3-2	No Significant Hit				
P3-3	No Significant Hit				
P3-4	Hordeum vulgare NBS-LRR type resistance protein (b6) gene, partial cds	gb AF032684.1	238	4e-74	425/519 (81%)
	Oryza sativa NBS-LRR-like protein (YR5) gene, partial cds	gb AF220731.1	44.1	0.062	25/26 (96%)
	Brassica napus chromosome N1 disease resistance gene homolog 9N (RPM1) gene	gb AF105140.1	44.1	0.062	22/22 (100%)
P3-5	No Significant Hit				
P3-6	No Significant Hit				
P4-2	No Significant Hit				
P4-3	No Significant Hit				
P4-4	Danio rerio mRNA for protein tyrosine kinase (jak2b), partial	emb AJ005691.1	40.1	0.45	20/20 (100%)
P4-5	Medicago sativa mitogen-activated protein kinase homologue (TDY1) gene	gb AF129087.1	42.1	0.23	21/21 (100%)
P4-6	No Significant Hit				
P5-1	Hordeum vulgare NBS-LRR type resistance protein (b1) gene, partial cds	gb AF032679.1	125	2e-26	66/67 (98%)
	Avena sativa cultivar Sun II NBS-LRR type disease resistance	gb AF078874.1	79.8	1e-12	116/141 (82%)
	Avena sativa cultivar Sun II NBS-LRR type disease resistance protein O1 gene, partial	gb AF078873.1	79.8	1e-12	116/141 (82%)
P5-2	Hordeum vulgare NBS-LRR type resistance protein (b1) gene, partial cds	gb AF032679.1	117	4e-24	65/67 (97%)
	Avena sativa cultivar Sun II NBS-LRR type disease resistance protein O2 gene, partial	gb AF078874.1	67.9	3e-09	58/66 (87%)
	Avena sativa cultivar Sun II NBS-LRR type disease resistance protein O1 gene, partial	gb AF078873.1	67.9	3e-09	58/66 (87%)
P6-1	No Significant Hit				
P6-4	Capsicum chacoense disease resistance protein BS2 (BS2) mRNA, complete cds	gb AF202179.1	46.1	0.015	29/31 (93%)
	Vigna unguiculata clone KIND11 disease resistance protein homolog gene, partial cds	gb AF141011.1	42.1	0.24	21/21 (100%)
P6-5	Hordeum vulgare genomic DNA fragment, clone MWG0851.rev	emb AJ234592.1	131	3e-28	238/295 (80%)
	Capsicum chacoense disease resistance protein BS2 (BS2) mRNA, complete cds	gb AF202179.1	46.1	0.016	29/31 (93%)
	Vigna unguiculata clone KIND11 disease resistance protein homolog gene, partial cds	gb AF141011.1	42.1	0.25	21/21 (100%)
P8-1	No Significant Hit				
P8-2	No Significant Hit				
P8-3	No Significant Hit				
P8-4	Triticum aestivum resistance protein (RGA1) gene, partial cds	gb AF087518.1	789	0.0	436/450 (96%)
	Hordeum vulgare NBS-LRR type resistance protein (b2) gene, partial cds	gb AF032680.1	529	e-148	398/443 (89%)
	Triticum aestivum resistance protein (RGA3) gene, partial cds	gb AF087520.1	232	1e-58	204/233 (87%)
P8-5	No Significant Hit				
A1-1	Hordeum vulgare NBS-LRR type resistance protein (b1) gene, partial cds	gb AF032679.1	317	3e-84	199/212 (93%)
	Avena sativa cultivar Sun II NBS-LRR type disease resistance protein O2 gene, partial	gb AF078874.1	311	2e-82	217/237 (91%)
	Zea mays disease resistance gene analog PIC17 (pic17) gene, partial cds	gb AF056157.1	60.0	1e-06	123/154 (79%)
A1-2	Hordeum vulgare NBS-LRR type resistance protein (b1) gene, partial cds	gb AF032679.1	125	2e-06	66/67 (98%)
	Avena sativa cultivar Sun II NBS-LRR type disease resistance protein O2 gene, partial	gb AF078874.1	79.8	1e-12	116/141 (82%)
	Avena sativa cultivar Sun II NBS-LRR type disease resistance protein O1 gene, partial	gb AF078873.1	79.8	1e-12	116/141 (82%)
A1-4	Hordeum vulgare NBS-LRR type resistance protein (b1) gene, partial cds	gb AF032679.1	270	3e-70	196/216 (90%)
	Avena sativa cultivar Sun II NBS-LRR type disease resistance protein O2 gene, partial	gb AF078874.1	240	3e-61	193/217 (88%)
	Zea mays disease resistance gene analog PIC17 (pic17) gene, partial cds	gb AF056157.1	58.0	2e-06	122/153 (79%)
A1-5	Hordeum vulgare NBS-LRR type resistance protein (b1) gene, partial cds	gb AF032679.1	309	7e-82	198/212 (93%)
	Avena sativa cultivar Sun II NBS-LRR type disease resistance protein O2 gene, partial	gb AF078874.1	303	4e-80	216/237 (91%)
	Zea mays disease resistance gene analog PIC17 (pic17) gene, partial cds	gb AF056157.1	60.0	1e-06	123/154 (79%)
A1-6	Hordeum vulgare NBS-LRR type resistance protein (b1) gene, partial cds	gb AF032679.1	309	7e-82	198/212 (93%)
	Avena sativa cultivar Sun II NBS-LRR type disease resistance protein O2 gene, partial	gb AF078874.1	299	7e-79	211/231 (91%)
	Zea mays disease resistance gene analog PIC17 (pic17) gene, partial	gb AF056157.1	52.0	3e-04	122/154 (79%)

CLONE	BLAST HITS	ACCESSION NUMBER	SCORE (BITS)	E VALUE	IDENTITIES
A2-1	Capsicum chacoense disease resistance protein BS2 (BS2) mRNA, complete cds	gb AF202179.1	46.1	0.015	29/31 (93%)
	Vigna unguiculata clone KIND11 disease resistance protein homolog	gb AF141011.1	42.1	0.23	21/21 (100%)
A2-2	Triticum aestivum resistance protein (RGA2) gene, partial cds	gb AF087519.1	870	0.0	490/507 (96%)
	Hordeum vulgare NBS-LRR type resistance protein (b4) gene, partial cds	gb AF032682.1	771	0.0	473/501 (94%)
	Oryza sativa NBS-LRR type resistance protein (r1) gene, partial cds	gb AF032688.1	325	1e-86	293/336 (87%)
A2-4	Triticum aestivum resistance protein (RGA2) gene, partial cds	gb AF087519.1	837	0.0	465/480 (96%)
	Hordeum vulgare NBS-LRR type resistance protein (b4) gene, partial cds	gb AF032682.1	743	0.0	447/471 (94%)
	Oryza sativa NBS-LRR-like protein (YR7) gene, partial cds	gb AF220733.1	363	5e-98	318/363 (87%)
A2-6	No Significant Hit				
A3-1	Hordeum vulgare NBS-LRR type resistance protein (b6) gene, partial cds	gb AF032684.1	315	1e-83	429/519 (82%)
	Oryza sativa NBS-LRR-like protein (YR5) gene, partial cds	gb AF220731.1	44.1	0.063	5/26 (96%)
	Brassica napus chromosome N1 disease resistance gene homolog 9N (RPM1) gene,	gb AF105140.1	42.1	0.25	21/21 (100%)
A3-2	Elaeis guineensis resistance protein (RGA4) gene, partial cds	gb AF197924.1	38.2	3.8	19/19 (100%)
	O.sativa L. putative disease-resistance gene, clone p8558-12, partial	emb Y09811.1	38.2	3.8	22/23 (95%)
	Vigna unguiculata DNA, nucleotide-binding site sequence (resistance gene analog), clone:CRGA 4	dbj AB020486.1	38.2	3.8	19/19 (100%)
A3-3	No Significant Hit				
A3-4	No Significant Hit				
A3-5	No Significant Hit				
A3-6	No Significant Hit				
A4-1	No Significant Hit				
A4-2	Hordeum vulgare resistance gene analog PIC25 (PIC25) gene, partial cds	gb AF146272.1	38.2	3.9	22/23 (95%)
	Hordeum vulgare NBS-LRR type resistance protein (b9) gene, partial cds	gb AF032687.1	38.2	3.9	22/23 (95%)
	Spinach latent virus putative polymerase and putative 2b protein genes, complete cds	gb U93193.1	38.2	3.9	22/23 (95%)
A4-3	Danio rerio mRNA for protein tyrosine kinase (jak2b), partial	emb AJ005691.1	40.1	0.73	20/20 (100%)
A4-4	No Significant Hit				
A4-5	No Significant Hit				
A4-6	No Significant Hit				
A5-1	Hordeum vulgare NBS-LRR type resistance protein (b1) gene, partial cds	gb AF032679.1	125	2e-26	66/67 (98%)
	Avena sativa cultivar Sun II NBS-LRR type disease resistance protein O2 gene, partial	gb AF078874.1	77.8	4e-12	100/120 (83%)
	Avena sativa cultivar Sun II NBS-LRR type disease resistance protein O1 gene, partial	gb AF078873.1	77.8	4e-12	100/120 (83%)
A5-2	No Significant Hit				
A6-1	Capsicum chacoense disease resistance protein BS2 (BS2) mRNA, complete cds	gb AF202179.1	46.1	0.016	29/31 (93%)
	Vigna unguiculata clone KIND11 disease resistance protein homolog gene, partial cds	gb AF141011.1	42.1	0.24	21/21 (100%)
A6-2	Capsicum chacoense disease resistance protein BS2 (BS2) mRNA, complete cds	gb AF202179.1	46.1	0.016	29/31 (93%)
	Vigna unguiculata clone KIND11 disease resistance protein homolog gene, partial cds	gb AF141011.1	42.1	0.24	21/21 (100%)
A6-4	Triticum aestivum resistance protein (RGA1) gene, partial cds	gb AF087518.1	914	0.0	494/508 (97%)
	Hordeum vulgare NBS-LRR type resistance protein (b2) gene, partial cds	gb AF032680.1	628	e-178	455/504 (90%)
	Triticum aestivum resistance protein (RGA3) gene, partial cds	gb AF087520.1	232	1e-58	204/233 (87%)
A6-6	No Significant Hit				
A8-1	No Significant Hit				
A8-2	No Significant Hit				
A8-3	No Significant Hit				
A8-4	No Significant Hit				
A8-5	No Significant Hit				
CS2-1	Triticum aestivum resistance protein (RGA2) gene, partial cds	gb AF087519.1	894	0.0	500/517 (96%)
	Hordeum vulgare NBS-LRR type resistance protein (b4) gene, partial cds	gb AF032682.1	801	0.0	488/516 (94%)
	Oryza sativa NBS-LRR type resistance protein (r1) gene, partial cds	gb AF032688.1	361	2e-97	308/350 (88%)
CS2-3	Triticum aestivum resistance protein (RGA1) gene, partial cds	gb AF087518.1	848	0.0	459/468 (98%)
	Hordeum vulgare NBS-LRR type resistance protein (b2) gene, partial cds	gb AF032680.1	561	e-157	422/467 (90%)
	Triticum aestivum resistance protein (RGA3) gene, partial cds	gb AF087520.1	224	3e-56	203/233 (87%)
CS2-4	Capsicum chacoense disease resistance protein BS2 (BS2) mRNA, complete cds	gb AF202179.1	46.1	0.014	29/31 (93%)
	Vigna unguiculata clone KIND11 disease resistance protein homolog gene, partial cds	gb AF141011.1	42.1	0.21	21/21 (100%)
CS2-5	Triticum aestivum resistance protein (RGA2) gene, partial cds	gb AF087519.1	42.1	0.24	21/21 (100%)
	Oryza sativa NBS-LRR-like protein (YR22) gene, partial cds	gb AF220744.1	38.2	3.7	22/23 (95%)
	Oryza sativa NBS-LRR-like protein (YR20) gene, partial cds	gb AF220742.1	38.2	3.7	22/23 (95%)
CS2-6	Triticum aestivum resistance protein (RGA1) gene, partial cds	gb AF087518.1	555	e-156	302/308 (98%)
	Hordeum vulgare NBS-LRR type resistance protein (b2) gene, partial	gb AF032680.1	406	e-111	275/297 (92%)
	Triticum aestivum resistance protein (RGA3) gene, partial cds	gb AF087520.1	190	2e-46	193/224 (86%)
CS5-1	Hordeum vulgare NBS-LRR type resistance protein (b1) gene, partial cds	gb AF032679.1	109	1e-21	61/63 (96%)
	Avena sativa cultivar Sun II NBS-LRR type disease resistance protein O2 gene, partial	gb AF078874.1	63.9	6e-08	114/141 (80%)
	Avena sativa cultivar Sun II NBS-LRR type disease resistance protein O1 gene, partial	gb AF078873.1	63.9	6e-08	114/141 (80%)
CS5-2	Hordeum vulgare NBS-LRR type resistance protein (b1) gene, partial cds	gb AF032679.1	117	4e-24	65/67 (97%)
	Avena sativa cultivar Sun II NBS-LRR type disease resistance protein O2 gene, partial	gb AF078874.1	67.9	3e-09	58/66 (87%)
	Avena sativa cultivar Sun II NBS-LRR type disease resistance protein O1 gene, partial	gb AF078873.1	67.9	3e-09	58/66 (87%)
CS6-1	Capsicum chacoense disease resistance protein BS2 (BS2) mRNA, complete cds	gb AF202179.1	46.1	0.016	29/31 (93%)
	Vigna unguiculata clone KIND11 disease resistance protein homolog gene, partial cds	gb AF141011.1	42.1	0.24	21/21 (100%)
CS6-4	No Significant Hit				
CS8-1	No Significant Hit				
TD3-1	Oryza longistaminata receptor kinase-like protein (Xa21) gene, complete cds and family member C, pseudogene	gb U72723.1	52.0	2e-04	35/38 (92%)
TD3-2	Oryza longistaminata receptor kinase-like protein gene, family member A1, complete	gb U72725.1	52.0	2e-04	35/38 (92%)
TD3-3	No Significant Hit				
TD3-12	No Significant Hit				
TD3-16	Hordeum vulgare NBS-LRR type resistance protein (b7) gene, partial cds	gb AF032685.1	87.7	4e-15	101/120 (84%)
	Canine herpesvirus unique short region, protein kinase US2 and serine/threonine protein kinase US3 genes, complete cds	gb U61387.1	40.1	0.79	20/20 (100%)
TM3-1	No Significant Hit				
TM3-2	No Significant Hit				
TM3-10	No Significant Hit				

Every generated sequence was analysed using blast algorithm for comparison with the genes and DNA sequences available in the GenBank databases. The homologous sequences were obtained by analysing the NCBI-data. According to most probable relevancies we grouped our sequences into 8 groups and the alignment and genetic similarities were determined for every group, too. The alignments were presented below in Boxshade format, in the gap generated for performed by using ClustalX (1.81) computer program.

1. Group 1 (RGA1)

```

RGA1      1  -----
A6-4      1  -----
P2-AB     1  -----
A2-3      1  -----
CS2-3     1  -----
PB-4      1  -----
CS2-6     1  -----
P2-3      1  -----

RGA1      48  -----
A6-4      48  -----
P2-AB     48  -----
A2-3      48  -----
CS2-3     48  -----
PB-4      110 -----
CS2-6     47  -----
P2-3      46  -----

RGA1      157 -----
A6-4      157 -----
P2-AB     157 -----
A2-3      157 -----
CS2-3     157 -----
PB-4      219 -----
CS2-6     156 -----
P2-3      155 -----

RGA1      266 -----
A6-4      266 -----
P2-AB     266 -----
A2-3      265 -----
CS2-3     265 -----
PB-4      327 -----
CS2-6     256 -----
P2-3      254 -----

RGA1      376 -----
A6-4      376 -----
P2-AB     375 -----
A2-3      374 -----
CS2-3     375 -----
PB-4      437 -----
CS2-6     -----
P2-3      364 -----

RGA1      486 -----
A6-4      486 -----
P2-AB     485 -----
A2-3      484 -----
CS2-3     -----
PB-4      -----
CS2-6     -----
P2-3      472 -----

```

2. Group 2 (RGA2)

```

A2-2      1  -----
P2-4      1  -----
CS2-1     1  -----
RGA2      1  -----
CS2-5     1  -----
A2-4      1  -----

A2-2      109 -----
P2-4      109 -----
CS2-1     110 -----
RGA2      110 -----
CS2-5     105 -----
A2-4      95 -----

A2-2      218 -----
P2-4      218 -----
CS2-1     219 -----
RGA2      219 -----
CS2-5     195 -----
A2-4      195 -----

```

A2-2	-----
P2-4	-----
CS2-1	-----
RGA2	533 CTTC-
CS2-5	513 CCTCA
A2-4	-----

P1-4	205	ACGCTGCTTGAGGAG	GAGCTTGCCTTGGCT	GGGATATG	ACATCATCCAGTGGGAAATTTCTCTCTACATCATGACGAGAGAGTTCACAAATCTTCTCTAG
P1-6	205	ACGCTGCTTGAGGAG	GAGCTTGCCTTGGCT	GGGATATG	ACATCATCCAGTGGGAAATTTCTCTCTACATCATGACGAGAGAGTTCACAAATCTTCTCTAG
Al-1	203	ACGCTGCTTGAGGAG	GAGCTTGCCTTGGCT	GGGATATG	ACATCATCCAGTGGGAAATTTCTCTCTACATCATGACGAGAGAGTTCACAAATCTTCTCTAG
P1-5	205	ACGCTGCTTGAGGAG	GAGCTTGCCTTGGCT	GGGATATG	ACATCATCCAGTGGGAAATTTCTCTCTACATCATGACGAGAGAGTTCACAAATCTTCTCTAG
Al-3	197	ACGCTGCTTGAGGAG	GAGCTTGCCTTGGCT	GGGATATG	ACATCATCCAGTGGGAAATTTCTCTCTACATCATGACGAGAGAGTTCACAAATCTTCTCTAG
P1-7	205	ACGCTGCTTGAGGAG	GAGCTTGCCTTGGCT	GGGATATG	ACATCATCCAGTGGGAAATTTCTCTCTACATCATGACGAGAGAGTTCACAAATCTTCTCTAG
CS-2	205	ACGCTGCTTGAGGAG	GAGCTTGCCTTGGCT	GGGATATG	ACATCATCCAGTGGGAAATTTCTCTCTACATCATGACGAGAGAGTTCACAAATCTTCTCTAG
P1-1	221	ACGCTGCTTGAGGAG	GAGCTTGCCTTGGCT	GGGATATG	ACATCATCCAGTGGGAAATTTCTCTCTACATCATGACGAGAGAGTTCACAAATCTTCTCTAG
P5-2	148	ACGCTGCTTGAGGAG	GAGCTTGCCTTGGCT	GGGATATG	ACATCATCCAGTGGGAAATTTCTCTCTACATCATGACGAGAGAGTTCACAAATCTTCTCTAG

P1-4
 P1-6 521 **ATATATATTC** **GAAGAGT** **GAA**
 A1-1 519 **ATATATATTC** **GAAGAGT**
 P1-5 521 **ATATATATTC** **GAAGAGT** **GAATACGAC**
 A1-5 512 **ATATATATTC** **GAA**
 P1-6 521 **ATATATATTC** **GAAGAGT** **GAATACGACGATGCGCTCAGGGAA**
 C95-2 512 **ATATATATTC** **GAAGAGT** **GAATACGACGATGCGCTCAGGGAGGATGTGCCAAAAAGAGATGCATGTATGT**
 P1-1 538 **ATATATATTC**
 P5-2

```

513 AGG GGGTCCCGTCTGGTCTG
514 AGG GGGTCCCGTCTGGTCTG
514 AGG GGGTCCCGTCTGGTCTG
514 AAG GGGTCCCGTCTGGTCTG
515 GAT GGGTCCCGTCTGGTCTG
514 CTT GGGTCCCGTCTGGTCTG
149 TCT GATGATGATGATGATG
408 CAC GATGATGATGATGATG
246 CTC GATGATGATGATGATG
479 CTC GATGATGATGATGATG
307 TCT GATGATGATGATGATG
497 TCT GATGATGATGATGATG
446 CTG GATGATGATGATGATG

```

6. Group 4a

A2-1	1	GGGGGGGTGGGAGACGACGTTTCCACCCAGACCTGTATATATGATGGAGACTGAGAGCACTTCTGCTTATCTTTGGCTGTGTATGACATCTTATGATGGT
A6-2	1	GGGGGGGGTGGGAGACGACGTTTCCACCCAGACCTGTATATGATGGAGACTGAGAGCACTTCTGCTGTATGCTGTGTATGACATCTTATGATGGT
P6-5	1	GGGGGGGTGGGAGACGACGTTTCCACCCAGACCTGTATATGATGGAGACTGAGAGCACTTCTGCTGTATGCTGTGTATGACATCTTATGATGGT
CS2-4	1	GGGGGGGTGGGAGACGACGTTTCCACCCAGACCTGTATATGATGGAGACTGAGAGCACTTCTGCTGTATGCTGTGTATGACATCTTATGATGGT
P6-4	1	GGGGGGGTGGGAGACGACGTTTCCACCCAGACCTGTATATGATGGAGACTGAGAGCACTTCTGCTGTATGCTGTGTATGACATCTTATGATGGT
A6-1	1	GGGGGGGTGGGAGACGACGTTTCCACCCAGACCTGTATATGATGGAGACTGAGAGCACTTCTGCTGTATGCTGTGTATGACATCTTATGATGGT
CS6-1	1	GGGGGGGTGGGAGACGACGTTTCCACCCAGACCTGTATATGATGGAGACTGAGAGCACTTCTGCTGTATGCTGTGTATGACATCTTATGATGGT

[illegible]

A2-1	221	CCGAAAGCTCTTAATGCTCTTGGATGATATATGGAAATATGCGAGATATGCGAAACCTCTGATCTGCTTATCATMAAGGGAGGCCCAAGGCCACATGATATAT
A6-2	221	CCGAAAGCTCTTAATGCTCTTGGATGATATATGGAAATATGCGAGATATGCGAAACCTCTGATCTGCTTATCATMAAGGGAGGCCCAAGGCCACATGATATAT
P6-5	221	CCGAAAGCTCTTAATGCTCTTGGATGATATATGGAAATATGCGAGATATGCGAAACCTCTGATCTGCTTATCATMAAGGGAGGCCCAAGGCCACATGATATAT
C32-4	221	CCGAAAGCTCTTAATGCTCTTGGATGATATATGGAAATATGCGAGATATGCGAAACCTCTGATCTGCTTATCATMAAGGGAGGCCCAAGGCCACATGATATAT
P6-4	221	CCGAAAGCTCTTAATGCTCTTGGATGATATATGGAAATATGCGAGATATGCGAAACCTCTGATCTGCTTATCATMAAGGGAGGCCCAAGGCCACATGATATAT
A6-1	221	CCGAAAGCTCTTAATGCTCTTGGATGATATATGGAAATATGCGAGATATGCGAAACCTCTGATCTGCTTATCATMAAGGGAGGCCCAAGGCCACATGATATAT
C36-1	221	CCGAAAGCTCTTAATGCTCTTGGATGATATATGGAAATATGCGAGATATGCGAAACCTCTGATCTGCTTATCATMAAGGGAGGCCCAAGGCCACATGATATAT

A2-1	331	TTTGTGCACTGCTGGATTCGGAAATCTTGCGAAGATGATGAAGAAACATGTTCAATGAGACACACGACAGCTTGGAGCTCTATACCTCTCTTCCAGATCTTTGACAGCGATGTA
A6-2	331	TTTGTGCACTGCTGGATTCGGAAATCTTGCGAAGATGATGAAGAAACATGTTCAATGAGACACACGACAGCTTGGAGCTCTATACCTCTCTTCCAGATCTTTGACAGCGATGTA
P6-5	331	TTTGTGCACTGCTGGATTCGGAAATCTTGCGAAGATGATGAAGAAACATGTTCAATGAGACACACGACAGCTTGGAGCTCTATACCTCTCTTCCAGATCTTTGACAGCGATGTA
P6-6	331	TTTGTGCACTGCTGGATTCGGAAATCTTGCGAAGATGATGAAGAAACATGTTCAATGAGACACACGACAGCTTGGAGCTCTATACCTCTCTTCCAGATCTTTGACAGCGATGTA
CS2-4	331	TTTGTGCACTGCTGGATTCGGAAATCTTGCGAAGATGATGAAGAAACATGTTCAATGAGACACACGACAGCTTGGAGCTCTATACCTCTCTTCCAGATCTTTGACAGCGATGTA
P6-4	331	TTTGTGCACTGCTGGATTCGGAAATCTTGCGAAGATGATGAAGAAACATGTTCAATGAGACACACGACAGCTTGGAGCTCTATACCTCTCTTCCAGATCTTTGACAGCGATGTA
A6-1	331	TTTGTGCACTGCTGGATTCGGAAATCTTGCGAAGATGATGAAGAAACATGTTCAATGAGACACACGACAGCTTGGAGCTCTATACCTCTCTTCCAGATCTTTGACAGCGATGTA
CS6-1	331	TTTGTGCACTGCTGGATTCGGAAATCTTGCGAAGATGATGAAGAAACATGTTCAATGAGACACACGACAGCTTGGAGCTCTATACCTCTCTTCCAGATCTTTGACAGCGATGTA

A2-1 441 ATTTCGACACATACGCCCGGATATATGAGATGACTGACTGATGATATGACAGCAAAATTCG
 A6-2 441 ATTTCGTGACATACGCCCGGGATATATGAGATGACTGACTGCTGATTCGACGCAATATGAGGCTTCGCC
 P6-5 441 ATTTCGTGACACATACGCCCGGGATATATGAGATGACTGACTGCTGATTCGACGCAATATGAGGCTTCGCC
 CS2-4 441 ATTTCGTGACACATACGCCCGGGATATATG
 P6-4 441 ATTTCGTGACACATACGCCCGGGATATATGATATTCGATATGACAGCAATATGAGGCTTCGCC
 A6-1 441 ATTTCGTGACACATACGCCCGGGATATATGAGATGACTGACTGCTGATTCGACGCAAAATTCG
 CS6-1 441 ATTTCGTGACACATACGCCCGGGATATATGAGATGACTGACTGCTGATTCGACAGCAATATGAGGCTTCGCC

7. Group 5 (RGA b6)

A3-1	1	GGGGGGTGGGGAGCGGCCCTTGT	CAAGATG	TATGATCGGAAAGTGCAT	TTCCTGCTCCACM	TCTTGGTGTGTTG	TCAAAGATATATGTC
P3-4	1	GGGGGGTGGGGAGCGGCCCTTGT	CAAGATG	TATGATCGGAGAAATGCAT	TTCCTGCTCCACM	TCTTGGTGTGTTG	TCAAAGATATATGTC
RGAb6-barley	1	GGGGGGTGGGGAGCGGCCCTTGT	CAAGATG	TATGATCGGAGAGAAAT	TTCCTGCTCCACM	TCTTGGTGTGTTG	TCAAAGATATATGTC
A4-2	1	GGGGGGTGGGGAGCGGCCCTTGT	CAAGATG	TATGATCGGAGAGAAAT	TTCCTGCTCCACM	TCTTGGTGTGTTG	TCAAAGATATATGTC
TD3-12	1	GGGGGGTGGGGAGCGGCCCTTGT	CAAGATG	TATGATCGGAGAGAAAT	TTCCTGCTCCACM	TCTTGGTGTGTTG	TCAAAGATATATGTC

[illegible]

A3-1	211	ATGACCAATATGACAGAGTGG	ATGAGGAGGAGAGCTGTTGAT	TGCTGT	TGATATGTGTCATG	TAGAT	CATATATGCAATATGCAATG	TCCACAGGTATATCC
P3-4	211	ATGACCAATATGACAGAGTGG	ATGAGGAGGAGAGCTGTTGAT	TGCTGT	ATGATGTGTGCTG	TAGAT	CATATATGCAATATGCAATG	TCCACAGGTATATCC
RGAb6-barley	211	ATGACCAATATGACAGAGTGG	ATGAGGAGGAGAGCTGTTGAT	TGCTGT	TGATATGTGTCATG	TAGAT	CATATATGCAATATGCAATG	TCCACAGGTATATCC
A4-2	214	ATGACCAATATGACAGAGTGG	ATGAGGAGGAGAGCTGTTGAT	TGCTGT	TGATATGTGTCATG	TAGAT	CATATATGCAATATGCAATG	TCCACAGGTATATCC
TD3-12	213	ATGACCAATATGACAGAGTGG	ATGAGGAGGAGAGCTGTTGAT	TGCTGT	TGATATGTGTCATG	TAGAT	CATATATGCAATATGCAATG	TCCACAGGTATATCC

[illegible][illegible]

A3-1	529	CCCTCTAGCCGA-----
P3-4	529	CCCTCT-----
RGAb6-barley	529	TATGG-----
A4-2	523	CCCTCTAGCCGAATTCCA
TD3-12	508	GA-----

[illegible]

The other important group, with 4 uninterrupted sequences (A2-2, P2-4, CS2-1, A2-4), is the Group 2 aligned together with the RGA2 GenBank sequence. A2-4 clone's sequence has 1-6 nucleotide deletions at 11 sites and insertions of 1-5 nucleotides at 9 different positions. While the remaining uninterrupted sequences have aligned more precisely with the RGA2, the A2-4 still conserves its polypeptide translation status. This group is specific to S1/AS2 degenerate primer set derived sequences and all the cultivars used, have generated at least one perfectly aligned uninterrupted sequence.

80

Group 4 and 4a are from the important classes. In Group 4 a total of 7 uninterrupted sequences were clustered, among them being the 2 *T. monococcum* sequences (TM3-2 and TM3-10). Though the Group 4 is a non-homogeneous group, it is worth to work with these sequences, even due to the presence of the wild type sequences. Group 4a on the other hand, was generated as a subgroup of Group 4 containing only the homologous S1/AS2 and S1/AS6 derived sequences. The A2-1, A6-2, A6-1 and P6-5 sequences were also present in Group 4 but taking them as a different class an almost 100% homologous group was obtained. The total number of uninterrupted sequences in Group 4a are 5 out of 7 sequences clustered.

Two sequences without any stop codon (A3-1 and A4-2), were clustered in Group 5 together with the barley RGA6 gene sequence. While *T. dicoccoides* (a promising wild type) derived sequence was clustered in this group, unfortunately its sequence was interrupted, however, since this sequence was clustered with A2-4, this 'Avocet' derived sequence may be utilised in our further analysis and experimentation.

The last group had some homologies with kinases genes and considering the only monocot resistance gene so far sequences is a receptor like protein kinase (*Xa21*) from rice the class was found interesting. Knowing that *T. dicoccoides* possesses disease resistance genes, the alignment of the only uninterrupted sequence of the group, TD3-16, together with the kinase related sequences was considered meaningful and worth of further analysis.

3.3.4 Phylogenetic Analysis of Wheat RGAs

The large amount of wheat RGA sequences generated in this thesis and grouped according to Section 3.3.3, enable us to construct phylogeny and compare the RGA groups (within group), based on nucleotide sequences aligned from ClustalX (1.81). The Phylogenetic trees in Figure 3.12 were obtained by the neighbor-joining algorithm implemented in ClustalX software. The reliability of the trees were then established by conducting 1000 bootstrap resampling steps. The GenBank derived homolog sequences were also incorporated in the analysis whenever available.

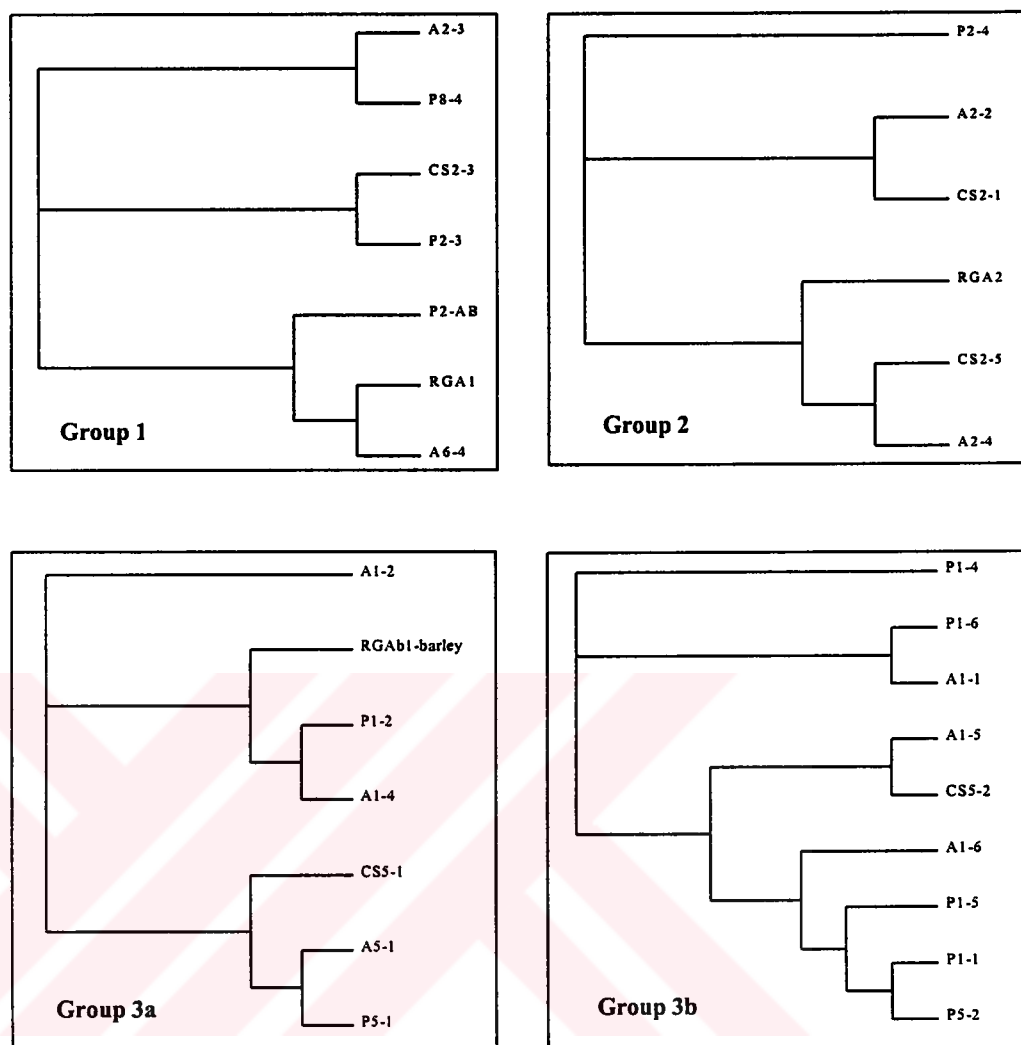


Figure 3.12 Genetic distances of the homologs of Groups: 1 (RGA1), 2 (RGA2), 3a (RGA b1), 3b (RGA b1), 4, 4a, 5 (RGA b6), Group kinase. Dendograms were generated based on neighbor-joining method of ClustalX 1.81. Dendogram was obtained from TreeView.

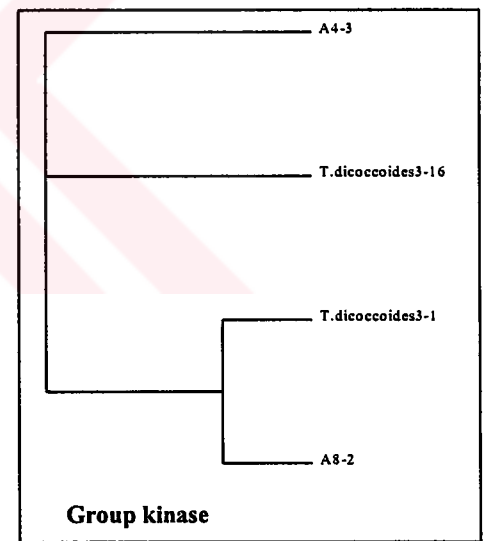
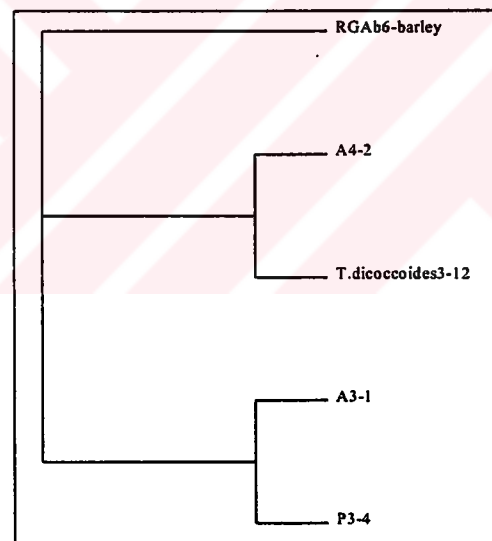
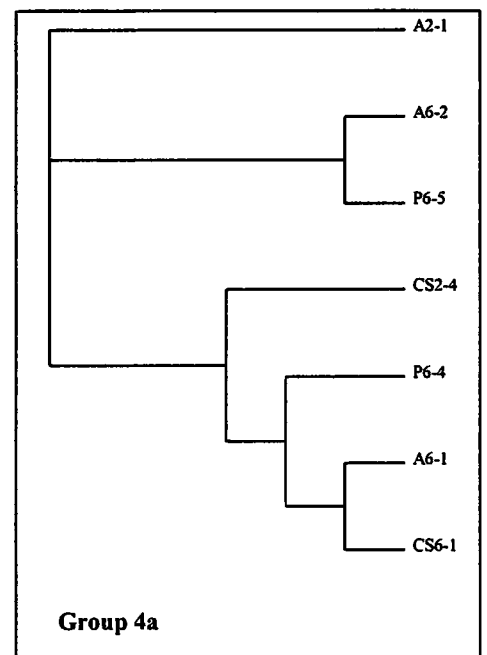
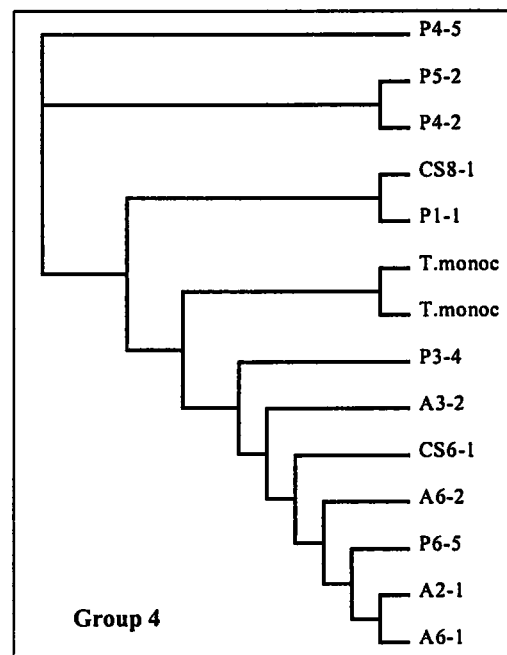


Figure 3.12 (Continued)

3.3.5 Alignment and Phylogenetic Analysis of Uninterrupted RGA Sequences Based on Amino Acid and Nucleotide Sequences

Knowing that not all the sequences are uninterrupted but some of them are probably pseudogenes we analyzed separately the ones that are protein coding without any stop codon. We are not unaware that this does not necessarily mean the upstream or downstream of our sequences do not contain any stop codon since R genes are evolving in a biased way with higher frequency than the remaining part of the genome. The Clustal amino acid sequence alignment of the RGA without any stop codon are presented in Figure 3.13. The genetic relationships of the sequences were presented in Figure 3.14.



Figure 3.13 Alignment of amino acid sequences of the RGAs without any stop codon.
Alignment generated with ClustalX (1.81) software

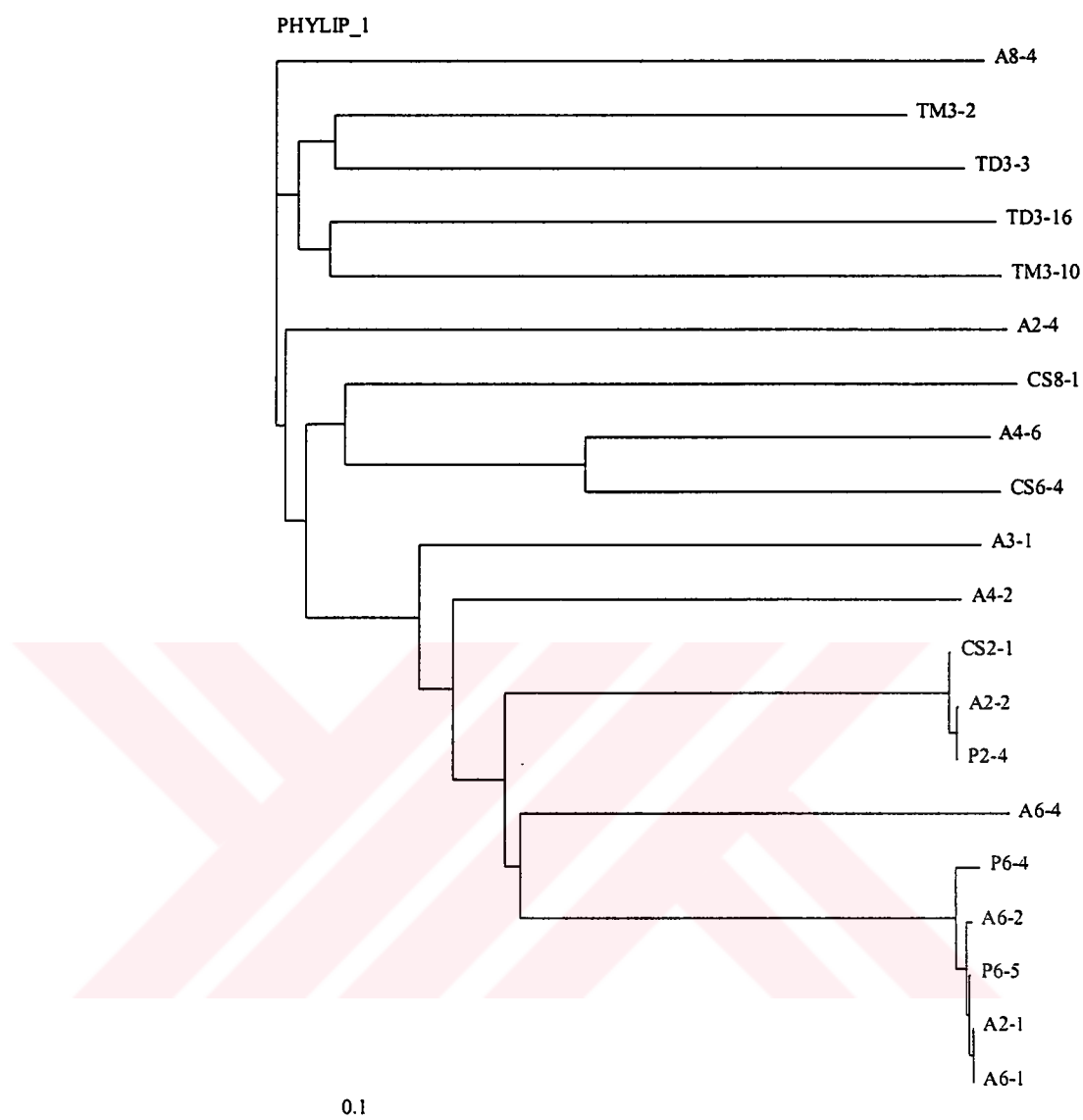


Figure 3.14 Genetic relation of the RGA amino acid sequences that are uninterrupted. Tree generated by neighbor-joining method of ClustalX 1.81 with default (1000) bootstrap values. Dendrogram was obtained using TreeView computer program.

For optimized alignment, the nucleotides coding for amino acid sequences without any stop codons were aligned in Boxshade program after the gaps are introduced in ClustalX 1.81 software and were presented below. The consensus sequence was also presented at the last line of each alignment column.

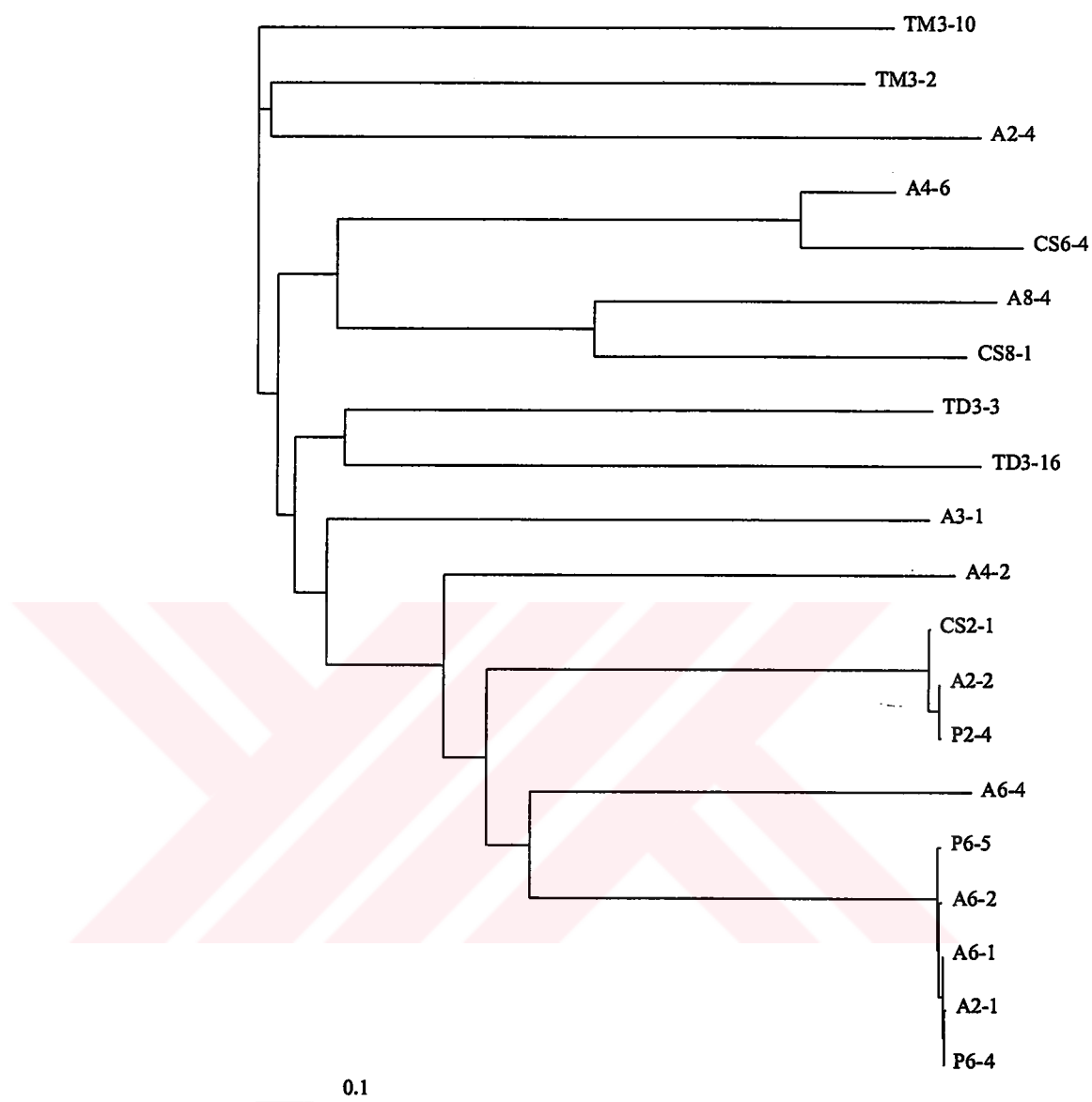


Figure 3.15 Genetic relationship of the RGA nucleotide sequences that do not have any stop codons. Tree generated by neighbor-joining method of ClustalX 1.81 with default (1000) bootstrap values. Dendrogram was obtained from TreeView.

[illegible][illegible][illegible][illegible][illegible]

```

A2-2      -----
P2-4      -----
CS2-1     -----
A6-4      -----
A4-2      520  TTGCGCCCTCAAGCGGAATTCCA
A2-1      -----
P6-4      525  -----
A6-1      526  CTCACCTTCA-----
P6-5      526  CTCGCGCTTCA-----
A6-2      526  CTCGCGCTTCA-----
A3-1      526  TTGCGCCCTCAAGCGGA-----
consensus 551  -----

```

88

CLUSTAL X (1.81) multiple sequence alignment (+ genes)	
A2-1	-----DAKGMILVTRFFPKLAK90CT-----VDSVLEQLGLENDFTTFFACIFGEHK-
P6-4	-----DAKGMILVTRFFPKLAK90CT-----VDSVLEQLGLENDFTTFFACIFGEHK-
A6-1	-----DAKGMILVTRFFPKLAK90CT-----VDSVLEQLGLENDFTTFFACIFGEHK-
P6-5	-----DAKGMILVTRFFPKLAK90CT-----VDSVLEQLGLENDFTTFFACIFGEHK-
A6-2	-----DAKGMILVTRFFPKLAK90CT-----VDSVLEQLGLENDFTTFFACIFGEHK-
RG1	-----IVKQZVLATRTKISVAMHGT-----HDEITLSGLGDEKFLPTACAFRND-
RG2	-----IVKQZVLATRTKISVAMHGT-----HDEITLSGLGDEKFLPTACAFRND-
A6-4	-----IVKQZVLATRTKISVAMHGT-----HDEITLSGLGDEKFLPTACAFRND-
Xa1	QVNSQGEATGMILVTRIQSLAKSLGT-----VQSLKLEALKDDINSKLYAFGNDK-
A2-2	-----GGGSLIVVTRNINVGKLMG-----MTPTYLQGLSNDGSLFRYAFVQGN-
P2-4	-----GGGSLIVVTRNINVGKLMG-----MTPTYLQGLSNDGSLFRYAFVQGN-
C52-1	-----GGGSLIVVTRNINVGKLMG-----MTPTYLQGLSNDGSLFRYAFVQGN-
Y92-1	-----GGGSLIVVTRNINVGKLMG-----MTPTYLQGLSNDGSLFRYAFVQGN-
A4-2	-----VGSFVILVTRNINVGKLMG-----MTPTYLQGLSNDGSLFRYAFVQGN-
C103	-----GGGSLIVVTRNINVGKLMG-----MTPTYLQGLSNDGSLFRYAFVQGN-
I2C-2	-----DLSKILVTRNINVGKLMG-----MTPTYLQGLSNDGSLFRYAFVQGN-
RF8	-----LQ-SNFIITVTRNINVGKLMG-----MTPTYLQGLSNDGSLFRYAFVQGN-
A3-1	-----HGSNMITTRNINVGKLMG-----MTPTYLQGLSNDGSLFRYAFVQGN-
b6	-----LQ-SNFIITVTRNINVGKLMG-----MTPTYLQGLSNDGSLFRYAFVQGN-
RF1	-----HGSNMITTRNINVGKLMG-----MTPTYLQGLSNDGSLFRYAFVQGN-
L4	-----LQ-SNFIITVTRNINVGKLMG-----MTPTYLQGLSNDGSLFRYAFVQGN-
N	-----HGSNMITTRNINVGKLMG-----MTPTYLQGLSNDGSLFRYAFVQGN-
b1	-----HGSNMITTRNINVGKLMG-----MTPTYLQGLSNDGSLFRYAFVQGN-
RF2	-----HGSNMITTRNINVGKLMG-----MTPTYLQGLSNDGSLFRYAFVQGN-
A4-6	-----LQ-SNFIITVTRNINVGKLMG-----MTPTYLQGLSNDGSLFRYAFVQGN-
C56-4	-----LQ-SNFIITVTRNINVGKLMG-----MTPTYLQGLSNDGSLFRYAFVQGN-
P2f	-----HGSNMITTRNINVGKLMG-----MTPTYLQGLSNDGSLFRYAFVQGN-
TD3-2	-----HGSNMITTRNINVGKLMG-----MTPTYLQGLSNDGSLFRYAFVQGN-
TD3-16	-----HGSNMITTRNINVGKLMG-----MTPTYLQGLSNDGSLFRYAFVQGN-
A8-4	-----HGSNMITTRNINVGKLMG-----MTPTYLQGLSNDGSLFRYAFVQGN-
A2-4	-----HGSNMITTRNINVGKLMG-----MTPTYLQGLSNDGSLFRYAFVQGN-
TD3-10	-----HGSNMITTRNINVGKLMG-----MTPTYLQGLSNDGSLFRYAFVQGN-
RF5	-----HGSNMITTRNINVGKLMG-----MTPTYLQGLSNDGSLFRYAFVQGN-
C58-1	-----HGSNMITTRNINVGKLMG-----MTPTYLQGLSNDGSLFRYAFVQGN-
A2-1	-----P-ROYEDDLACIAREIA-----
P4-4	-----P-ROYEDDLACIAREIA-----
A6-1	-----P-ROYEDDLACIAREIA-----
P4-5	-----P-ROYEDDLACIAREIA-----
RG1	-----P-ROYEDDLACIAREIA-----
RG2	-----P-ROYEDDLACIAREIA-----
A6-4	-----P-ROYEDDLACIAREIA-----
Xa1	-----P-ROYEDDLACIAREIA-----
A2-2	-----P-ROYEDDLACIAREIA-----
P2-4	-----P-ROYEDDLACIAREIA-----
C52-1	-----P-ROYEDDLACIAREIA-----
Y92-1	-----P-ROYEDDLACIAREIA-----
A4-2	-----P-ROYEDDLACIAREIA-----
C103	-----P-ROYEDDLACIAREIA-----
I2C-2	-----P-ROYEDDLACIAREIA-----
RF8	-----P-ROYEDDLACIAREIA-----
A3-1	-----P-ROYEDDLACIAREIA-----
b6	-----P-ROYEDDLACIAREIA-----
RF1	-----P-ROYEDDLACIAREIA-----
L4	-----P-ROYEDDLACIAREIA-----
N	-----P-ROYEDDLACIAREIA-----
b1	-----P-ROYEDDLACIAREIA-----
RF2	-----P-ROYEDDLACIAREIA-----
A4-6	-----P-ROYEDDLACIAREIA-----
C56-4	-----P-ROYEDDLACIAREIA-----
P2f	-----P-ROYEDDLACIAREIA-----
TD3-1	-----P-ROYEDDLACIAREIA-----
TD3-2	-----P-ROYEDDLACIAREIA-----
TD3-16	-----P-ROYEDDLACIAREIA-----
A8-4	-----P-ROYEDDLACIAREIA-----
A2-4	-----P-ROYEDDLACIAREIA-----
TD3-10	-----P-ROYEDDLACIAREIA-----
RF5	-----P-ROYEDDLACIAREIA-----
C58-1	-----P-ROYEDDLACIAREIA-----

Figure 3.16 Amino acid sequence alignment of uninterrupted RGAs combined with available resistance genes in GenBank. Alignment generated with ClustalX (1.81) software.

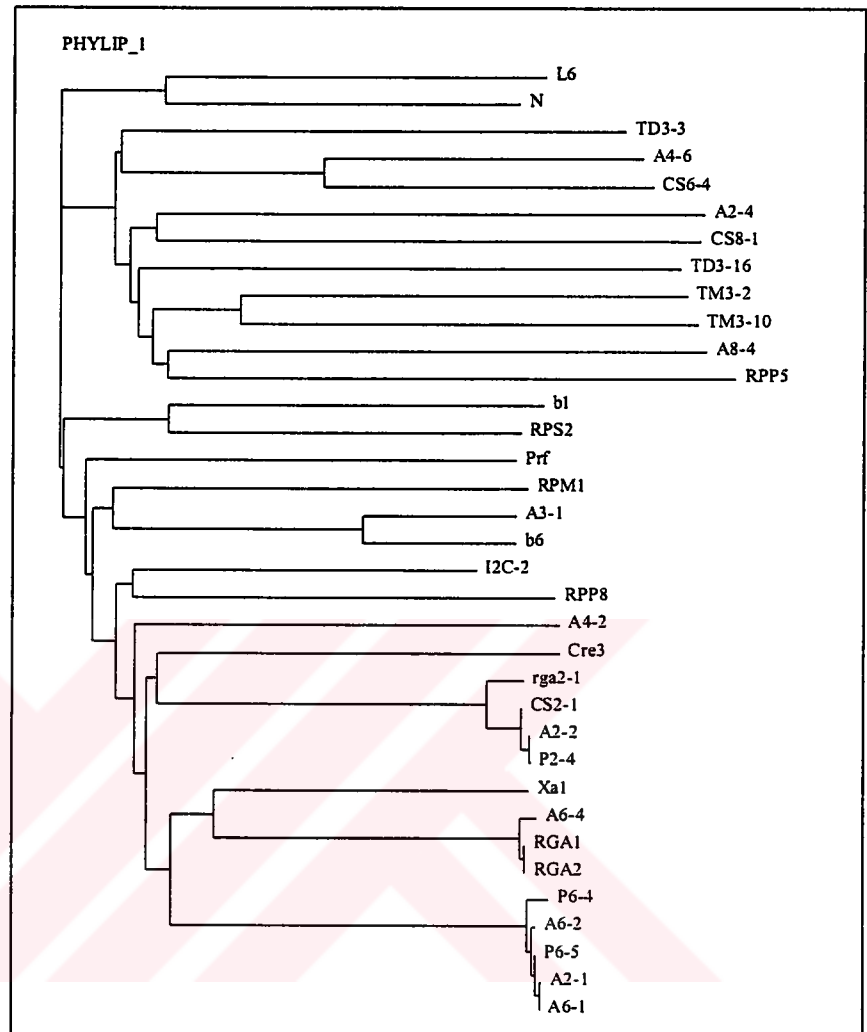


Figure 3.17 Genetic relationship of RGAs in amino acid level without any stop codons combined with available resistance genes from GenBank. Tree generated by neighbor-joining method of ClustalX 1.81 with default (1000) bootstrap values. Dendrogram was obtained from TreeView.

The sequences generated from this thesis are going to make a great contribution to the currently accepted resistance breeding strategy after they will be delivered to the GenBank where 574 resistance genes/gene homologs are available by 19/10/20. The uninterrupted RGA sequences, generated from this thesis, alone will make a 1/30 addition to the available sequences. By using the genetic similarity/difference information of our homologous gene sequences to the known genes, the isolation of the genes related to the rust pathogenesis will be attempted.

On the other hand, while expression of the selected resistance gene analogs will be followed, the development of AFLP and or SSR molecular markers linked to the resistances of the very important wheat rust disease, yellow rust, will be conducted.

As it was observed in the analysis above, molecular geneticists have techniques to classify lines, populations, and landraces; to establish genetic linkages with traits of agronomic and economic interest; and to detect the effect of genetic variation on those traits. Once genes and alleles related to the expression of a trait are identified, the allelic frequencies in a segregating population can be described by a statistical analysis, the genetic variation within and between populations, races, or cultivars can be summarized and compared.

CHAPTER V

CONCLUSIONS

Wheat is the most important crop of Turkey and the entire world. We are among the highest per capita wheat consuming population worldwide. The national mean grain yield is quite low (around 2 t ha⁻¹) even though a considerable yield increase was attained especially during the 1970s (In 1925 yield was 344 kg ha⁻¹) with the advanced breeding tools. In 1999 from the total area of 9 100 000 ha planted, 19 000 000 tons were produced with a mean yield of 2 087 kg ha⁻¹. Thus, Turkey uses 42 % of its total arable land for wheat production which is very high when compared with developed countries (USA uses 22% and Germany 20% of their land) (Braun, 1999). This huge land utilization necessitates continues survey and research as well as controlled production, since the production directly concerns a great proportion of the Turkish population. When there is an epidemic the economic consequences become very important.

Progress in wheat biotechnology and molecular genetics generally lags behind that in maize and barley. Nevertheless, exciting progress has been made in doubled-haploid (DH) production protocols, the development of molecular markers [Simple Sequence Repeats (SSRs) or microsatellites and Amplified Fragment Length Polymorphisms (AFLPs)], and transformation protocols in wheat and barley in recent years. Incorporation of these techniques in the yearly breeding cycle now warrants serious consideration.

The implications of the routine use of DH technologies and molecular marker assisted selection (MAS) on the organization of the wheat breeding projects should now be considered. A combination of DH production and MAS, based on markers for each target disease resistance gene segregating in a population, potentially offer very substantial gains in time and efficiency with a reduced number of backcrossing steps. Generally a minimum of two to three generations can be saved using DH technology while MAS can be used to ensure that only those genotypes carrying the target genes for disease resistance are tested in the field. Both DH

technologies and MAS currently have technical limitations that reduce their cost competitiveness and limit their applicability to a relatively small number of selected crosses. However, these technologies are now being used by commercial companies for both wheat and barley improvement in Europe. Major advances in technology are expected to be more competitive and widely applicable in the next few years.

At the molecular level, genetic diversity can be detected with biochemical and molecular markers. Biochemical markers include seed storage proteins and isozymes, proteins with an enzymatic function (also used in genetic fingerprinting). Isozyme markers are comparatively inexpensive and powerful methods of measuring allele frequencies for specific genes, but because there are few isozyme systems perfected per species (not more than 30), there are relatively few markers available. Molecular markers are enabling the study of a much larger number of the genes that code for plant expression, as well as other non-coding segments of the chromosome. Thus, they are much more informative.

Molecular geneticists have techniques to classify lines, populations, and landraces; to establish genetic linkages with traits of agronomic and economic interest; and to detect the effect of genetic variation on those traits. Once genes and alleles related to traits identified, the allele frequencies in a segregating population can be described by a standard set of summary statistics, and the apportionment of genetic variation within and between populations, races, or cultivars can be summarized and compared by multivariate analysis.

Turkey is considered to be the center of genetic diversity and origin of wheat. The archeological and most currently molecular evidence strongly support that the wheat domestication was in the South-East Turkey, a part of the Fertile Crescent, and great civilizations started from this region. Analysis on the Turkish wheat germplasm with the advanced molecular genetic tools is absent. Not only the extremely rich wild progenitors but also the widely cultivated wheat varieties are not properly characterized. Microsatellites are probably the most powerful molecular markers currently available and are ready to serve on the analysis of the wheat genome. They are very useful for developing detailed linkage maps with obvious utility for the identification of markers linked to genes of agronomic importance. Beyond this practical utility this marker system can also provide detailed information regarding to the structure of the genome and the evolutionary relatedness of species within the family as well as the sequence rearrangements during evolution and the relationship between physical and genetic distances. SSRs are highly polymorphic and informative markers. While we can add the availability of these markers *via* publications of the primer sequences to their advantages their disadvantage is that they require the generation of sequence information for

their application. In order to develop these markers a labor-intensive work is needed, specifically related to the cloning, screening and sequencing of the positive clones. Not only the utilization of expensive radio-nucleotides but handling with thousands of clones is time consuming process. In spite of this disadvantage, numerous eukaryotic genomes are being characterized and the genetic maps are generated with this marker system on almost every important species including human basically due to this marker system's superiorities over the classical morphological and biochemical RFLP and RAPD analysis. In this thesis:

1. We developed a new shortcut strategy for the isolation of SSRs, based on AFLP utilization (which is another important marker system, also applied to the wheat analysis studies in our lab). As it was explained in Chapter I, a novel method for isolation of microsatellite markers were developed efficiently by using wheat as source of DNA but the method is equally applicable to every other eukaryotic genome.

This method, using different AFLP selective primer combinations, would be almost endless source for isolating sufficiently long fragments containing any target repeat units. In the methodology, further improvements can be done by taking some precautions such as, longer electrophoresis of the AFLP gel in order to separate the fragments further apart, also applying higher stringency conditions during enrichment.

We believe this short-cut isolation procedure for finding and sequencing of microsatellite loci is a viable alternative to current practice and may prove to be highly advantageous and productive for most laboratories.

2. We analyzed the doubled haploid Turkish durum wheat varieties 'Çakmak-79', 'Berkmen-469' and 'Kunduru-1149' by utilizing 15 SSR markers, 10 of which were highly polymorphic (two markers amplifying two different loci). Populations of doubled-haploid lines of landrace selected cultivars 'Çakmak-79' and 'Kunduru-1149', were the two most distant populations with $(\delta\mu)^2=1.42$. 'Berkmen-469' x 'Çakmak-79' and 'Berkmen-469' x 'Kunduru-1149' yielded similar genetic distances, $(\delta\mu)^2$ of 0.84 and 0.85, respectively.

Populations of doubled-haploid lines of landrace selected cultivars 'Çakmak-79' and 'Kunduru-1149', were the two most distant populations with $(\delta\mu)^2=1.42$. In the pairwise comparisons 'Berkmen-469' x 'Çakmak-79' and 'Berkmen-469' x 'Kunduru-1149' yielded similar genetic distances, $(\delta\mu)^2$ of 0.84 and 0.85, respectively.

In the analysis, some data generated in this thesis (data from seven out of fifteen markers used) study was combined with the previous work of a fellow graduate student (Dograr, 1999) in the lab and the genetic relation of the DH lines with the other durum cultivars, including the progenitors of the DH samples were determined.

This first molecular level analysis of Turkish durum wheat DH lines and the genetic relationship data within and between the 3 cultivars 'Berkmen-469', 'Çakmak-79' and 'Kundur-1149' in combination with their derivatives (DH lines) were obtained in our lab. 'Çakmak-79' variety was shown to be relatively pure while 'Berkmen-469' and 'Kundur-1149' were genetically highly mixed.

3. Being in a region where wheat was domesticated and evolved, most wheat diseases and pests occur in Turkey. The most important production limiting disease is yellow rust with major recent epidemics occurring in six of the past 25 years and causing estimated losses exceeding 1 million tons in 1991 and 1997. The epidemic comes especially in the suitable climatic conditions with more rainfall and thus, much more production and welfare increase to the farmer is expected. Even if the losses due to the diseases are compensated with the yield increase the mean production yield does not increase.

In this thesis, wheat varieties 'Pastor' and 'Avocet' with genes conferring resistance to yellow rust were utilized together with susceptible variety 'Chinese spring' and the potential resistance harboring wild types *T. monococcum* and *T. dicoccoides*, to isolate and genetically analyze resistance gene analogs. From the 8 different primer sets utilized 7 of them amplified DNA fragments suitable for cloning and about 30 positive clones were selected for further analysis. Representative clones derived from every primer set were selected for sequencing. A total of 84 wheat resistance gene homologues were sequenced. The sequences were aligned and genetic relationships were determined within the sequences, and our sequences combined with the already sequenced resistance genes available in GenBank by using suitable genetic analysis software. We also determined the relation of the sequences according to the homologies that we observed with the GenBank sequences. About 25% of these uninterrupted sequences were found to encode amino acids most of them related to the NBS-LRR regions which was originally targeted by using degenerate PCR primers specifically designed to amplify this most abundant resistance motive. Our results representing the monocot species, differ from the studies relating to the dicot species where up to 75% uninterrupted sequences were observed. While the

complicated (high level of repetitive sequences) hexaploid wheat genome should contribute to this result, the fact that the currently available gene sequences are from few model plant species like arabidopsis and tomato, hence wheat is not genetically related to these plants, is another factor. One more thing to consider, should be the fact that resistance gene clusters appear to evolve more rapidly than other regions of the genome through sequence divergence and numerous nonfunctional genes or pseudogenes do arise.

The information obtained from the wheat RGA sequences will probably be utilized to clone and isolation of entire genes, by hybridization and or different PCR strategies, related to disease resistances. The mRNA expression analysis of the selected RGA sequences in the resistant and susceptible genotypes will provide valuable information on the function of these genes and/or gene homologues. The publication of these sequences will contribute to the future evolutionary studies of the RGA sequences especially due to the limited data available on monocot species in the literature.

REFERENCES

- Aggarwal, R. K., Brar, D. S., Nandi, S., Huang, N. and Khush, G. S. (1999) "Phylogenetic relationships among *Oryza* species revealed by AFLP markers" *Theoretical and Applied Genetics* **98**: 1320-1328
- Akkaya, M.S., Bhagwat, A.A. and Cregan, P.B. (1992) "Length polymorphism of simple sequence repeat DNA in soybean" *Genetics* **132**:1131-1139
- Akkaya, M.S., Shoemaker, R. C., Specht, J. E., Bhagwat, A.A. and Cregan, P.B. (1995) "Integration of simple sequence repeat DNA markers into a soybean linkage map" *Crop Science* **35**: 1439-1445
- Altschul, S.F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990) "Basic Local Alignment Search Tool" *Journal of Molecular Biology* **215**(3): 403-410
- Anderson, K., Dimaranan, B., Hertel, T. and Martin, W. (1997) "Asia-Pacific food markets and trade in 2005: a global, economy-wide perspective" *The Australian Journal of Agricultural and Resource Economics* **41**(1): 19-44
- Baker, B., Zambryski, P., Staskawicz, B. and Dinesh-Kumar, S. P. (1997) "Signaling in plant-microbe interactions" *Science* **276**: 726-733
- Baxevanis, F. F. and Ouellette, A. D. B. (1998) "Bioinformatics - A practical guide to the analysis of genes and proteins" New York, John Wiley & Sons, Inc.
- Becher, S. A., Steinmetz, K., Weising, K., Boury, S., Peltier, D., Renou, J. P., Kahl, G. and Wolff, K. (2000) "Microsatellites for cultivar identification in *Pelargonium*" *Theoretical and Applied Genetics* **101**:643-651
- Bell, C.J. and Ecker, J.R. (1994) "Assignment of 30 microsatellite loci to the linkage map of *Arabidopsis*" *Genomics* **19**:137-144
- Bender, W. H. (1997) "How much food will we need in the 21st century" *Environment* **39**(2): 7-11 and 27-28

Bennetzen, J. L. and Freeling, M. (1997) "The unified grass genome: synergy in syntheny" *Genome Research* 7: 301-306

Bevan, M. and Murphy, G. (1999) "The small, the large and the wild - the value of comparison in plant genomics" *Trends in Genetics* 15(6): 211-214

Bisby, F. A. (2000) "The Quiet Revolution: Biodiversity Informatics and the Internet" *Science* 289: 2309-2312

Bishop, A. J. R., Louis, E. J. and Borts, R. H. (2000) "Minisatellite variants generated in yeast meiosis involve DNA removal during gene conversion" *Genetics* 156: 7-20

Bohn, M., Utz, H. F. and Melchinger, A. E. (1999) "Genetic similarities among winter wheat cultivars determined on the basis of RFLPs, AFLPs, and SSRs and their use for predicting progeny variance" *Crop Science* 39: 228-237

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

Bredemeijer, G. M. M., Arens, P., Wouters, D., Visser, D. and Vosman, B. (1998) "The use of semi-automated fluorescent microsatellite analysis for tomato cultivar identification" *Theoretical and Applied Genetics* 97: 584-590

Briggs, S. P. and Helentjaris, T. (1997) "Plant genomics moves into the limelight" *Genome Research* 7: 856-857

Brunel, D., (1994) "A microsatellite marker in *Helianthus annuus* L." *Plant Molecular Biology* 24:397-400

Bryan, G.J., Collins, A.J., Stephenson, P., Orry, A., Smith, J.B. and Gale, M.D. (1997) "Isolation and characterization of microsatellites from hexaploid bread wheat" *Theoretical and Applied Genetics* 94: 557-563

Buchanan, F.C., Adams, L.J., Littlejohn, R.P., Maddox, J.F. and Crawford, A.M. (1994) "Determination of evolutionary relationships among sheep breeds using microsatellites" *Genomics* 22:397-403

Bureau, T. E. Ronald, P. C. and Wessler, S. R. (1996) "A Computer-Based Systematic Survey Reveals the Predominance of Small Inverted-Repeat Elements in Wild-Type Rice Genes" *Proc. Natl. Acad. Sci. USA* 93(16): 8524-8529

Cervera, M. T., Gusmao, J., Steenackers, M., Gysel, A. V., Montagu, M. V. and Boerjan, W. (1996) "Application of AFLP™-based molecular markers to breeding of *Populus* spp." *Plant Growth Regulation* 20: 47-52

Chen, X., Temnykh, S, Xu, Y., Cho, Y. G. and McCouch, S. R. (1997) "Development of a microsatellite framework map providing genome-wide coverage in rice (*Oryza sativa* L.)" *Theoretical and Applied Genetics* 95: 553-567

Cohen, J. (1997). "Corn genome pops out of the pack" *Science* 276: 1960-1962

Cregan, P.B., Bhagwat, A.A., Akkaya, M.S. and Rongwen, J. (1994) "Microsatellite Fingerprinting and Mapping of Soybean" *Methods in Molecular and Cellular Biology* 5: 49-61

Deng, Z., Huang, S., Ling, P., Chen, C., Yu, C., Weber, C. A., Moore, G. A. and Gmitter Jr., F. G. () "Cloning and characterization of NBS-LRR class resistance-gene candidate sequences in citrus" *Theoretical and Applied Genetics* 101: 814-822

Devos, K. M. Bryan, G.J., Collins, A.J., Stephenson, P. and Gale, M. D. (1995) "Application of two microsatellite sequences in wheat storage proteins as molecular markers" *Theoretical and Applied Genetics* 90: 247-252

Devos, K. M. and Gale, M. D. (1997) "Comparative genetics in the grasses" *Plant Molecular Biology* 35: 3-15

Diamond, J. (1999) "Guns, Germs, and Steel" W.W. Norton & Company, Inc., N.Y. ISBN 0-393-31755-2

Diamond, J. (1997) "Location, location, location: the first farmers" *Science* 278: 1243-1244

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

Dograr, N., Akin-Yalin, S. and Akkaya, M.S. (2000) "Discriminating durum wheat cultivars using highly polymorphic simple sequence repeat DNA markers" *Plant Breeding* 119:360-362

Dograr, N. (1999) "Assessment of genetic relationship between durum wheat genotypes using highly polymorphic microsatellite markers" *Thesis submitted to the Graduate School of Natural and Applied Sciences of the Middle East Technical University*

Edwards, K. J., Barker, J. H. A., Daly, A., Jones, C. and Karp, A. (1996) "Microsatellite libraries enriched for several microsatellite sequences in plants" *BioTechniques* 20: 758-760

Edwards, J. L., M. A. L., Nielsen, E. S. (2000) "Interoperability of Biodiversity Databases: Biodiversity Information on Every Desktop" *Science* **289**: 2312-2314

Feuillet, C. and Keller, B. (1999) "High gene density is conserved at syntenic loci of small and large grass genomes" *Proc. Natl. Acad. Sci. USA* **96**: 8265-8270

Fisher, D., Bachmann, K. (2000) "Onion microsatellites for germplasm analysis and their use in assessing intra- and interspecific relatedness within the subgenus *Rhizirideum*" *Theoretical and Applied Genetics* **101**:153-164

Fisher, D., Bachmann, K. (1998) "Microsatellite enrichment in organisms with large genomes (*Allium cepa* L.)" *BioTechniques*, **5**:796-800

Garcia-Mas, J., Oliver, M., Gomez-Paniagua, H. and Vicente, M. C. (2000) "Comparing AFLP, RAPD and RFLP markers for measuring genetic diversity in melon" *Theoretical and Applied Genetics* **101**: 860-864

Goldstein, D.B., Linares, A. R., Cavalli-Sforza, L.L. and Feldman, M.W. (1995). Genetic absolute dating based on microsatellites and the origin of modern humans. *Proc. Natl. Acad. Sci. USA* **92**: 6623-6727

Goodfellow, P. N. (1993) "Microsatellites and the new genetic maps" *Current Biology* **3**(3): 149-151

Goodman, J.S., (1997) "R-St CALC-A collection of computer-programs for calculating estimates of genetic differentiation from microsatellite data and determining their significance" *Molecular Ecology* **6**: 881-885

Hakki, E. E. and Akkaya, M. S. (2000) "Microsatellite isolation using amplified fragment length polymorphism markers: no cloning, no screening" *Molecular Ecology* **9**(12): 000-000

Hamada, H., Kakunaga, T., Petrino, M. G. (1982) "A novel repeated element with Z-DNA-forming potential is widely found in evolutionarily diverse eukaryotic genomes" *Proc. Natl. Acad. Sci. USA-Biological Sciences* **79**(21): 6465-6469

Hansen, M., Kraft, T., Christiansson, M. and Nilsson N. O. (1999) "Evaluation of AFLP in Beta" *Theoretical and Applied Genetics* **98**: 845-852

Harlan, J. R. (1995) "The living fields - our agricultural heritage" Cambridge, UK, Cambridge University Press

Harlan, J. R. (1975) "Crops and man" Madison, Wisconsin, American Society of Agronomy - Crop Science Society of America

Hayes, A.J., Yue, Y. G. and Saghai Maroof, M. A. (2000) "Expression of two soybean resistance gene candidates shows divergence of paralogous single-copy genes" *Theoretical and Applied Genetics* **101**: 789-795

Heun, M., Schafer-Pregl, R., Klawan, D., Castanga, R., Accerbi, M., Borghi, B. and Salamini, F. (1997) "Site of einkorn wheat domestication identified by DNA fingerprinting" *Science* **278**: 1312-1314

Incirli, A., Akkaya, M.S. (2000) "Assessment of genetic relationships in durum wheat cultivars using AFLP markers" *Accepted for publication*

Innan, H., Terauchi, R. and Miyashita, N. (1997) "Microsatellite polymorphism in natural populations of the wild plant *Arabidopsis thaliana*" *Genetics* **146**:1441-1452

Ito, T., Smith, C. L. and Cantor, C. R. (1992) "Sequence-specific DNA purification by triplex affinity capture" *Proc. Natl. Acad. Sci. USA* **89**: 495-498

Jeanmougin, F., Thompson, J. D., Gouy, M., Higgins, D. G. and Gibson, T. J. (1998) "Multiple Sequence Alignment with Clustal-X" *Trends in Biochemical Sciences* **23**(10): 403-405

Jeffreys, A.J., Wilson, V. and Thein, S.L. (1985) "Hypervariable minisatellite regions in Human DNA" *Nature* **314**: 67-73

Johnson, D. G. (1997) "On the resurgent population and food debate" *The Australian Journal of Agricultural and Resource Economics* **41**(1): 1-17

Jones, N., Ougham, H. and Thomas, H. (1997) "Markers and mapping: we are all geneticists now" *New Phytol.* **137**: 165-177

Karagyozov, L., Kalcheva, I. D. and Chapman, V. M. (1993) "Construction of random small-insert genomic libraries highly enriched for simple sequence repeats" *Nucleic Acids Research* **21**(16): 3911-3912

Kashi, Y., King, D. and Soller, M. (1997) "Simple sequence repeats as a source of quantitative genetic variation" *Trends in Genetics* **13** (2): 74-78

Keim, P., Schupp, J. M., Travis, S. E., Clayton, K., Zhu, T., Shi, L., Ferreira, A. and Webb D. M. (1997) "A high-density soybean genetic map based on AFLP markers" *Crop Science* **37**: 537- 543

Korzun, V., Börner, A., Worland, A.J. and Röder, M.S. (1997) "Application of microsatellite markers to distinguish inter-varietal chromosome substitution lines of wheat [*Triticum aestivum* L.]" *Euphytica* **95**: 149-155

Lagercantz, U., Ellengren, H. and Andersson, L. (1993) "The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates." *Nucleic Acids Research* **21**:1111-1115

Lavi, U., Akkaya, M.S., Bhagwat, A.A., Lahav, E. And Cregan, P. (1995) "Generation of simple sequence repeat DNA markers in avocado (*Persea americana* M.)" *Euphytica* **80**:171-177

Lagudah, E. S., Moullet, O. and Appels, R. (1997) "Map-based cloning of a gene sequence encoding a nucleotide-binding domain and a leucine-rich region at the *Cre3* nematode resistance locus of wheat" *Genome* **40**: 659-665

Leister, D., Kurth, J., Laurie, D. A., Yano, M., Sasaki, T., Devos, K., Graner, A. and Schulze-Lefert, P. (1998) "Rapid reorganization of resistance gene homologues in cereal genomes" *Proc. Natl. Acad. Sci. USA* **95**: 370-375

Leister, D., Ballvora, A., Salamini, F. and Gebhardt, C. (1996) "A PCR-based approach for isolating pathogen resistance genes from potato with potential for wide application in plants" *Nature Genetics* **14 December**: 421-429

Li, Y. C., Fahima, T., Peng, J. H., Roder, M. S., Kirzhner, V. M., Beiles, A., Korol, A. B. and Nevo E. (2000) "Edaphic microsatellite DNA divergence in wild emmer wheat, *Triticum dicoccoides*, at a microsite: Tabigha, Israel" *Theoretical and Applied Genetics* **101**: 1029-1038

Litt, M. and Luty J.A. (1989) "A hypervariable microsatellite revealed by in-vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene" *American Journal of Human Genetics* **44**: 397-401

Liu, Z.W., Biyashev, R.M. and Saghai-Marooof, M.A. (1996) "Development of simple sequence repeat DNA markers and their integration to a barley linkage map" *Theoretical and Applied Genetics* **93**:869-876

Maheswaran, M., Subudhi, P. K., Nandi, S., Xu, J. C., Parco, A., Yang, D. C. and Huang, N. (1997) "Polymorphism, distribution, and segregation of AFLP markers in doubled haploid rice population" *Theoretical and Applied Genetics* **94**: 39-45

Maughan, P. J., Saghai Marooof, M. A., Buss, G. R. and Huestis, G. M. (1996) "Amplified fragment length polymorphism (AFLP) in soybean: species diversity, inheritance, and near-isogenic line analysis" *Theoretical and Applied Genetics* **93**: 392-401

Melchers, L. S. and Stuiwer, M. H. (2000) "Novel genes for disease-resistance breeding" *Current Opinion in Plant Biology* 3: 147-152

Michelmore, R. (2000) "Genomic approaches to plant disease resistance" *Current Opinion in Plant Biology* 3: 125-131

Morgante, M., and Olivieri, (1993) "PCR-amplified microsatellites as markers in plant genetics" *Plant Journal* 3:175-182

Pakniyat, H., Powell, W., Baird, E., Handley, L. L., Robinson, D., Scrimgeour, C. M., Nevo, E., Hackett, C. A., Caligari, P. D. S. and Forster B. P. (1997) "AFLP variation in wild barley (*Hordeum spontaneum* C. Koch) with reference to salt tolerance and associated ecogeography" *Genome* 40: 332-341

Pan, Q., Liu, Y-S., Budai-Hadrian, O., Sela, M., Carmel-Goren, L., Zamir, D., and Fluhr, R. (2000) "Comparative genetics of nucleotide binding site-leucine rich repeat resistance gene homologues in the genomes of two dicotyledons: tomato and arabidopsis" *Genetics* 155: 309-322

Peil, A., Korzun, V., Schubert, V., Schumann, E., Weber, W. E. and Roder, M. S. (1998) "The application of wheat microsatellites to identify disomic *Triticum aestivum*-*Aegilops markgrafii* addition lines" *Theoretical and Applied Genetics* 96: 138-146

Pejic, I., Ajmone-Marsan, P., Morgante, M., Kozumplick, V., Castiglioni, P., Taramino, G. and Motto, M. (1998) "Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs, and AFLPs" *Theoretical and Applied Genetics* 97: 1248-1255

Pestsova, E., Korzun, V., Goncharov, N. P., Hammer, K., Ganai, M. W. and Röder, M. S. (2000) "Microsatellite analysis of *Aegilops tauschii* germplasm" *Theoretical and Applied Genetics* 101: 100-106

Pfeiffer, A., Olivieri, A. M. and Morgante, M. (1997) "Identification and characterization of microsatellites in Norway spruce (*Picea abies* K.)" *Genome* 40:411-419

Plaschke, J., Ganai, M.W. and Röder, M.S. (1995) "Detection of genetic diversity in closely related bread wheat using microsatellite markers" *Theoretical and Applied Genetics* 91: 1001-1007

Powell, W., Morgante, M., Andre, C., Hanafey, M., Vogel, J., Tingey, S. and Rafalski, A. (1996) "The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis" *Molecular Breeding* 2: 225-238

Powell, W., Thomas, W. T. B., Baird, E., Lawrence, P., Booth, A., Harrower, B., McNicol, J. W. and Waugh, R. (1997) "Analysis of quantitative traits in barley by the use of amplified fragment length polymorphisms" *Heredity* 79: 48-59

Prasad, M., Varshney, R. K., Roy, J. K., Balyan, H. S. and Gupta, P. K. (2000) "The use of microsatellites for detecting DNA polymorphism, genotype identification and genetic diversity in wheat" *Theoretical and Applied Genetics* 100:584-592

Provan, J., Powell, W. And Waugh, R. (1996) "Microsatellite analysis of relationships within cultivated potato (*Solanum tuberosum*)" *Theoretical and Applied Genetics* 92:1078-1084

Qi, X. and Lindhout, P. (1997) "Development of AFLP markers in barley" *Molecular and General Genetics* 254: 330-336

Raybould, A. F., Mogg, R. J., Clarke, R. T., Gliddon, C. J. and Gray A. J. (1999) "Variation and population structure at microsatellite and isozyme loci in wild cabbage (*Brassica oleracea* L.) in Dorset (UK)" *Genetic Resources and Crop Evolution* 46: 351-360

Rholf, F.J. (1992) "NTSYS-pc numerical taxonomy and multivariate analysis system" Version 1.70, Exeter Publications, New York

Richter, T. E. and Ronald, P. C. (2000) "The evolution of disease resistance genes" *Plant Molecular Biology* 42: 195-204

Rongwen, J., Akkaya, M.S., Bhagwat, A.A., Lavi, U. and Cregan P.B. (1995) "The use of microsatellite DNA markers for soybean genotype identification" *Theoretical and Applied Genetics* 90:43-48

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

Röder, M.S., Plaschke, J., König, S.U., Börner, A., Sorrells, M.E., Tanksley, S.D. and Ganal, M.W. (1995). "Abundance, variability and chromosomal location of microsatellites in wheat" *Molecular and General Genetics* 246: 327-333

Russell, J., Fuller, J., Young, G., Thomas, B., Taramino, G., Macaulay, M., Waugh, R., and Powell, W. (1997) "Discriminating between barley genotypes using microsatellite markers" *Genome* 40: 442-450

Saghai-Maroofof, M.A., Soliman, K. M., Jorgensen, R. A. (1984) "Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics" *Proc. Natl. Acad. Sci. USA* 81: 8014-8018

Saitou, N. and Nei, M. (1987) "The Neighbor-Joining Method - A New Method for Reconstructing Phylogenetic Trees" *Molecular Biology And Evolution* 14(4): 406-425

Sambrook, J., Fritsch, E.F., Maniatis, T. (1989) "Molecular Cloning A Laboratory Manual" ed. Nolan, C., Second Edition, vols. 1,2 and 3, *Cold Spring Harbor Laboratory Press*, USA

Savaskan, C., Ellerbrook, C., Fish, L.J. and Snape, W. (1997) "Doubled haploid production in Turkish durum wheats using crosses with maize" *Plant Breeding* 116:1-3

Schlotterer, C., Amos, B. And Tautz, D. (1991) "Conservation of polymorphic simple sequence loci in cetacean species" *Nature* 354:63-65

Schwarz, G., Michalek, W., Mohler, V., Wenzel, G. and Jahoor, A. (1999) "Chromosome landing at the *Mla* locus in barley (*Hordeum vulgare* L.) by means of high-resolution mapping with AFLP markers" *Theoretical and Applied Genetics* 98: 521-530

Schut, J. W., Qi, X. and Stam, P. (1997) "Association between relationship measures based on AFLP markers, pedigree data and morphological traits in barley" *Theoretical and Applied Genetics* 95: 1161-1168

Senior, M.L. and Heun, M. (1993) "Mapping maize microsatellites and polymerase chain reaction confirmation of the targeted repeats using a CT primer" *Genome* 36:884-889

Shan, X., Blake, T. K. and Talbert, L. E. (1999) "Conversion of AFLP markers to sequence-specific PCR markers in barley and wheat" *Theoretical and Applied Genetics* 98: 1072-1078

Slatkin, M. (1995). "A measure of population subdivision on microsatellite allele frequencies" *Genetics* 139:457-462

Spielmeyer, W., Huang, L., Bariana, H., Laroche, A., Gills, B. S. and Lagudah, E. S. (2000) "NBS-LRR sequence family is associated with leaf and stripe rust resistance on the end of homoeologous chromosome group 1S of wheat" *Theoretical and Applied Genetics* 101: 1139-1144

Spielmeyer, W., Moullet, O., Laroche, A., and Lagudah, E. S. (2000) "Highly recombinogenic regions at seed storage protein loci on chromosome 1DS of *Aegilops tauschii*, the D-genome donor of wheat" *Genetics* 155: 361-367

Spielmeyer, W., Robertson, M., Collins, N., Leister, D., Schulze-Lefert, P., Seah, S., Moullet, O. and Lagudah, E. S. (1998) "A superfamily of disease resistance gene analogs is located on all homoeologous chromosome groups of wheat (*Triticum aestivum*)" *Genetics* 155: 361-367

Sosinski, B., Gannavarapu, M., Hager, L. D., Beck, L. E., King, G.J., Ryder, C. D., Rajapakse, S., Baird, W. V., Ballard, R. E. and Abbott, A. G. (2000) "Characterization of microsatellite markers in peach (*Prunus persica* (L.) Batsch)" *Theoretical and Applied Genetics* 101:421-428

Stephenson, P., Bryan, G., Kirby, J., Collins, A., Devos, K., Busso, C. and Gale, M. (1998) "Fifty new microsatellite loci for the wheat genetic map" *Theoretical and Applied Genetics* 97: 946-949

Struss, D. and Plieske, J. (1998) "The use of microsatellite markers for detection of genetic diversity in barley populations" *Theoretical And Applied Genetics* 97(1-2): 308-315

Szewc-McFadden, A.K., Kresovich, S., Bliet, S.M., Mitchell, S.E. and McFerson, J.R. (1996) "Identification of polymorphic conserved simple sequence repeats (SSRs) in cultivated *Brassica* species" *Theoretical and Applied Genetics* 93:534-538

Taramino, G. and Tingey, S. (1996) "Simple sequence repeats for germplasm analysis and mapping in maize" *Genome* 38:277-287

Tautz, D. (1989) "Hypervariability of simple sequences as a general source for polymorphic markers" *Nucleic Acids Research* 17:6463-6471

Tautz, D., Trick, M. and Dover, G.A. (1986) "Cryptic simplicity in DNA is a major source of genetic variation" *Nature* 322:652-656

Tautz, D. and Renz, M. (1984) "Simple sequences are ubiquitous repetitive components of eukaryotic genomes" *Nucleic Acids Research* 12 (10): 4127-4138

Thomas, M.R. and Scott, N.S. (1993) "Microsatellite repeats in grapevine reveal DNA polymorphisms when analysed as sequence-tagged-sites (STSs)" *Theoretical and Applied Genetics* 86:985-990

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. and Zabeau, M. (1995) "AFLP: a new technique for DNA fingerprinting." *Nucleic Acids Research* 23:4407-4414

Wang, Z., Weber, J.L., Zhong, G. and Tanksley, S.D. (1994) "Survey of plant short tandem DNA repeats" *Theoretical and Applied Genetics* 88:1-6

Weber, J.L. and May, P.E. (1989) "Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction" *American Journal of Human Genetics* 44:388-396

Welsh, J. and McClelland, M. (1990) "Fingerprinting genome using PCR with arbitrary primers" *Nucleic Acid Research* 18:7213-7218

Williams, J.G.K., Kubelik, A.R., Livak, K.J., Rafalski, J.A. and Tingey, S.V. (1990) "DNA polymorphisms amplified by arbitrary primers are useful as genetic markers" *Nucleic Acids Research* 18:6531-6535

Wu, K. and Tanksley, S. D. (1993) "Abundance, polymorphism and genetic mapping of microsatellites in rice" *Molecular and General Genetics* 241:225-235

Yang, G. P., Saghai-Marooof, M. A., Xu Qifa Zhang, C. G., and Biyashev, R. M. (1994) "Comparative analysis of microsatellite DNA polymorphism in landraces and cultivars of rice" *Molecular and General Genetics* 245: 187-194

Yu, K., Park, S.J. and Poysa, V. (1999) "Abundance and variation of microsatellite DNA sequences in beans (*Phaseolus* and *Vigna*)" *Genome* 42:27-34

Yu, Y. G., Buss, G. R. and Saghai-Marooof, M. A. (1996) "Isolation of a superfamily of candidate disease-resistance genes in soybean based on a conserved nucleotide-binding site" *Proc. Natl. Acad. Sci. USA* 93: 11751-11756

Zabeau, M. and Vos, P. (1993) "Selective restriction fragment amplification: a general method for DNA fingerprinting" *European Patent Application*, publication no. EP 534858-A1, No.92402629.7

Zencirci, N., Aktan, B. and Atli, A. (1994) "Genetic relationships of Turkish durum wheat cultivars" *Turkish Journal of Agricultural and Forestry* 18: 187-192

Zhao, X. and Kochert, G. (1993) "Phylogenetic distribution and genetic mapping of a (GGC)_n microsatellite from rice (*Oryza sativa* L.)" *Plant Molecular Biology* 21:607-614

Zohary, M. and Hopf, D. (1988) "Domestication of plants in the old world - The origin and spread of cultivated plants in West Asia, Europe, and the Nile Valley" New York, *Clarendon Press-Oxford University Press*

APPENDIX A

T. dicoccoides (Acc. No. TR02771) Resistance Gene Analog Sequences

TD3-1

GGGGGGGTGGGGAAGACGGCGAAGAAGAAGACAGGACAATGACAACATATGTGCCCCCTGAATACGGTGATGCTGCAACAGGGGGAGCTGC
TGAAAGATCAAGAGGAACCCGACAGTACGTACGCAGCAAAGAAGGTCGTATGCCCTCTAGGATTGGATGTGCAGAAGATACATGCATGC
CCTAAAGACTACATCCTCTACCGCGGTGCGTACGAGGATTGAACGCATGCCCGGTATGCGGTGCATTGCGGTATAAGATCAGACGAG
ATGACCCCTGGTGATGTTGACGGTGAGCCCCACAGGAAGAGAGTTCATGCGAAGGTGATGTGGTATGCTCCTATAATAGCACGGTTGAA
ATGACTATTCAGAAACAAAGAGCATGCCAAGTTGATGCGATGGAACAGTGAGGATCGTAAGAAAGATGGGAAGTTGAGAGCACCCGCT
GACGGGTCGCAGTGGGA

Observed Homology:

- Disease resistance protein RPM1 homolog" / protein_id=AAD27570.1
- A receptor kinase-like protein encoded by the rice disease resistance gene, Xa21
JOURNAL Science 270 (5243), 1804-1806 (1995) / MEDLINE 96106403

TD3-2

GGGGGGGTGGGGAAGACGGCCGACTGTGGTATTTCTTTAAAGCCCTTTCTGTGATGTCGCCCCGACGTCCGACTGTGGCATTCTTT
TAAAGCCCTTTATGTGGTGTGCGTAAGACGCCCGACTGTGGTATTTCTTTAAAGCCCTTTCTGTGATGTCGCCCCGACGTCCGACTG
TGGCATTCTTTTAAAGCCCTTTATGTGGTGTGCGTAAGACGCCCGACTGTGGTGTCTTTTAAAGCCCTTTCTGTGATGTCGCCCC
GACGTCCGACTGTGGCATTCTTTTAAAGCCCTTTATGTGTTGTCGGTAAGACGCCCGACTGTGGCATCATTTCTTATATTTATTT
TTCACCTTTCGAGGCGTCGCTCAGACTCCCGATCGGTCTTCATTATGTT

TD3-3

GGGGGGGTGGGGAAGACGGCATCTGGTGGATCCGACGACAAAAGTGCGCCAGTGGTGGATTGCGCTGCATTGTTAGCCATGGAAGCG
CGGGAGGGCAGGCAGCTACGCGCAGACCGCATGGGGTTATAGGCGGTAGTGGTGACCGGAGGGATGGAGGGGATAGTGGGA

TD3-12

GGGGGGGTGGGGAAGACGACACTGGCAGATCAAGTCTATCGCAAGATACATGGACACTTTCATTGTGCGGGCTTTCTGTGTCAGTCTCAC
AAAAGCCAGATATAAAGAAAATTGTGAAGGATGTGATCTCTCAAGTATCTTCCAAGATAAATCCAACAAGATATCAGTGAGTGGGA
CGAAAAGAAGTCTATTGCAAACTAAGAGAACTGCTACAAGATAAGAGGTAATCTCAGCTCTGTCTCTTCTGCAAAACATGTGTCA
TGAGTATATTTCTGTAATATTTCTGCATATACTGTCACTAAAGCGGGGATGATGCATGTGCAAGTAACCGCATATGGAACATAATC
AATAAGCATGATATATACTGATACCTCAGTTAACAAGTTGAGCGTATTAATCTCATATACACTTTTGGTGTTCATGGCATGTCTAC
TGATCTACTCTTAATTTCTGACTGTTCAAATGTATATTGTGAATAGTATTTCGACATAAAAACCAA

Observed Homology:

- Hordeum vulgare NBS-LRR type resistance protein (b7) gene, partial cds. / ACCESSION AF032685
- Homo sapiens DMBT1 candidate tumour suppressor gene, exons 1 to 55. / ACCESSION AJ243211

TD3-16

GGGGGGGTGGGGAAGACGGCGCCACCCGTCATGGCCGAGAGGGCCGGGCTCGGCAGCTGATCTAGCAAAGATCCTGTCTGGCTTGTA
CAAGGGGGAGGAGGAGGACTTTCTCCGCACGAATCCATTGAACGGATTCTCCATGAATAACCTCGGAGCTGGGTAAAGCGGACGTTTA
TCTACCCGATCCTTACTTCCAAAGTCAAGTATCTTACTTTGTGATTTCAACGCAGGAGCTGCACCGGGACGTGGAGGGCATAACAAAGC
CTAACTCCACAGCCCGAGGATCCAGAACGATCCCTTATCTGACCTCCAAAGAGGATCCGGACATAAAGGTGGAGCTGATTGACGGGGT
GTTCCACCAACTTAGCATGGACAATGCCCTAGTCGCCATTACGGCTGACTACCCCGATTGTTTCCGGCCTCCCTTCACCCCTCA

Amino acid sequence:

"WGGGEDGATRHRRGPGSAADLAKILSGLYKGEEDFLRTNPLNGFSMNNPRSWVSGRLSTRSLLPKSSILLCDFNAGAAPGRGGHT
KPNSTARGSRITPYLTSKEDPDIKVELIDGVFHLQSLMDNALVAITADYPDLFPASPSPS"

Observed Homology:

- protein kinase US2 and
- serine/threonine protein kinase US3 genes / protein_id="AAB93490.1"

APPENDIX B

T. monococcum (Acc. No. TR36938) Resistance Gene Analog Sequences

TM3-1

GGGGGGTGGGAAGACGGCCGACTGTGGCATGATTTTCTGTTATTATTGTTACCTTTCGAGGCGTCGCTTCAGACACCCGATCG
GCTTTCATTTATATTTTGTGGGATTTTAGGCGGACTACCGCCGTTCCCTCGTACTAACTTCGTTATGCAAAATTTTAAATGTG
CATTGGAAGATTAATCATGGTGCATTCATGCATGCATTGTTATATCTTACATCCGAACATCTTGCAGTACTTCTCAAAGTACTC
ACTGGCTTGTGATTTGGCGAGATGCTGACGAAGGTGATCTCATGGATGAAGAGTTCGATAGCGAGTCCGACGCCTAGAGGAATCCCA
GTCAGTCTCGCGCACCCTGGACTTGGTCACTGTATTATATTCGCTTCCGCAACCCGAATAAATTTATCGAGCCTCACCTCGACACTC
GATGAGATGTCAGTTTGAGAGTCGTCTCCCCCA

TM3-2

GAGGGTGAAGGGGAGGCCAGGAAAGGTTGGGTCCATGGAGTAATGACTGCCGTGATGAGATGCACCGACGATGTTATTTGGTGGCTG
AGCCTCCGATGATGTCAGTGTGTTTTGAAATCTGGGGTTATGTTCCGATCCGGACTAATCTGCCGTCTTCCCACCCCCCA

TM3-10

GAGGGTGAAGGGGAGGCCAGGCACCCGGTGAACCTCTGGATCTCTCTGAGCTTGGTTGGGACCGCCATCCTCTCGATGGCTTTGATCT
TCACAGGGTTGCACTCTATGCCGCTTCAGAGACCAGGAAGCCTAGGAGCTGGCCGGCTGGCACTCCGAACACGCATTTCTCTGGGTT
GAACTTAATTCGGAACCGCGCAGGTTCTCGAATGTTTCCTTCAAGTCTTCCAGCAGGTTGCCATGCTTCGCCGTCTTCCCACCCCC
CCA

APPENDIX C

Resistance Gene Analog Sequences from 'Avocet' Variety

A1-1

GGGCGAGGGGAGGCCTCCACAACCTTTGGGCTATTTTCATAGCGTGCTCTCTAGTACCATTCTGGGAACCAAGCCATCCAACCTGATGC
ACTAGCCTCCTTGCTCAGCTTGCCCAAGAACAGTTCCCATGAAGCATCGTTACCCAAAATTCGCATCTCAATCAAGCTTCTTTGTGCA
TTCATCTGAAAGCATACTTCTCGGTAACGTGTTGTGAGGATCAGCTTGCTCTGGCTGTGGGATACCAACATACTCCAGTTGGAATTC
TTCCTTACATCATCAAGCAAGATTACAAATCTTTCTAGTAAGTGCCCTTATCAAGAATTTGGCCCGTTTGGCAATTGGCTCCGCAT
CATTCACGGTAAGTTGAGCCTTTTCAGAGATAGTCTGTTGTATCTCCACTGTGTTAGCGTCTCTGAATTGGATACTTCAATAAAGAT
AACAACCTGGCACAAGCAAATTAAGGAAAATGGGAACATCAATTATAGGTAGAAGGTAATTTTCAGATATGAATAGAAATATAATCTG
AAAAGTA

A1-2

GGGGGGTGGGGAAGACGACGCTTCTACATGTTTTCAACAATGATCTTGAAAAGAAGGCCCATGATTATCAGGTACTATGTTGCTTAC
TCCAAGACTGCAATTTATGCCTCAGAGTACATGTTTGTCTGTACTTTAATATTGCTATTTAACTTACGAGAATACTTATTTCACTGAC
CACGCTTATATTTGTAACGATGCTCAAAGAAATTTATGTATAAACAAGAGAGAAAATTTGTGCTGTGCATGCAATTTAAGATGGCA
GCAAGTTTCATTGTAGAACAGCTACCCAGTGCACTGAAGTCTGAAGAAGTAATGTTGATGCTCTTGTACAGGCATATGAAATTTAACA
CGGCTTTACACGGAAAGGCAGTTAATACCATTACCATGACAATTTGACCTGTTACAGAATTATAGGTACACATTAGCCGGTAGAT
ACGGACATGAAAATGCATTGTTCTTGGATGCAATTAGTGGTGGTCATTTGCTGGTTTTCCACAACCTCTTAAAGTT

A1-4

AGGGCGAGGGGAGGCCTCCACAACCTTTGGGCTATTTTCATAGCGTGCTCTCTAGTATCATTCTGGGAACCAAGCCATCCAACCTGATG
CACTAGCCTCCTTGCTCAGCTTGCCCAAGAACAGTTCCCATGAAGCATCGTTACCCAAAATTCGCATCTCAATCAAGCTTCTTTGTGC
ATTCATCTGAAAGCATACTTCTCGGTAACGTGTTGTGAGGATCAGCTTGCTCTGGCCGTGGGATACCAACATACTCCAGTTGGAATTT
CTTCTTACATCATCAAGCAAGATTACA

A1-5

GGGGGAGGCCTCCACAACCTTTGGGCTATTTTCATAGCGTGCTCTCTAGTATCATTCTGGGAACCAAGCCATCCAACCTGATGCAC
CTCCTTGCTCAGCTTGCCCAAGAACAGTTCCCATGAAGCATCGTTACCCAAAATTCGCATCTCAATCAAGCTTCTTTGTGCATTATC
TGAAAGCATACTTCTCGGTAACGTGTTGTGAGGATCAGCTTGCTCTGGCTGTGGGATACCAACATACTCCAGTTGGAATTTCTTCCTT
ACATCATCAAGCAAGATTACAAATCTTTCTAGTAAGTGCCCTTATCAAGAATTTGGCCCGTTTGGCAATTGGCTCCGCATCATTC
ACGGTAAGTTGAGCCTTTTCAGAGATAGTCTGTTGTATCTCCACTGAGTTAGCGTCTCTGAATTGGATACTTCAATAAAGATAACAAC
CTGGCACAAGCAAATTAAGGAAAATGGGAACATCAATTATAGGTAGAAGGTAATTTTCAGATATGAATAGAAATATAATCTGAAA

A1-6

GGGCGAGGGGAGGCCTCCACAACCTTTGGGCTATTTTCATAGCGTGCTCTCTAGTATCATTCTGGGAACCAAGCCATCCAACCTGATGC
ACTAGCCTCCTTGCTCAGCTTGCCCAAGAACAGTTCCCATGAAGCATCGTTACCCAAAATTCGCATCTCAATCAAGCTTCTTTGTGCA
TTCATCTGAAAGCACACTTCTCGGTAACGTGTTGTGAGGATCAGCTTGCTCTGGCTGTGGGATACCAACATACTCCAGTTGGAATTC
TTCCTTACATCGTCAAGCAAGATTACAAATCTTTCTAGTAAGTGCCCTTATCAAGAATTTGGCCCGTTTGGCAATTGGCTCCGCAT
CATTCACGGTAAGTTGAGCCTTTTCAGAGATAGTCTGTTGTATCTCCACTGTGTTAGCGTCTCTGAATTGGATACTTCAATAAAGAT
AACAACCTGGCACAAGCAAATTAAGGAAGATGGGAACATCAATTATAGGTAGAAGGTAATTTTCAGATATGAATAGAAATATAATCTT
GAAAGTAGAAACGACGAGATGGCGTCAGGGAA

A2-1

GGGGGGTGGGGAAGACGACTTTACCCAACACCTGTATAATGATGGTAGGACTGAAGAGCACTTTGCTGTTATGGTTTGGGTGTGTG
TATCAACTGATTTTGATGTGTTGAACTCACCCAACAGATTCTTAAGTGCATACCTGCAACTGACAATGAAATTGCAATTGAAACAAC
CAATCTAGACAGACTTCAAAAATCCATTGCGCAGAGACTGAAGACCAAAAGGTTCTTAATTGTCTTGGATGATATATGAAATGCGAT
AGTGAGGATGAGTGGAAAACCTTGCTAGCTCCATTACAAAAAGGGGACGCCAAAGGCAGCATGATACTTGTACAACTCGATTCCAA
AACTAGCAAGAATGATGAAAACAGTTGATTAGTAGAAGTGAAGCTTGGAGTCTAATGACTTCTTCACATTCTTTGAAGCATGTAT
ATTTGGTGAACATAAGCCCAGGGATTATGAAGATGACTTAGCTTGATTGCAAGAGAAATTGCA

A2-2

GGGGGGTGGGGAAGACGGCTCTAACACAGCTTGTGTACAATGATACAAGAATAAAGGAGCATTTCAGCTAAGGGTCTGGCTGTGTGT
 TTCTGAAAATTTTGATCAGATGAAGCTTACCAAGGAAACCATTGAATCAGTTGCCAGCGAGTTTGAATCAACCATCAGTTGGGTCTCA
 TCCGTACGACCAACATGAACCTTGCTCCAAGAAGACCTCTCAAAAAGCTAAAAGGTAAAGGTTTCTTCTAGTACTTGATGATGCAT
 GGAATGAGGACCTGAAAAGTGGGGTACATATCGCCGTGCTCTACTTACTGGAGGAAAGGGAAGCAGAATAGTAGTAACCAAGGAA
 CAAAATGTGGGAACTAATGGGTGGGATGACTCCATACATCTAAATCAGTTATCCGACAATGATTGCTGGTCTTTGTTTCAGGAGC
 TATGCATTTGTGGATGGTAACTCCAATGCACACCCAAATTTGGAATGATAGGCATGGAAATTGTGA

A2-3

GGGGGGTGGGGAAGACGGCGCTTGTGTAGATATGTATACCATGCTAAAAGAATTAAGATCAATTTGATTGCTAATATGGGTCTGTG
 TATCCACTAATTCGATGTAGTGGGGCTAACACTTCAGATCCTAGATCATGTATGTGAAGATAGACCATATGAAAAAAGCAGTTTG
 AATAAATTGACGGAGATACCTTCAAGAAATATTAGAAACAAAAGATTTCTGCTTGTATGGATGATATGTGGGAAGAGAAGGACAGGG
 GTGGATGGATTAACTCTTGGCTCCATTGAAAAGAATAAAGTAAAAGGTTGCATGGTACTAGCAACAACCAGAACAATAATCAGTGGCA
 AAAATGATAGGAATATGGATGAAATCACATTGAGTGGTCTGGATGAGAAGGAGTTTGGCTGTTTTCAAGGCATGTGCATTTCCGA
 ATGACAACCGCAGCACCATCCTAGTTTGAATCTATTGGGAAACA

A2-4

GGGCGAGGGGGAAGCCTTTCAACTTCTTCAAAATTTCCATGCCTATCATTTCCAAATTTGGGTGTGCATGGAGTTACCATCCACAAA
 TGCATAGCTCCTGAACAAAGACAGCAATCATTGTGCGATAACTGATTTAGATAGTATGGAGTCATCCACCCATTAGTTTCCCCACA
 TTTTGTCTCTTGTGGTTACTACTATTCTGCTCCCTTTCTCCAGTAAAGTAGAGCAGCGGATATGTACCCCACTTTTCAGGGTCTCT
 CATTCATACATCATCAAGTACTAGAAGAAACCTTTACCTTTTAGCTTTTGGAGAGTCTTCTTGGAGCAAGCTCATGTTGGTCTGT
 GACGGATGAGACCAACTGATGGTTGATTCAAACCTCGCTGGCAACTGATTCATGGTTTCTTGGTAAGCTTCATCTGATCAAAATTT
 TCAGAAACACACAGCCAGACCTTAACCTGGAATGCTCCTT

A2-6

GGGGGGTGGGGAAGACGGCGGGGAGAGATACAAGACTCTAAACTGTCTATCTCATCGAGTGTGAGGTGAGGCTCGATGAATTCA
 TCTGGTAGCGGAAGCGGATATAATACAGTGACCAAAATTCAGGGTTCGACGAGACTAAGTGGGACTCCTCTAGGCGTCCGACTCGCTAT
 CAAATCTTCATCCATGAGATCGCCTTCGTGAGCATCTGGCCAAATCAACAAGCCAGTGAGTACTTTGAAAGTACTCGCAAG

A3-1

GGGGGGTGGGGAAGACGGCCCTTGTCAAGAATGTGTATGATCGGGAGAAAGTCAACTTCCCTGATGCACATGCTTGGATTGTGGTGT
 CAAAGGAATATGATGTAATTGATTGTTGGGGACATTGCTCACCAAGATACACACAGACAAGAGTCAACGACACCACCTCCTCTTAG
 CGTGGGTGCCAAATCTGATGTGTATGACTTAACAGAAGCAATAAACAAGATACTACAAGACAAGAAGTGTGATTGTGCTTGATGAT
 GTGTGAGTAAAGATGCATATAATCAAATGTGCAATGCTTTCCAGGGTATCCATGGAAGCCGTGTTATGATCAGACACGGATGGAAG
 ATGTTGACGCTCTTGTCTCAACCAAGCGTCCGCTTGTACTCCAGCCATTGGGTAGCACTGAGTCATTCAAGCTATTTTGTCTCAAGGGC
 TTTCCACAGCAACCTGAACCGCGAGTGCCCTCTGGAGCTCAAGACGGTGGCTACAAAAGTTGTTGAGAGGTGTGCTGGCCCTCCCTTC
 GCCCTCAAGCCGA

A3-2

GGGGGGTGGGGAAGACGACCCTTGCTCAAGTCATTTACAATCATGGAAGGTGAAGGACCACCTTACTCAGAGAGCATGGGTGTATG
 TTTCAAACACTTTGGTGTCAAGAGAACATTGCAAGAGATGTTGCGCTCTTTGAAAGGAAATGATTCATCTTTTGACTATGCTGATAG
 TCTGGAACAAATTGTCAACAACATTCAAATGTCTCCAGCAGGAAGGAAATTCCTTCTTGTGCTAGACAGTGTGGGATGAGATG
 TGTGACCAAGTGGAGTAGTTTGCTTATTGCCATAGCTTGTGAACCTGCCGGAAGTGTAGTTCTTGTCAACAACACAAGCAAGGGGTTG
 CAAAACAGTGGCAACAGTGTGCAAGTTCACCTAGAACCCTTGGCATGGGAGAGTTTGGCCAGTTTTCCAGTATTATGCATTTGG
 CACAACCAATGTTGAGGTTGAGAACAACCAAGACTCTTCTGTTAATTTGGTGAACAAATAGCAAAAAAAGTGGTGGCCTCCCTTCGCC
 CTCA

A3-3

GGGGGGTGGGGAAGACGGCCGACTGTGGCATCATTTCTTATATTTATTTATTTACCTTTTCGAGGCGTCTTCAGACACCCGATCG
 GTATTCAATTTATGTTTGTAGGGATTTCAAGCGGACTACCGCGTTTCCCTTTTACTTACCTTGTATGCTAATTTTGAATGTACAT
 TGGTGAATTTATTCATGACGCATTCATGCATGCATTTGTTATATCTTACGTCCAAACTGTCTTGCAGTACTTTCAAAGTACTCACTGG
 CTTGTTGATTGCGCAGATGCTGACGAAGGCATCTCATGGATGAAGAGTTTGATAGTGAGTCCGACACCTAGAGGAGTCCCAGTCAG
 TCTCGTGGACCTGGATTGTTGCTACTGTTTATATCCGCTTCCGCTACCAATAAATTCATCGAGCCTCACCTCGACGCTCGAGGAG
 ATGACAGACTTAGAGTCTGTATCTCTCCCGCGCTCTTCCCCACC

A3-4

GGGGGGTGGGGAAGACGGCTGGGAGAGAAATACAAGACTCTAAGTCTGTCTCATCTCATCGAGCGTGGAGGTGAGGCTCGATGAATTTA
 TTTGGTAGCGGAAGCGGATATAAAATAGTGACCAAAATCCAGGGTTCGACGAGACTGACTGGGACTCCTCTAGGTGTGCGACTCGCTAT
 CAAACTCTTCATCCATGAGATCGCCTTCGTGAGCATCTGGCCAAATCAACAAGCCAGTGAGTACTTTGAAAGTACTCGCAAGACAGTT
 CGGACGTAAGATATAACAAACGCATGCATGAATGCATCATGAATAATCCACCAATGTACATTCAAAAATTACCACAAC

A3-5

GGGGGGTGGGGAAGACGGCGGGGAGAGAAATACAAGACTCTAAGTCTATCATCTCATCAAGCATCGAGGTGAGGCTCGATGAATTTA
 TTTGGTAGCGGAAGCGGATATAAAACAGTGACCAAAATCCAGGGTTCGACGAGACTGACTGGGACTCCTCTAGGCGTCAGACTCGCTAT
 CATATTCTTCATCCATGAGATCGCCTTCGTGAGCATCTGGCCAAATCAACAAGCCAGTGAGTACTTTGAAAGTACTCGCAAGACAGTT
 CGGATGTAAGATATACCAATGCATGCATGAATGCGTCATGATTAAATTCACCAATGTACATTCAAAAATTAGCATAACGAGGTAAAGTAA
 AAGGGAACACAGGTATTCCGCTGAAATCCACAAAACATAATGAAGACCGATCGGGTGTCTGAAGCGACGCTCGAAAGGTGAAC
 AATAAATATAAGGAAATGATGCCACAGTCGGCGCTCTTCCCCACC

A3-6

GGGGGGTGGGGAAGACGGCCCAAGACTCTTGATAGAGTGAAGATAATAGTAAAAGATAGGCCCTTCGACAGGGGAAGCAGAGGTT
 TGCATGCCCTTTTATGGTTGGATGCACAAAATCTTAATGCCGAAAGATGTCACTTTATGTTGCCCTTGTGATATGAACCTTTATTA
 TGCAGTCCATCGCCTTTTATTACTTCGACCTCACAAAATCGTATAAGCTTATTTCTCCACACCAATCAATCATACATATTAGAGC
 AATTTTATTGCTTT

A4-1

GGGGGGGTGGGGAAGACGGCGGGGAGAGAAATACAAGACTCTAAGTCTGTCATCTCATCGAGAGTCGAGGTGAGGCTCGATGAATTTA
TTTGGTAGTGGAAGCGGATATAAAACAGTGACCAAAATCCAGGGTCGCACGAGACTGACTGGGACTCCTCTAGGTGTCGGACTCGCTAT
CAAACCTCTTCATCCATGAGATCGCCTTCGTGAGCATCTGGCCAAATCAACAAGCCAGTGAATACCTTTGAAAGTACTCGCAAGACAGTT
CGGACGTAAGATATAACAAATGCATGCATGGATGTGTCATGAATAATTCACCAATGTACATTCAAAAATTAGCATAACAAGGTAAGTT
AAAGGGTAACAGCGGTAGTCCGCCTGAAATCCCAACAAAACATAAATACTACCGATCGGGTGTCTGAAGCGACGCCTCGAAAGGTGA
ACAAATAAATAAGGAAATGATGCCTCAGTCGGCCGTCTTCCCACCC

A4-2

GGGGGGGTGGGGAAGACGGCCCTTGCTCAAGTCATTTACAATCATGAAAATGTGAAGGGTCACCTTTGGGCACAGAGCATGGGTGTATG
TTTCAAAACATTTTCGGAGTGAAAAGGACGTTGCGAGGAGATGTTGTGCTCCATTAAAGGAAATGATTTCATCTTTTGACCGTGTGATGG
TCTAGAGACGGTTGTCAACAATATTTCAAATGTCATCCGGCATGGAAGATTCTTTCTGTGCTTGATAGTGTGTTGGGATGAGATGTGT
GATCAGTGGAGTAGTTTGCTTACTGCCATAGCTGGTGAAGTACCGGGGAAGTGAATCTTTGTCACAAACACAAAGCAAAGGGTTGCAG
ACACAGTAGCAACAATGTGTCAAGTTCACCTAGCACCTTTGCCATGGGAGAGTTTTGGTCAGCCTTCCAGTATATGCTTTTGGCAC
CACCCTGCTGTGGCAGAGAACAATCAGACTCTTCTGTTAATTGGTGAGCAAAATAGCAAAAAAATAGATGGCTTCCCCTTCGCCCTC
AAGCCGAATTCCA

A4-3

GGGGGGGTGGGGAAGACGGCGGGGAGAGAAATACAAGACTCTAAGTCTATCATCTCATCGAGCGTCGAGGTGAGGCTCGAAGAATTTA
TTTGGTAGCGGAAGCGGATATAACAGTGACCAAAATCCAGGGTCGCATGAGACTGACTGGGACTCCACTAGGTGTCGGACTCCTCATC
AAACTCTTCATCCATGAGATCGCCTTCGTGAGCATCTGGCCAAATCAACAAGCCAGTGAATCTTTGAAAGTACTCGCAAGACAGTT
GGACGTAAGATATAACAAATGCATGCATGGATGCGTCATGAATAATTCACCAATGTACATTCAAAAATTAGCATAACAAGGTAAGTAA
AAGGGAAACGACGGTAGTCCGCCTGAAATCCCAACAAAACATAAATACTACCGATCG

A4-4

GGGGGGGTGGGGAAGACGGCCGACTGTGGCATCATTTCCCTTATATTTATTTGTTTACCTTTTCGAGGGCTCGCTTCAGACACCCGATC
GGTCTTCATTTATGTTTTGTGGGATTTTCAGGCGGACTACCGCCGTTTCCTTTTACTTACCTCGTTATGCTAATTTTAAATGTACATT
GGTGAATTAATCATGACGCATTCATGCATGCTTTGTTATATCTTGTCCGAAGTGTCTTGGAGTACTTTTCAAGTACTCAGTGGCTT
TTTGATTTGGCCAAATGCTGACGAAGGGCATCTCATGGATGAAGAATTTGATAGCGAGTCCGACGCGCTAGAGGAGTCCAGTCGGTCT
CGTACGACCTGGATTTGGTCACTGTATATATCTGCTTCCGCTACCAATAAATTCATCGAGCCTCACCTCGACGCTCGATGAGATG
ACAGTTTTAGAGACTTGATCTCTCTCCTGCCGTCTTCCCACCC

A4-5

GGGGGGGTGGGGAAGACGACTCTCAAACCTGGCATCTCATCGAGCGTCGAGGTGAGGCTCGATAAATTTATTCGGGTTCGGAAGCGGC
TATAATACAGTGACCAAAATCCAGGGTCGCATGAGACTGACTGGGATCCTCTAGGCGTCGAACCTCGCTATTGAAGTCTTCATCCATGA
GATCGCCTTCGTGAGCATCTGGCCAAATCAACAAGCCAGTGAATCTTTGAAAGTACTCGCAAGACAGTTTCGGAGTAAAGATATAACT
AAGCATGCATGAATCGGTGATGAATAATCCACCAATGTACATTCAAAAATTAGCACAACAAGGTAAGTAAAAGGGAAACGGCGGTAG
TCCGCCTGAAATCCCAACAAAACATAAATGAAGACCGATCGGGTGTCTGAAGCGATGCCTCGAAAGGTGAACAAATAAATAAAGGAA
ATGATGCCACAGTCGGTCTCTTCCCACCC

A4-6

GGGGGGGTGGGGAAGACGGCGGGGAGAGATACAAGACTCTAAGTCTGTCATCTCATCGAGCGTCGAGGTGAGGCTTGATGAATTTA
TTTGGTAGCGGAAGCGGATATAAAATAGTGACCAAAATCCAGGGTCGCACGAGACTGACTGGGACTCCTCTAGGTGTCGGACTCGCTAT
CAAACCTCTTCATCCATGA

A5-1

GGGGGGGTGGGGAAGACGCGCTTCTACATGTTTTCAACATGATCTTGAAGAAGGCCCATGATTATCAGGTACTATGTTGCTTAC
TCCAAGACTGCAATTTATGCCTCAGAGTACATGTTGTCTGTACTTTAATCTGCTATTTAACTTACGAGAATACTTATTTAGTGAC
CACGCTTATATTTGTAACGATGCTCAAAGAAATTTATGTATAAACAAGAGAGAAATTTGTGCTGTATGCAATTTAAGATGGCA
GCAAGTTTTCATTGTAGAACAGTACCCAGTGCATGAAGTCTGAAGAAGTAATGTTGATGTTCTTGTACAGGCATATGAATTTAACA
CGGCTTTCACACGGAAGGCAGTTAATACCATTACCATGACAATTTGACCTGTTACAGAATTATAGGTACACATTAGCCGGTAGAT
ACGGACATGAAAATGCATTGTTTCTTGATGCAATTAGTGTGTGTCATTTTGTGTTTTCCACAACTCCTTAAA

A5-2

GGGGGGGTGGGGAAGACGGCGGGGAGCTACAGTGTGGGCTGGGTGCGTGTAGCTGGGCGGCCAGCTGGCTGGGCGCACAGGTA
AGTCTTTCCCCTTTCTTTCTGTTTCTGTTTTATTAAATATATCTGCAACTTTGTTGAATTAATTAATAATACCTAGGCAACTCCAAA
ATTACCAAACTACTCCTGGTCCATAGTTGGATTATTTCAACATGAACATTTTGTGAGATATTTGAGCATTTAGATATTTTA
TATTATTTTAAATGCCAAATTCAAATATCTATGATTAAATTCAGAGCCCTATGATGGCCTAGAAAAATGTGCACCACTTTTGTGAG
AGGTTCTGAACCAAGGCTAAAATGATGGACATTTTAGAAGGGCATTTAGGTTTCAATGAAAAATTTTATAGTAAACCTAGTTGGTTT
CAGAGGGGACTGGGGTCTGTGATCCCCATTTCAAGTTTCTGATGAAAAAGTAAACATGATGCAACACTCTAATGCATGACTAGCTA
GGGTGTGACATCTAGCCCCTGACAGCTGACAGGTGTTCTTTCGCCAATCAGACTT

A6-1

GGGGGGGTGGGGAAGACGACTTTACCCAACACCTGTATAATGATGGTAGGACTGAAGAGCACTTTGCTGTTATGGTTTGGGTGTGTG
TATCAACTGATTTTGTGTGTTGAAACTACCCAACAGATTCTTAAGTGCATACCTGCAACTGACAATGAAATTGCAATTGAAACAAC
CAATCTAGACAGACTTCAAAAATCCATTGCGCAGAGACTGAAGACCAAAAGGTTCTTAATTTGCTTGGATGATATATGGAAATGCGAT
AGTGAGGATGAGTGGAAAACCTTGCTAGTCCATTCAAAAAGGGGACGCCAAAGGCAGCATGATACTTGTACAACTCGATTCCCAA
AACTAGCAAGAAATGATGAAAACAGTTGATTTCAGTAGAAGCTGAAGGCTTGGAGTCTTAATGACTTCTTACATCTTTGAGCATGTTAT
ATTTGGTGAACATAAGCCAGGGATTATGAAGATGACTTAGCTTGTATTGCAAGAGAAATGCAAGCAAGTTAAAGGGCTTCCCCTC
ACCTTCA

A6-2

GGGGGGGTGGGAAGACGACTTTCACCCAACGCCTGTATAATGATGGTAGGACTGAAGAGCACTTTGCTGTTATGGTTGGGTGTGTG
TATCAACTGATTTTGTATGTGTGAAACTCACCCAACAGATTCTTAAGTGCATACCTGCAACTGACAAATGAAATTGCAATTGAAACAAC
CAATCTAGACAGACTTCAAAAATCCATTGCGCAGAGACTGAAGACCAAAAGTTCTTAATTGCTTGGATGATATATGGAATGCGAT
AGTGAGGATGAGTGGAAAACCTTGCTAGCTCCATTACAAAAGGGGACGCCAAAGGCAGCATGATACTTGTACAACTCGATTCCCAA
AACTAGCAAGAAATGATGAAACAGTTGATTAGTAGAAGTGAAGGCTTGGAGTCTAATGACTTCTTACATTCTTTGAAGCATGTAT
ATTTGGTGAACATAAGCCCAGGGATTATGAAGATGACTTAGCTTGTATTGCAAGAGAAATTGCAAGCAAGTTAAAGGGCTTCCCCCTC
GCCTTCA

A6-4

GGGGGGGTGGGAAGACGACGCTTGTAGATATGCATACCATGCTAAAAGAATTAAAGATCATTTTGATTGCTAATATGGTCTGTG
TATCCACTAACTTCGATGTAGTGGGGCTAACACTTCAGATCCTAGATCATGTATGTGAAGATAGACCATATGAAAAAAAAGCAGTTT
GAATAAATTGCAGGAGATACTTCAAGAAAATATTAGAAACAGAAGATTCTGCTTGTAAATGGATGATATGTGGGAAGAGAAGGACAGG
AGTGGATGGATTAACTCTTGGCTCCATTGAAAAGTAATAAAGTAAAGGTTGCATGGTACTAGCAACACCAGAACAAAATCAGTGG
CAAAAATGATAGGAACATATGGATGAAATCACATTGAGTGGTCTGGATGAGAAGGAGTTTGGCTGTTTTCAAGGCATGTGCTTTTCG
CAATGACAACCGCAGCACCATCTGGTTTGAATCTATTGGGAAACAAATTGCTAAAGCACTAAAGGCTTCCC

A6-6

GGGGGGGTGGGAAGACGCGGGGAGAGAGATACAAGTCTCAAAAATGTCTCATCTCGAGCGTCGAGGTGAGGCTCGATGAATTTA
TTTGGTAGAGGAAGCGGATATAATACAGTGACCTAATCCAGGGTCGTACGAGACTGACTGGGACTCCTCTAGGCGTCGGACTCGCTAT
CAAATCTTTCATCCATGAGATCGCTTCGTCAGCATCTGGCCAAATCAACAAGCCAGTGAGTACTTTGAAAGTACTCGCAAGACAGTT
CGGACGTAAGATATAACAAATGCATGATGAATGCATCATGATTAAATCACCATGTACATTCAAAAATTAGCATAACGAGGTAAGTA
AAAGGGAAACGGCGGTAGTCCGCTGAAATCCCAAAAACATAAATGAAGACCGATCGGGTGTCTGAAGCGACGCTCGAAAGGTGAA
CAATAAATATAAGGAAAGGATGCCATAGTCGGCGCTCTTCCCCACC

A8-1

GGGGGGGTGGGAAGACGCGGGGAGAGAGATACAAGACTCTAAGTCTGTCTCATCTCACCGAGCGTCGAGGTGAGGCTTGATGAATTTA
TTTGGTAGCGGAAGCGGATATAAATAGTGACCAATCCAGGGTCGACGAGACTGGCTGGGACTCCTCTAGGTGTGCGGACTCGCTAT
CAAATCTTTCATCCATGAGATCGCTTCGTCAGCATCTGGCCAAATCAACAAGCCAGTGAGTACTTTGAAAGTACTCGCAAGACAGTT
CGGACGTAAGATATAACAACGCGATGCATGAATGCCGTCATGAATAATCCACCAATGTACATTCAAAAATTA

A8-2

GGGGGGGTGGGAAGACGACGAAGAAGAAGAGGACGATGACAACTATGTGCCCTTGAATACGGTGATGCTGCAACGGGGGAACTGC
TGAAGATCAAGAGGAACAGACGATGTGCCAGATGATGCTGCAACGGGGGAAGCTGCTGAAGATCAAGAGGAACAGACGATGTGCCC
GATCATGATGTTCTCCGCGGGTCATTGTTGATGCAATGACGCAATGCGAAAGTCAAAAAGAAGCTGAAGTTCGATCGCATGTTAG
AGGATCACA AAAAAGGATTGTACCCCAATTGCCAAGATGGCAACACAAGCTCGGTACCGTACTGGAATTGATGCAGTGGAAGGAGG
AGAATGCTATGCCTGATAAAGGATTTGAGAAGCTACTGAAAATATTGAAGAAGAAGCTTCCAAAGGATAACGAATTGCCCGACAGTAC

A8-3

GGGGGGGTGGGAAGACGCGGAAGAAGAAGAGGACGATGACAACTATGTGCCCTTGAATACGGTGATGCTGCAACGGGGGAACTGC
TGAAGATCAAGAGGAACAGACGATGTGCCAGATGATGCTGCAACGGGGGAAGCTGCTGAAGATCAAGAGGAACAGACGATGTGCCC
GATGATGATGATCTCCGCGGGTCATTGTTGATGCAATGACGCAATGCGAAAGTCAAAAAGAAGCTGAAGTTCGATCGCATGTTAG
AGGATCACA AAAAAGGATTGTACCCCAATTGCCAAGATGGCAACACAAGCTCGGCACCGTACTGGAATTGATGCAGTGGAAGCCAGA
GAATGCTATGCCTGATAAAGGATTTGAGAAGCTACTGAAAATATTGAAGAAGAAGCTTCCAAAGGATAACGAATTGCCCGACAGTAC
GTACT

A8-4

GAAGGCGAAGGGGAAGCCACGGTGCAATAAAAAACAAACGTCAAAATAAGGAAGAAAGTCCGGACAACCCAACGGTCAACGCCGGATTTC
AGGGGCTCTCGACCCGGTCAACAAAAGAATCCATTTAAAGGCAGCAAGAGGGACCGTCCGGCTTGAACAAAGTCTTGGACAGACTGT
GCCAAATTCATGGCACCCCGGAAAACCTGCGAACCACACTCACAGAGAATGCTGGGTCTTCAAGCAGGCCGGCAAATTAATGTCTGA
ACACAAGGGAGGGGAAACCAAGTGGAGACGAGGATGAGCCTCACCGCCGAACACTGGAGGACAGAAGAAATTCCTGCCAGAGGTA
AAAACAGTTAACATGATCCACGCGACTCACATATCCACGAGAATGCGCGAATGCGCGCTCAGAGACGCATACGCTATAGAGCCCGTCG
CACCCAAACTTAACCACTGGCCGTCTTCCCCACCCCCCCA

A8-5

GGGGGGGTGGGAAGACGACTCTCAAACTGGCACCTCATCGAGCGTCGAGGTGAGGCTCGATAAATTTATTGGGTTGCGGAAGCGGC
TATAATACAGTGACCAATCCAGGGTCGATGAGACTGACTGGGATTCTCTAGGCGTCGAACTCGCTATTGAACTCTTCATCCATGA
GATCGCTCTTCTGTCAGCATCTGGCCAAATCAACAAGCCAGTGAGTACTTTGAAGTACTCGCA

APPENDIX D

Resistance Gene Analog Sequences from 'Chinese spring' variety

CS2-1

GGGGGGGTGGGGAAGACGGCTCTAACACAGCTTGTGTACATGATACAAGAATAAAGGAGCATTCCAGTTAAGGGTCTGGCTGTGTG
TCTCTGAAAATTTTGATCAGATGAAGCTTACCAAGGAAACCATTGAATCAGTTGCCAGCGAGTTTGAATCAACCATCAGTTGGGTCTC
ATCCGTCACGACCAACATGAACCTTGCTCCAAGAAGACCTCTCAAAAAGCTAAAAGGTAAAGGTTCTTCTAGTACTTGATGATGA
TGGAATGAGGACCTGAAAAGTGGGTACATATCGCCGTGCTCTACTTACTGGAGGAAAGGGAAGCAGAATAGTAGTAACCAAGGA
ACAAAATGTGGGGAACATAATGGGTGGGATGACTCCATATCTATAATCAGTTATCCGACAATGATTGCTGGTCTTTGTTTCAGGAG
CTATGCATTTGTGGATGGTAACCAATGCACACCCAAATTGGAAATGATAGGCATGGAATTTGAAGAAGTTGAAGGCTT

CS2-3

GGGGGGGTGGGGAAGACGGCGCTTGTAGATATGTATACCAAGCTAAAAGAATTAAAGATCATTTTGATTGCTAATATGGGTCTGTG
TATCCACTAAGTTAGATGTAGTGGGGCTAACACTTCAGATCCTAGATCATGTATGTGAAGATAGACCATATGAAAAGAAGCAGTTTG
AATAAATTGCAGGAGATACTTCAAGAAAATATTAGAAAACAAAAGATTCTGCTTGTAAATGGATGATATGTGGGAAGAGAAGGACAGGG
GTGGATGGATTAAACTCTTGGCTCCATTGAAAAGTAATAAGTAAAAGTTGCATGGTACTAGCAACAACAGAACAAAATCAGTGGC
AAAAATGATAGGAATATGGATGAAATCACATTGACTGGTCTGGATGAGAAGGAGTTTGGCTGTTTTCAAGGCATGTGCATTTCCG
AATGACAGCCGCGAGCACCATCTAGTTGCAAT

CS2-4

GGGGGGGTGGGGAAGACGACTTTCACCAACACCTGTATAATGATGGTAGGACTGAAGAGCACTTTGCTGTTATGGTTGGGTGTGTG
TATCAACTGATTTTGATGTGTTGAAACTCACCAACAGATTCTTAATGTCATACCTGCAACTGACAATGAAATGCAATTGAAACAAC
CAATCTAGACAGACTTCAAAAATCCATTGCGCAGAGACTGAAGACCAAAAGGTTCTTAATTGCTTGGATGATATATGAAAATGCCAT
AGTGAGGATGAGTGGAAAACCTTGCTAGCTCCATTCAAAAAGGGGACGCCAAAGGCAGCATGATACTTGTCACAACTCGATTCCCAA
AATAGCAAGAAATGATGAAAACAGTTGATTCAGTAGAATGCAAGGCTTGGAGTCTAATGACTTCTTACATTCTTTGAAGCATGTAT
ATTTGGTGAACATAAGCCCAGGGATTATGAA

CS2-5

GGGGGGGTGGGGAAGACGGCACTCACACAACACATATTTAGAGAACTAGAGAGTTCTTTCCAAGTTTCGGTTTGGATATGTGTCTCTC
TCGATTTTAATGCAGATAGGTTGCTACAAGAGATTGTGGACAAGATCCCTAAAGTAAATGGTGAAAAGGTTAATACTAGCAACCAAGA
GTTAATTGAACAAAGGTTGAAATCTAAGCGGTTGTTGCTTGTGTTTAGATGATGAGTGGACATATCATGAGGATGAGTGGAAAACTA
TTAGCTCCATTAAACATAACAATGGAGACAAGGGTAATGTGGTTATAGTCACTACACGAATTCCAAAGGTAGCAAGTATGGTCACCA
CAACTAATCTTCAATCGACGTGGAACGTCCTAAGTCATGGGATACTATGCTTTCTTCAAGTATGTGTATTTGGTGACCAACGACC
ATGGGAAGACCACCCTGAATTACGTGATGTTGGGAATACAATAGTAACAAAATTGAAAGGCTTCCCCCTCGCCCTCA

CS2-6

GAGGGCAGGGGGAAGCCTTTTAGTGCTTTAGCAATTTGTTTCCAATATATTGCAAACTAGGATGGTGCTCGCAGTTGTCATTGCGA
AATGCACATGCCTTGAAAACAGCCAAAACCTCTTCTCATCCAGACCTCAATGTGATTTCATCCATAGTTCTTATCATTTTTTGCCA
CTGATTTTGTCTGTTGTTGCTAGTACCATGCAACCTTTTACTTTATTTACTTTTCAATGGAGCCAAAAGTTTATCCATCCACCCC
TGTCTTCTCTTCCGCATATCATCCATTACAAGCAGAAATCTTTTG

CS5-1

GGGGGGGTGGGGAAGACGGCGCTTCTACATGTTTTCAACAATGATCTTGAAAAGAAGGCCCATGATTATCGGGTACTATGTTGCTTAC
TCCAAGACTGCAATTTATGCCTCAGAGTACATGTTTGTCTGTACTTTAATATTGCTATTTAACTTACGAGAATACTTATTTCAAGTGAC
CACGCTTATATTTGTAACGATGCTCAAAGAAATTTATGTATAAACAAGAGAGAAAATTTGTGCTGTGATGCAATTTTAAGATGGCA
GCAAGTTTCAATTGTAGAACAGCTACCCAGTGCAGTGAAGTCTGAAGAAGTAATGTTGATGTTCTTGTACAGGCATATGAAATTTAACA
CGGCTTTCACACGGAAAGGAGTTAATACCATTACCATGACAATTTGACCTGTTACAGAAATTTAGGTACACATTAGCCGGTAGAT
ACGGACATGAAAATGCATTGTTCTTGTGATGCAATTAGTGTGTGTCATTTGCTGTTTTCCACAA

CS5-2

GAAGGTGAGGGGAGGCTCCACAACCTTTGGGCTATTTTCATAGCGTGTCTCTAGTATCATTTCTGGGAACCAAGCCATCCAACCTGAT
GCATAGCCTCCTTGCTCAGCTTGCCCAAGAACAGTTCCCATGAAGCATCGTTACCCAAAATTCGATCATCAATCAAGCTTCTTTGTG
CATTCATCTGAAAGCATACTTCTCGGTAACGTGTTGTGAGGATCAGCTTGCTGCTGGCTGTGGGATACCAACATACTCCAGTTGGAATT
TCTTCTTACATCATCAAGCAAGATTACAAATCTTTTCTAGTAAGTGCTTTATCAAGAATTTGGCCCGTTTGGCAATTGGCTCCGC
ATCATTCACGGTAAGTTGAGCCTTTAGAGATAGTCTGTTGTATCTCCACTGTGTTTAGCGTCTCTGAATTGGATACCTCAATAAAG
ATAACAACCTGGCACAAGCAAAATTAAGGAAAATGGGAACATCAATTTAGGTAGAAGTAATTTTCAGATATGAATAGAAATATAATC
TGAAAAGTAGAACGACGAGATGGCGTCAGGAAGGATGTGCCAAAAGAGATGCATGTATGT

CS6-1

GGGGGGTGGGAAGACGGCTTCACCCAACACCTGTATAATGATGGTAGGACTGAAGAGCACTTTGCTGTTATGGTTGGGTGTGTG
TATCAACTGATTTTGATGTGTGAAACTCACCAACAGATTCTTAAGTGCATACCTGCAACTGACAATGAAATTGCAATTGAAACAAC
CAATCTAGACAGACTTCAAAAATCCATTGCGCAGAGACTGAAGACCAAAAGGTTCTTAATTGTCTTGGATGATATATGGAAATGCGAT
AGTGAGGATGAGTGGAAAACCTTGCTAGCTCCATTTACAAAAGGGGACGCCAAAGGCAGCATGATACTTGTCAAACTCGATTCCCA
AACTAGCAAGAATGATGAAAACAGTTGATTCAGTAGAACTGCAAGGCTTGGAGTCTAATGACTTCTTCACATTCTTTGAAGCATGTAT
ATTTGGTGAACATAAGCCCAGGGATTATGAAGATGACTTAGCTTGTATTGCAAGAGAATTGCAAGCAAGTTAAAGGGCTTCCCCTCAC
CTTCA

CS6-4

GGGGGGTGGGAAGACGACTCTCAAAGTGGCATCTCATACGAGCGTCGAGGTGAGGCTCGATAAATTATTTCGGGTGCGGAAGCGG
CTATAATACTGACCAAATCCAGGGTCGCATGAGACTGACTGGGATTCTCTAGGCGTC

CS8-1

GAAGCGAAGGGGAAGCCACGATGGAACAAGAACAATGGCGCCACAACAATGGCGAAGCTAACGAGATGGCGGTAAATGCCGAATTT
TGTAACCCGTAGTCCGGGCCATGGAAAAAGCCCTCTAAAGGCAACCAAGAGGGCCCATCCAAGTTGGATGCAATCCTAGATAGGCCTT
GCCAGATCCACGACACCTTGAGAAGCCCGCCAAACACACCAACCATAACTGTTGGGTGATAAAACAAGCGGGCAAGCTGACAGCCGAA
GACCCTAGCAAAAGCCCTCAAAGCGAGGACGATTAGGAGCCCTGCAGGCCGAACAGTGGGAAGTCAAATC



APPENDIX E

Resistance Gene Analog Sequences from 'Pastor' Variety

P1-1

TGTTAGCGTCTCTGAATTGGATACTTCAATAAAGATAACAACCTGGCACAAGCAAATTAAGGAAATGGGAACATCAATTATAGGTA
GAAGGTAATTTTCAGATATGAATAGAAATATAATCTGAAAAGTAGAAACGACGAGATGGCGTCAGGGAAGGATGCCCAAAAAGAGAT
GCATGTATGTGTCAAAAATAAATTACACAAAAGGCAACAATAACAATTTACGACGTCGATTTTATGCATATTTACTTGGTTTACG
CAAAATATGCTTAAAAAGAAAACCTTTTAAGGAGTTTGTGGAAAACAGCAAATGACACACACTAATTGCATCAAGAAACAATGCAT
TTTTCATGTCCGTATCTACCGCTAATGTGTACCTATAATTCTGTGAACAGGTCAAATTGTCTGGTGAATGGTATTAAGTGCCTTTC
CGTGTGAAGCCGTGTTAAATTTTATATGCCTGTACAAGAACATCAACATTACTTCTTCAGACTTCAGTGCCTGGTAGCTGTTCTAC
AATGAAACTTGCT

P1-2

GAGGGCGAGGGGGAGGCTCCACAACCTTTGGGCTATTTTCATAGCGTGCTCTCTAGTATCATTCTGGGAACCAAGCCATCCAACCTGAT
GCACTAGCCTCCTTGCTCAGCTTGCCCAAGAACAGTTCCCATGAAGCATCGTTACCCAAAATACGCATCTCAATCAAGCTTCTTTGTG
CATTCACTGAAAGCATACTTCTCGGTAACGTGTTGTGAGGATCAGCTTGCTCTGGCTGTGGGATACCAACATACTCCAGTTGGAATT
TCTTCCTTACATCATCAAGCAAGATTACAAATCTTTTCCTAGTAAGTGCCCTTTATCAAGAATTGGCCCGTTGGCAATTGGCTCCGC
ATCATTCCACGGTAAGTTGAGCCTTTCAGAGATAGTCTGTTGTATCTCCACTGTGTTTAGCGTCTCTGAATTGGATACTTCAATAAAG
ATAACAACCTGGCACAAGCAAATTAAGGAAAATGGGAACATCAATTATAGGTAGAAGGTAATTTTCAGATATGAATAGAAATATAATC
TGAAAAGTAGAAAC

P1-4

GAGGGCGAGGGGGAGGCTCCACAACCTTTGGGCTATTTTCATAGCGTGCTCTCTAGTATCATTCTGGGAACCAAGCCATCCAACCTGAT
GCACTAGCCTCCTTGCTCAGCTTGCCCAAGAACAGTTCCCATGAAGCATCGTTACCCAAAATTCGCATCTCAATCAAGCTTCTTTGTG
CATTCACTGAAAGCATACTTCTCGGTAACGTGTTGTGAGGATCAGCTTGCTCTGGCTGTGGGATACCAACATACTCCAGTTGGAATT
TCTTCCTTACATCATCAAGCAAGATTACAAATCTTTTCCTAGTAAGTGCCCTTTATCAAGAATTGGCCCGTTGGCAATTGGCTCCGC
ATCATTCCACGGTAAGTTGAGCCTTTCAGAGATAGTCTGTTGTATCTCCACTGTGTTTAGCGTCTCTGAATTGGATACTTCAATAAAG
ATAACAACCTGGCACAAGCAAATTAAGGAAAATGGGAACATCAATTATAGGTAGAAGGTAATTTTCAGATATGAATAGAAATATAATC
TGAAAAGTAGAAAC

P1-5

GAGGGTGAGGGGGAGGCTCCACAACCTTTGGGCTATTTTCATAGCGTGCTCTCTAGTATCATTCTGGGAACCAAGCCATCCAACCTGAT
GCACTAGCCTCCTTGCTCAGCTTGCCCAAGAACAGTTCCCATGAAGCATCGTTACCCAAAATTCGCATCTCAATCAAGCTTCTTTGTG
CATCCATCTGAAAGCATACTTCTCGGTAACGTGTTGTGAGGATCAGCTTGCTCTGGCTGTGGGATACCAACATACTCCAGTTGGAATT
TCTTCCTTACATCATCAAGCAAGATTACAAATCTTTTCCTAGTAAGTGCCCTTTATCAAGAATTGGCCCGTTGGCAATTGGCTCCGC
ATCATTCCACGGTAAGTTGAGCCTTTCAGAGATAGTCTGTTGTATCTCCACTGTGTTTAGCGTCTCTGAATTGGATACTTCAATAAAG
ATAACAACCTGGCACAAGCAAATTAAGGAAAATGGGAACATCAATTATAGGTAGAAGGTAATTTTCAGATATGAATAGAAATATAATC
TGAAAAGTAGAAACGAC

P1-6

GAGGGCGAGGGGGAGGCTCCACAACCTTTGGGCTATTTTCATAGCGTGCTCTCTAGTATCATTCTGGGAACCAAGCCATCCAACCTGAT
GCACTAGCCTCCTTGCTCAGCTTGCCCAAGAACAGTTCCCATGAAGCATCGTTACCCAAAATTCGCATCTCAATCAAGCTTCTTTGTG
CATTCACTGAAAGCATACTTCTCGGTAACGTGTTGTGAGGATCAGCTTGCTCTGGCTGTGGGATACCAACATACTCCAGTTGGAATT
TCTTCCTTACATCATCAAGCAAGATTACAAATCTTTTCCTAGTAAGTGCCCTTTATCAAGAATTGGCCCGTTGGCAATTGGCTCCGC
ATCATTCCACGGTAAGTTGAGCCTTTCAGAGATAGTCTGTTGTATCTCCACTGTGTTTAGCGTCTCTGAATTGGATACTTCAATAAAG
ATAACAACCTGGCACAAGCAAATTAAGGAAAATGGGAACATCAATTATAGGTAGAAGGTAATTTTCAGATATGAATAGAAATATAATC
TGAAAAGTAGAAAC

P2-1

GAGGGTGAGGGGGAGGCAAAACCTCAGCCGAAGCTGACGAAGTTATCGTCTTCACTGATCAGATGAACCGGGGCTTCTCACC GCCG
GATCAAAATTTCTCCGTGATGTTCTGCACTTTATGCAACTTCATCTCAAGATATCGGGCCAACTCCATTTCACAGCTCTGCAACTTC
CAAGTCTTCTGTGAAGTATCCCTTCAAGAAAGAACTACTGTTGATCTCTTCAGAGAGTATTTCTATTGTAACCGCCAGAAATGATTAC
CGATGGCCCATCCAAGAACTCGGAGGGATCTCAATTCAGCATCATAGGGATGCCATTTCCCTTATGCTTCCCTTCAAGCCATCCC
ATAGACTGGAATCAAACTGATTCTATTGCAAGACACATCTCCTCTGATGAAAATCCTTTACAGGCTATCGTGATCAACGGCTCA
ACCAAAGCATCTCTTCTCGGCGGCTTCCCTT

P2-2

GAGGGCGAGGGGGAAGCCTTTACTGGAGCACGGGAAGACTAAGTGGAGCACACAACACGACAACCATTTGCGGTGTACGTGTTATCA
CACATGTGTTAACATAAGGAAACGTATGTGTGACTTACTAATCATCGCACACGAGTACTCGCACACGCATCATGTGCGAGACTCCTAG
CCATCGCAAGCAACTAGTATGCATTTTTCAAAGTTTTTACTAGACATAGAATCACACACGAATTAACCTAGTAACAGTATGTGTTAT
ATTTTCTCATCGCAAATAGTTCATCCGAGTGGGCTATTTGTGCGGTATCACACACATCTTGTTTACACGAATCATTTGTGTTCTTTTG
GCTCATAGCAAATAGTTCATCTATGTGAACGTGCATGCCGACATCACACACATCTTGTAAATGAACCATTTCTGTTGTGTTCTTAG
TCGAAATAGTTTGTCCGAGTGAACCGTATGCCGTTCTCC

P2-3

AGGGCGAGGGGGAAGCCTTTTAGTGCTTTAGCAATTTGTTTCCCAATAGATTGCAAACTAGGATGGTGTGCGGGTGTGCTATGCGAA
ATGCACATGCCTTGAAAAACAGCCAAACTCCTTCTCATCCAGACCACTCAATGTGATTTTCATCCATAGTTCCTATCATTTTTGCCAC
TGATTTTGTCTGCTGTTGCTAGTACCATGCAACCTTTTACTTTATTACTTTTCAATGGAGCCAAGAGTTTAAATCCATCCACCCCTG
TCCTTCTCTTCCACATATCATCCATTACAAGCAGAAATCTTTTGTCTTAATATTTTCTTGAAGTATCTCCTGCAATTTATTCAAAC
TGCTTTTTTTTCATATGGTCTATCTTACATACATGATCTGAGATCTGAAGTGTGAGCCCACTACATCGAAGTTAGTGGATACACAG
ACCCATATTAGCAAATCAAATGATCTTTAATCTTTTAGCATGGTATACATATCTAA

P2-4

GGGGGGTGGGGAAGACGGCTCTAACACAGCTTGTGTACAATGATACAAGATAAAGGAGCATTTCAGCTAAGGGTCTGGCTGTGTGT
TTCTGAAAAATTTTGATCAGATGAAGCTTACCAAGGAAACCATTTGAATCAGTTGCCAGCGAGTTTGAATCAACCATCAGTTGGGTCTCA
TCCGTACGACCAACATGAACCTTGCTCCAAGAGACCTCTCAAAAAAGCTAAAAGGTAAGGTTTCTTCTAGTACTTGATGATGCAT
GGAATGAGGACCTGAAAAAGTGGGTACATATCGCCGTGCTCTACTTACTGGAGGAAAGGGAAGCAGAAATAGTAGTAACCAAGGAA
CAAAAAATGTGGGAACTAATGGGTGGGATGACTCCATATCTAAATCAGTTATCCGACAATGATTGCTGGTCTTTGTTTCCAGGAGC
TATGCATTTGTGGATGGTAACTCCAATGCACACCCAAATTTGGAATGATA

P2-5

GGGGGGTGGGGAAGACGACTTGACAAGGAAACCGCAAAGCTTTTGATAGTATGGTGATGCTTGTGTGCTGGCATTATGGAAGCAA
ATCAATGATCAGGTTTTCGCTCGGTGGCCATACCAATATTGGTTACAGATCTTGTAAATAGATTCTAGATGAGCTTCGCACTTGGG
CGATAGCGGGAGGACGTGGAGTTGAAATCATTTTGTGAGGAGTTAATGTAGCGTGTGAGTTGGAGTTTGCCTGCCCTCAGTCGGCA
TGTGGCTGTTGACTGTGTACAAAATATTCTTGTATATTGTTTCTCTCTTTTACTCCCTCCGTCCGAATTACTTGTGCGCAGGTATAG
ATATATCTAGAACTAAAGCGGAAAGAGTATAAGCTAATTTGCGTACCTCAAATAATAATAATAAAATATGCGTTTAGGAGTGCTTTT
GGAGATGCTGTTGGAACAAGCCAAGTCAA

P2-8

GGGGGGTGGGGAAGACGACGCTTGTTAGATATGTATACCATGCTAAAAGATTAAAGATCATTTTGATTGCTAATATGGGTCTGTG
TATCCACTAACTTCGATGTAGTGGGGCTAACACTTCAGATCCTAGATCATGTATGTGAAGATAGACCATATGAAAAAAAAGCAGTTT
GAATAAATTGCAGGAGATACTTCAAGAAAAATTAGAAACAAAGATTCTGCTTGTAAATGGATGATATGTGGGAAGAGAGGACAGG
GGTGGATGGATTAACTCTTGGCTCCATTGAAAAGTAATAAAGTAAAGGTTGCATGGTACTAGCAACAACAGAACAAAATCAGTGG
CAAAAATGATAGGAATATGGATGAAATCACATTGAGTGGTCTGGATGAGAGGAGTTTGGCTGTTTTTCAAGGCATGTGCATTTTCG
CAATGACAACCGGAGCACCATCTAGTTTGAATCTATTGGGAAACAAATTGCTAAAGCACTAAAAGGCTTCCCCCTCGCCCTCA

P3-1

GGGGGGTGGGGAAGACGCGAGCAAGGTTTTCTGGATGAGTGAGGAAGATTGGGGAGCGGGGGGAATTGGGGGGAGATGGACGCGGG
AGAGCGAGTTGTTTATGTTTATAAGGCCCGCTGCACTAACCAATCGCTTTTCCCAAGAACGGCTCTCTTCTTTTGGCTTGTGCTGCT
GGAGCTGGCCATAAATTTGGCTGACTGGCCATGCATGCAGGATGCATGTGACCGCGCGGCACATGCATGCGCTGTTTGTACTAGG
CCTATTTAATAGGGTGGTTATGGAGTAAATTAAGGAGATTTATTTTCTCTGTGTGCATCCACCAATTAGAAGATGCGGTACTAGATG
CAGACACTCACATTAAATAACATTCAAGAAACAAAGTTTCAAGATGCTATGCATCTCAATGATTCCCACTACTATAACCAATACA
AAGCTGTGAGAAGGTGTGGAATAAAGAAAAATCGGTGATGCTTCACTGAGCAAGGGCTTCCCCCTCGCCCTCA

P3-2

GGGGGGTGGGGAAGACGACATACGAACGAGAGATAGGCGATGGTGTAGAGTCAAGGAATATGGTGTAGGAGTGAGGATGGCGACCA
TGATCCGTGCAAGGGGCGGTGGCGGCGTCGGAAGGCGGCACATGATCTGAGGGATGCGATGACGGAGCTCGACTGTGCGCTCAAAGCA
GGGGCAACTGTTGCAAAAATGGATGAATGTGGGTGGGCTATGGTTTGGGGGGCGGTAGCACTAGATTTTATTACCGGAGACAGGCC
AATCGACCGGTGCACCGTCTTCTGCTTAATGCAAGAGGGAGTAGTCTCCCCATTTCTACTAAGAGAATGCCCACTACTAAAGAA
AAAATTGGCTA

P3-3

GGGGGGTGGGGAAGACGGCCGACTGTGGCATATTTCTTATATTTATATGTTTACCTTTTCGAGGCGTCTGCTTACAGACACCCGATC
GGTATTCATTTATGTTTGGTGGGATTTCAAGCGGCACACTGCGGTTTCCCTTTTACCTACCTCGTTATGCTAATTTTTGATTGTAC
ATTGGTGAATTATTCATGACGATCCATGCATGATTTGTTATATCTTACGTCCGAACTGTCTTGCAGTACTTTCAAAGTACTCACT
GGCTTGTGATTTGGCCAGATGCTGACGAAGGCTGATCTCATGGAT

P3-4

GGGGGGTGGGGAAGACGGCCCTTGTCAAGAAATGTGTATGATCGGGAGAAAGTCAACTTCGCTGATGCACATGCTTGGATTGTGGTAT
CAAAGGAATATGATGTAATTGATTGCTGGGACATTGCTCAAGATACGACACACACAAGAGTCAACGGGACCCCTCCTCTTAG
CGTGGGTGCCAAATCTGATGTGTATGACTTAACAGAAGCAATAAACAAGATAGTACAAGAGAAGAAGTGTGATTGTGCTTGTATGAT
GTGTGGAATAGAGATGCATATAATCAAAATGTGAATGCTTTCCAGGGTATCCATGGAAGCCGTGTTATGATCACAACCGGATGGAG
ATGTTGCGGCTCTTGCTCAACCGAAGCGTCGCTTGTACTCCAGCCATAGGGTAGCACTGAGTCATTCAAGCTATTTTGTCTCAAGGCT
TTTCCATAGCAACTTGAACCGTGAGTGCCCTCTAGAGCTCAATACAGTGGCTACAAAAGTTGTTGAGAGGTGTGCTGGCCTCCCTTC
GCCCTCA

P3-5

GGGGGGGTGGGGAAGACGGCCGACTGTGGCACCTTTTCTTTTATCTGTTACCTTTTCGAGGCGTCGCTTCAGACACCCGATCGGTTTT
 CATTATAGTTTTGTTGGGATTCAGGCGGACTACCGCCGTTTTCCCTTTTACTTACCTCATTATGCTAATTTTGAATATACATTGGT
 GAATTAATCATGACACATTCATGCATGCATTGTTATATCTTACATCCGAACGTGCTTGCGAGTACTTTCAAAGTACTCACTGGCTTG
 TTGATTTGGCCAGATGCTGACGAAGGCGATCTCATGGATGAGGAATTTGATAGCGAGTCCGACGCTAGAGGAGTCCAGTCAGTCTC
 GTATGACCTGGATTTGGTCACTGTATTATATCCGCTTCCGCTACCAAATAAATTCATCGAGCCTCACCTCGACGCTCGAAGAGATGA
 CAGTTTTAGAGACTTGATCTCTCTCCCGCCGCTTCTCCACC

P3-6

GGGGGGGTGGGGAAGACGGCCGATTGTGCCATCATTCTTATATTATTGTTACCTTTTCGAGGCGTCGCTTCAGACACCCGATCT
 GTAGTTATTATGTTTTGTTGGGATTCAGGCGGACTACCGCCGTTTTCCCTTTTACTTACCTTGTATGCTAATTTTGAATGTACAT
 TGGTGAATTATTCATGACGCATCCATGCATGAATTTGTTATATCTTATGTCCGAACGTGCTTGCGAGTACTTTCAAAGTACTCACTGG
 CTGTTGATTTGGCCAAATGCTGACGAAGGTGATCTCATGGATGAAGAGTTTGATAGCGAGTCCGACACCTAGAGGAGTCCAGTCAG
 TCTCGTGCGACCTGGATTGG

P4-2

GGGGGGGTGGGGAAGACGACTTGGTAAAGGTATCGATGAGAGGCCATGTAGGAGTACATGGTGGGTTGTTTCATTGAAGCCGACCTTA
 AGCACTGAGATCTGTATGTGATTTAAGAATCAGTACTACCATGCATTTGGGCCGAAACCAATGGACCTCTCGGCTTCTTATTCA
 CCCTAGTTCTCCGTCCAGGAGTGCAGTAGTTTCTGGTGTGTTAGCCTACTGGAGGCGGTGACACGCTGACCGTAGGGGTGGGC
 TGTGATGCGGTAGGTACGTGGCCGGGTGTACCGAATACCCGTTAGGTATCTCGGGGAACCTGTTACATCGTTTGGGGACGTATGGGA
 AACCTCGGCCGGACTCCCTGCGGATGGAACCTGGATAGGCGATAAACCTGGACTGGAGGCTTAGGTGATTAGGTAGGTGTTGGCCGAC
 ACCGCGTTGGGCTTCCCTTCGCCCTCA

P4-3

GGGGGGGTGGGGAAGACGACAACTTGGATCCAGTATTCGCTCCAGCTGGTGGCGCGTGAGTGGCGATCGCTGAGTCCAGGGTGGAC
 ATAATAAGGAGGCAACCTATGATCCACCTCAGGATCGACACCGGGGTGCCGTGCTGCTGGAGGACTGATACGCTCTCCACCGT
 ATCTATTTTCAAATACTTTTGGTCTGTTTGGACTCTAATTGCAACAATTTGTATGAACTAAATCGGACTCGTGTGTTATTAGC
 AGAACTAACATGGTGTGTT

P4-4

GGGGGGGTGGGGAAGACGGCGGGGAGAGAAATACAAGACTCTAAGTCTATCATCTCATCGAGCGTCGAGGTGAGGCTCGATGAATTTA
 TTTGGTAGCGGAAGCGGATATAAAACAGTGACCAAATCCAGGGTCGACGAGACTTACTGGAACCTCCTTAGGTGTCGGACTCGCTAT
 CAAACTCTTCATCCATGAGATCGCCTTCGTGAGCATCTGGCCAAATCAACAAGCCAGTGAGTACCTTGAAAGTACTCGAAGACTG

P4-5

GAGGGCGAAGGGGAAGCCAGATTGCGTCACAGAAGCTTTGGAGATTTAAAGGGAATAGAGGAATACACTGGATTTGATAGCATCACGA
 CCTTTCAAATTTGTCAATTCACATGGAGGGGAGCAATGGACTATATATAGAGACATGTATAACTCAAGTATGCCATACAATAG
 CTAATTTTGATGCACATCAATCTAAAATCTACTACAGATCTACAAATGAAATCGTACACTATGATAATTTATGAAATTTGCTAAA
 TATTAATGCCAAATGTGTAGGAAATATTGAAATAAATGCATCATACCTTTTCTTAATCCTGGAAGATAGAAATATATTTATGCCA
 TGCAACAATCTAAGATTCATTGTTTGTGCGGAAGAAACAACATACAATATCAATCTGTATAGCATAGACAATACTCCCTCCGTCCTT
 TAATTCATGTCGCGGCAGTGTTTTGCAATTTAGCCCTCGCCGTCTTCCCAACCCCCCA

P4-6

GGGGGGGTGGGGAAGACGGCGGGGAGAGATACAAGACTCTAAGTCTGTCTCATCTCATCGAGCGTCGAGGTGAGGCTCGATGAATTTA
 TTTGGTAGCGGAAGCGGATATAAAACAGTGACCAAATCCAGGGTCGACGAGACTGACTGGGACTCCTCTAAGTGTTCGGACTCGCTA
 TCAAACCTCTTCATCCATGAGATCGCCTTCGTGAGCATCTGGCCAAATCAACAAGCCAGTGAGTACTTTGAAAGA

P5-1

GGGGGGGTGGGGAAGACGACGCTTCTACATGTTTTCAACAATGATCTTGAAAGAAGGCCCATGATTATCAGGTACTATGTTGCTTAC
 TCCAAGACTGCAATTTATGCCTCAGAGTACATGTTGTCTGTACTTTAATATTGCTATTTAACTTACGAGAATACTTATTTCAAGTGAC
 CACGCTTATATTGTAAAGATGCTCAAGAAATTTATGTATAAACAAGAGAGAAAATTTGTGCTGTATGCATTTTAAAGATGGCA
 GCAAGTTTCATTGTAGAACAGCTACCCAGTGCATGAAGTCTGAAGAAGTAATGTTGATGTTCTGTACAGGCATATGAAATTTAACA
 CGGCTTTCACACGGAAGGCAGTTAATACCATTCACCATGACAATTTGACCTGTTACAGAATATAGGTACACATTAGCCGGTAGAT
 ACGGACATGAAATGCAATGCTTCTTGATGCAATTAGTGTGTGTCATTTTGTCTGTTTTGCCAAACCTCCTTAAAGGTTTTCTTTTT
 TAAGCATATTTTGCCTAAACAGTAAATATGCATAAATGCAGTGCCTAAATGT

P5-2

GGGGGGGTGGGGAAGACGGCGCTTCTACATGTTTTCAACAATGATCTTGAAAGAAGGCCCATGATTATCAGGTACTATGTTGCTTAC
 TTCAAGACTGCAATTTATGCCTCAGAGTACATGTTGTCTGTACTTTAATATTGCTATTTAACTTACGAGAATACTTATTTCAAGTGAC
 CACGCTTATATTGTAAAGATGCTCAAGAAATTTATGTATAAACAAGAGAGAAAATTTGTGCTGTATGCATTTTAAAGATGGCA
 GCAAGTTTCATTGTAGAACAGCTACCCAGTGCATGAAGTCTGAAGAAGTAATGTTGATGTTCTGTACAGGCATATGAAATTTAACA
 CGGCTTTCACACGGAAGGCAGTTAATACCATTCACCATGACAATTTGACCTGTTACAGAATATAGGTACACATA

P6-1

GGGGGGGTGGGGAAGACGACACCCGTGGAATACCGTGCGTGAGGCGCAGGAGTATAGAGGTGTTACATGCTAGACCTATGTGGCAT
 CGAGTCGGGGTCTGACAGAGCGTTTCGTGATGGTCATCTTCGCCGTATGCTAGGGAGGCGTGGCGGCCCTCCACACTAGTTTC
 GCCTCTCATTCTCAGGCAGTGCACACTCTATACGACCGAGGTGGGTGAGACCAGCTTGGAGATCTCAGCATCATAGAGTACTTCA
 ACAAGATGACGAGCCTCGCCGACACCTTGGCCTCTATTGGTCAGCCGCTTTAGAAGGAGT

P6-4

GGGGGGGTGGGGAAGACGACTTTCACCCAACACCTGTATAATGATGGTAGGACTGAAGAGCACTTTGCTGTTATGGTTTGGGTGTGTG
TATCAACTGATTTTGATGTGTTGAAACTCACCCAACAGATTCTTAAGTGCATACCTGCAACTGACAATGAAATTGCAATTGAAACAAC
CAATCTAGACAGACTTCAAAAATCCATTGCGCAGAGACTGAAGACCAAAAGGTTCTTAATTGTCTTGGATGATATATGGAAATGCGAT
AGTGAGGATGAGTGGAAAACCTTGCTAGCTCCATTACAAAAGGGGACGCCAAAGGCAGCATGATACTTGTCACAACCTCGATTCCCAA
AACTAGCAAGAATGATGAAAACAGTTGATTAGTAGAAGTGAAGGCTTGGAGTCTAATGACTTCTTACATTCTTTGAAGCATGTAT
ATTTGGTGAACATAAGCCAGGGATTATGAAGATGACTTAGCTTGTATTGCAAGAGAATTGCAAGCAAGTTAAAGGGCTTCCCC

P6-5

GGGGGGGTGGGGAAGACGACTTTCACCCAACACCTGTATAATGATGGTAGGACTGAAGAGCACTTTGCTGTTATGGTTTGGGTGTGTG
TATCAACTGATTTTGATGTGTTGAAACTCACCCAACAGATTCTTAAGTGCATACCTGCAACTGACAATGAAATTGCAATTGAAACAAC
CAATCTAGACAGACTTCAAAAATCCATTGCGCAGAGGCTGAAGACCAAAAGGTTCTTAATTGTCTTGGATGATATATGGAAATGCGAT
AGTGAGGATGAGTGGAAAACCTTGCTAGCTCCATTACAAAAGGGGACGCCAAAGGCAGCATGATACTTGTCACAACCTCGATTCCCAA
AACTAGCAAGAATGATGAAAACAGTTGATTAGTAGAAGTGAAGGCTTGGAGTCTAATGACTTCTTACATTCTTTGAAGCATGTAT
ATTTGGTGAACATAAGCCAGGGATTATGAAGATGACTTAGCTTGTATTGCAAGAGAATTGCAAGCAAGTTAAAGGGCTTCCCCCTC
GCCTTCA

P8-1

GGGGTGGGGAAGACGACGGGGAGAGAGATACAAGACTCTAAGTCTGTCATCTCATCGAGCGTCGAGGTGAGGCTTGATGAATTTATTT
GGTAGCGGAAGCGGATATAAATAGTGACCAATCCAGGGTCGACGAGACTGACTGGGACTCCTCTAGGTGTCGGACTCGCTATCAA
ACTCTTCATCCATGAGATCGCCTTCGTCAGCATCTGGCCAAATCAACAAGCCAGTGAGTACTTTGAAAGTACTCGCAAGACAGTTGCGG
ACGTAAGATATAACAACGCATGCATGAATGCGTCATGAATAATCCACCAATGTACATTCAAAAATAGCACACAAGGTAAGTAAAA
GGGAAACGGCGGTAGTCCGCCTGAAATCCCAACAAAACATAAATGAAGACCGATCGGGTGTCTGAAGCGATGCCCCGAAAGGTGAACA
AATAAATATAAGGAATGATGCCACAGTCGGCCGTCTTCCCCACCCCCCA

P8-2

GAAGGCGAAGGGGAAGCCACGATGGAACAAGAACAATGGCGCCACAACAATGGCGAAGCTAACGAGATGGCGGTAATGCCGAATTT
TGTAACCCGTAGTCCGGGCCATGGAAAAGCCCTCTAAGGCAACCAAGAGGGCCATCCAAGTTGGATGCAATCCTAGATAGGCCTT
GCCAGATCCACGACACCTCTGAGAAGCCCGCAACACACCCGACCACTAAGTGTGGGTGATAAAACAAGCGGGCAAGCTGACAGCCGA
AGACCCCTAGCAAAAGCCCTCAAAGCGAGGACGATTAGGAGCCCTGCAGGCCGAACAGTGAAGCCAAAATCCCTTCTCGCGAGGTC
AAGACGGTGAACACGATCTACGCCACCCATATACCAAGAAAGGAGGGAAGCGCATGCTATGAAATGTTTACGCCGTGAGCCCGTAA
TGCCAAAGTTCTACCCATGGTCGTCTCTCCCAACCCCCCA

P8-3

GGGGGGGTGGGGAAGACGCGGGGAGAGAAATACAAGACTCCAGGTCCGTCATCTCATCGAGCGTCGAGGTGAGACTCGATGAATTTA
TTTGGTAGCGTAAGCGGATATAAATAGTGACCAATCCAGGGTCGACGAGACTGACTGGGACTCCTCTAGGTGTGCGACTCGCTAT
CAAACCTTTCATCCATGAGATCGCCTTCGTCAGCATCTCGCCAAATCAACAAGCTAGTGAGTACTTTGAAAG

P8-4

GGGGGGATGGGGAAGACGGCCATTACTAAGTGGGGGGTGGGGAAGACGGCCATTACTAAGTGGGGGGTGGGGAAGACGGCGCTTGT
TAGATATGTATACCATGCTAAAAGAATTAAAGATCATTTTGATTTGCTAATATGGGCTGTGTATCCACTAACTTCGATGTAGTGGGG
CTAACACTTCAGATCCTAGATCATGTACGTGAAGATAGACCATATGAAAAAAGCAGTTTGAATAAATGCAGGAGATACTTCAAGA
AAATATTAGAAACAAAAGATTTCTGCTTGTAAATGGATGATATGTGGGAAGAGAAGGACAGGGGTGGATGGATTAAACTCTTGGCTCCA
TTGAAAGTAATAAAGTAAAGGTTGCATGGTACTAGCAACAACCAGAACAAAATCAGTGGCAAAAATGATAGGAACATATGGATGAAA
TCACATTGAGTGGTCTGGATGAGAAGGAGTTTGGCTGTTTTTCAAGGCATGTGCAATTCGCAATGACAACCGCAGC

P8-5

GGGGGGGTGGGGAAGACGGCCGACTATGGCATCATTTCTTTTATTTAGTCATTACCTTTTCGAGGCGTCGCTTCAGACACCCGATCG
GTATTCATTTATGTTTGTGGGATTTTCAGGCGGACTACCGCGTTTCCCTTTTACTTACCTTGTATGCTAATTTTGAATGTACAT
TGGTGAATTAATCATGACGCATCCATGCATGCATTTGTTATATCTTACGTCCGAAGTGTCTTGGCAGTACTTTCAAAGTACTCACTGG
CTTGTGACTTGGCCAGATGCTGACGAAGGCGATCTCATGGATGAAGAATTTGATAGCGAGTCCGACACCTAGAGGAGTCCCAGTCAG
TCTCGTGCACCCCTGGATCTGGTCACTGTTTTATATCCGCTTCCGCTACCAAAATAAATTCATCGAGCCTCACCTCGACTCTCGATGAG
ATGACAGACTTAGAGTCTTGTATTCTCTCCC

APPENDIX F

CLUSTAL X (1.81) multiple sequence alignment of all wheat RGA sequences

```

P1-4 -----GAG-----
P1-5 -----GAG-----
A1-1 -----G-----
P1-2 -----GAG-----
A1-5 -----GAG-----
CS2-2 -----GAG-----
P1-3 -----GAG-----
A1-6 -----G-----
A3-4 -----AGG-----
P1-1 -----G-----
P1-1 -----G-----
A3-4 -----G-----
A1-1 -----G-----
A1-3 -----G-----
A3-5 -----G-----
A6-6 -----G-----
P1-6 -----G-----
P1-1 -----G-----
A1-6 -----G-----
P1-3 -----G-----
AC-6 -----G-----
A1-5 -----G-----
A8-5 -----G-----
CS6-4 -----G-----
AC-1 -----G-----
P1-1 -----G-----
CS2-4 -----G-----
CS6-1 -----G-----
A6-1 -----G-----
A6-2 -----G-----
P1-5 -----G-----
CS2-5 -----G-----
AC-3 -----G-----
CS2-3 -----G-----
FC_1-9_2AB -----G-----
A6-4 -----G-----
P1-4 -----GGGGGATGGGGGAAGCGCCATTACTTAAGTGGGGGGTGGGGGAAGACGCCATTACTTA-----
A1-1 -----G-----
A1-2 -----G-----
A1-2 -----G-----
A1-2 -----G-----
FC-4 -----G-----
CS2-1 -----G-----
A5-1 -----G-----
P5-1 -----G-----
CS8-1 -----G-----
P1-2 -----G-----
A1-2 -----G-----
CS8-1 -----G-----
P1-2 -----G-----
A8-4 -----G-----
A3-3 -----G-----
P1-5 -----G-----
P3-6 -----G-----
A1-4 -----G-----
P3-5 -----G-----
P3-3 -----G-----
TD3-3 -----G-----
TD3-1 -----G-----
A8-2 -----G-----
A8-3 -----G-----
A3-1 -----G-----
P3-4 -----G-----
CS2-6 -----G-----
FC-3 -----G-----
AC-4 -----G-----
TD3-2 -----G-----
TD3-1 -----G-----
TD3-3 -----G-----
P3-2 -----G-----
P1-3 -----G-----
TD3-2 -----G-----
P1-1 -----G-----
A3-6 -----G-----
P1-5 -----G-----
P6-1 -----G-----
TD3-12 -----G-----
A5-2 -----G-----
P2-2 -----G-----
P2-5 -----G-----
P3-1 -----G-----
TD3-10 -----G-----
P1-2 -----G-----
TD3-16 -----G-----
P2-1 -----G-----

```

P1-4 GCGAGGGGGAGGCC--TCGAC--AACTTTGGGCTATTTCATGAGGGTCTCTATGATATCA
P1-6 GCGAGGGGGAGGCC--TCGAC--AACTTTGGGCTATTTCATGAGGGTCTCTATGATATCA
AL-1 GCGAGGGGGAGGCC--TCGAC--AACTTTGGGCTATTTCATGAGGGTCTCTATGATATCA
P1-2 --GGGGGGGCC--AACTTTGGGCTATTTCATGAGGGTCTCTATGATATCA
P1-3 --GGGGGGGCC--AACTTTGGGCTATTTCATGAGGGTCTCTATGATATCA
C95-2 GTGAGGGGGAGGCC--TCGAC--AACTTTGGGCTATTTCATGAGGGTCTCTATGATATCA
P1-5 GTGAGGGGGAGGCC--TCGAC--AACTTTGGGCTATTTCATGAGGGTCTCTATGATATCA
AL-6 GCGAGGGGGAGGCC--TCGAC--AACTTTGGGCTATTTCATGAGGGTCTCTATGATATCA
A8-1 GCGAGGGGGAGGCC--TCGAC--AACTTTGGGCTATTTCATGAGGGTCTCTATGATATCA
A8-1 --GGGGGGGTGGGG--AGAGC--GGCGGGGAGAGKATACACAGACTTAAGTGTCTATCATCT
P6-1 --GGGGGTGGGG--AGAGC--GGCGGGGAGAGATACACAGACTTAAGTGTCTATCATCT
A3-4 --GGGGGGGTGGGG--AGAGC--GGCTGGGAGAGATACACAGACTTAAGTGTCTATCATCT
P6-1 --GGGGGGGTGGGG--AGAGC--GGCTGGGAGAGATACACAGACTTAAGTGTCTATCATCT
A4-3 --GGGGGGGTGGGG--AGAGC--GGCGGGGAGAGATACACAGACTTAAGTGTCTATCATCT
A3-5 --GGGGGGGTGGGG--AGAGC--GGCGGGGAGAGATACACAGACTTAAGTGTCTATCATCT
A6-6 --GGGGGGGTGGGG--AGAGC--GGCGGGGAGAGKATACACAGACTTAAGTGTCTATCATCT
P6-1 --GGGGGGGTGGGG--AGAGC--GGCGGGGAGAGATACACAGACTTAAGTGTCTATCATCT
P4-4 --GGGGGGGTGGGG--AGAGC--GGCGGGGAGAGATACACAGACTTAAGTGTCTATCATCT
A4-6 --GGGGGGGTGGGG--AGAGC--GGCGGGGAGAGATACACAGACTTAAGTGTCTATCATCT
P8-3 --GGGGGGGTGGGG--AGAGC--GGCGGGGAGAGATACACAGACTTAAGTGTCTATCATCT
P6-1 --GGGGGGGTGGGG--AGAGC--GGCGGGGAGAGATACACAGACTTAAGTGTCTATCATCT
A4-5 --GGGGGGGTGGGG--AGAGC--GGCGGGGAGAGATACACAGACTTAAGTGTCTATCATCT
A8-5 --GGGGGGGTGGGG--AGAGC--GGCGGGGAGAGATACACAGACTTAAGTGTCTATCATCT
C92-4 --GGGGGGGTGGGG--AGAGC--GGCGGGGAGAGATACACAGACTTAAGTGTCTATCATCT
A2-1 --GGGGGGGTGGGG--AGAGC--GGCGGGGAGAGATACACAGACTTAAGTGTCTATCATCT
P6-4 --GGGGGGGTGGGG--AGAGC--GGCTTTCACCAACAGCTGTATTAATGATGTAGAGTGT
C92-4 --GGGGGGGTGGGG--AGAGC--GACTTTTCAOCCACAGCTGTATTAATGATGTAGAGTGT
C92-6 --GGGGGGGTGGGG--AGAGC--GGCTTTTCAOCCACAGCTGTATTAATGATGTAGAGTGT
P6-1 --GGGGGGGTGGGG--AGAGC--GGCTTTTCAOCCACAGCTGTATTAATGATGTAGAGTGT
A6-2 --GGGGGGGTGGGG--AGAGC--GACTTTTCAOCCACAGCTGTATTAATGATGTAGAGTGT
P6-5 --GGGGGGGTGGGG--AGAGC--GACTTTTCAOCCACAGCTGTATTAATGATGTAGAGTGT
C92-5 --GGGGGGGTGGGG--AGAGC--GGCACTTCAOCCACAGCTGTATTAATGATGTAGAGTGT
P6-1 --GGGGGGGTGGGG--AGAGC--GGCACTTCAOCCACAGCTGTATTAATGATGTAGAGTGT
P6-1 --GGGGGGGTGGGG--AGAGC--GGCACTTCAOCCACAGCTGTATTAATGATGTAGAGTGT
P1_1-5_2AB --GGGGGGGTGGGG--AGAGC--GGCACTTCAOCCACAGCTGTATTAATGATGTAGAGTGT
A6-4 --GGGGGGGTGGGG--AGAGC--GGCACTTCAOCCACAGCTGTATTAATGATGTAGAGTGT
P6-1 --GGGGGGGTGGGG--AGAGC--GGCACTTCAOCCACAGCTGTATTAATGATGTAGAGTGT
A3-2 --GGGGGGGTGGGG--AGAGC--GACCTCTTGTAGATATGATACACAGCTTAAGAGTGT
A4-2 --GGGGGGGTGGGG--AGAGC--GGCCCTCTGTAGATATGATACACAGCTTAAGAGTGT
A2-2 --GGGGGGGTGGGG--AGAGC--GGCCCTCTGTAGATATGATACACAGCTTAAGAGTGT
P6-1 --GGGGGGGTGGGG--AGAGC--GGCCCTCTGTAGATATGATACACAGCTTAAGAGTGT
C92-1 --GGGGGGGTGGGG--AGAGC--GGCCCTCTGTAGATATGATACACAGCTTAAGAGTGT
A5-1 --GGGGGGGTGGGG--AGAGC--GAGCCTCTGTAGATATGATACACAGCTTAAGAGTGT
P5-1 --GGGGGGGTGGGG--AGAGC--GAGCCTCTGTGTGTAGATATGATACACAGCTTAAGAGTGT
P6-1 --GGGGGGGTGGGG--AGAGC--GAGCCTCTGTGTGTAGATATGATACACAGCTTAAGAGTGT
P5-2 --GGGGGGGTGGGG--AGAGC--GAGCCTCTGTGTGTAGATATGATACACAGCTTAAGAGTGT
AL-2 --GGGGGGGTGGGG--AGAGC--GAGCCTCTGTGTGTAGATATGATACACAGCTTAAGAGTGT
C98-1 --GAGGCGCAGAGG--GAGGC--CAGCATGTCT--AATGAAGCA--ATATGGCGCCACACATGT
A8-4 --GAGGCGCAGAGG--GAGGC--CAGCATGTCT--AATGAAGCA--ATATGGCGCCACACATGT
A3-3 --GAGGCGCAGAGG--GAGGC--CAGCATGTCT--AATGAAGCA--ATATGGCGCCACACATGT
P8-5 --GAGGCGCAGAGG--GAGGC--CAGCATGTCT--AATGAAGCA--ATATGGCGCCACACATGT
A4-4 --GAGGCGCAGAGG--GAGGC--CAGCATGTCT--AATGAAGCA--ATATGGCGCCACACATGT
P3-5 --GAGGCGCAGAGG--GAGGC--CAGCATGTCT--AATGAAGCA--ATATGGCGCCACACATGT
P3-3 --GAGGCGCAGAGG--GAGGC--CAGCATGTCT--AATGAAGCA--ATATGGCGCCACACATGT
TH3-1 --GAGGCGCAGAGG--GAGGC--CAGCATGTCT--AATGAAGCA--ATATGGCGCCACACATGT
A8-2 --GAGGCGCAGAGG--GAGGC--CAGCATGTCT--AATGAAGCA--ATATGGCGCCACACATGT
A8-3 --GAGGCGCAGAGG--GAGGC--CAGCATGTCT--AATGAAGCA--ATATGGCGCCACACATGT
P3-4 --GAGGCGCAGAGG--GAGGC--CAGCATGTCT--AATGAAGCA--ATATGGCGCCACACATGT
C92-6 GCGCAGCGGGGAGGCGCTTT--AGTGGCTTATGACATTTTGTTCGCAATATATGGAACATGT
P2-3 GCGCAGCGGGGAGGCGCTTT--AGTGGCTTATGACATTTTGTTCGCAATATATGGAACATGT
A8-4 GCGCAGCGGGGAGGCGCTTT--AGTGGCTTATGACATTTTGTTCGCAATATATGGAACATGT
T93-2 --GGGGGGGTGGGG--AGAGC--GGCC--GACTTGTGGATTTATGATGTAGTGTGTGT
T93-1 --GGGGGGGTGGGG--AGAGC--GGCC--GACTTGTGGATTTATGATGTAGTGTGTGT
T93-3 --GGGGGGGTGGGG--AGAGC--GGCC--GACTTGTGGATTTATGATGTAGTGTGTGT
T93-4 --GGGGGGGTGGGG--AGAGC--GGCC--GACTTGTGGATTTATGATGTAGTGTGTGT
T93-2 --GGGGGGGTGGGG--AGAGC--GGCC--GACTTGTGGATTTATGATGTAGTGTGTGT
P1-1 --GGGGGGGTGGGG--AGAGC--GGCC--GACTTGTGGATTTATGATGTAGTGTGTGT
P4-5 --GGGGGGGTGGGG--AGAGC--GGCC--GACTTGTGGATTTATGATGTAGTGTGTGT
P93-12 --GGGGGGGTGGGG--AGAGC--GGCC--GACTTGTGGATTTATGATGTAGTGTGTGT
P2-2 --GGGGGGGTGGGG--AGAGC--GGCC--GACTTGTGGATTTATGATGTAGTGTGTGT
P2-5 --GGGGGGGTGGGG--AGAGC--GGCC--GACTTGTGGATTTATGATGTAGTGTGTGT
P3-1 --GGGGGGGTGGGG--AGAGC--GGCC--GACTTGTGGATTTATGATGTAGTGTGTGT
P3-10 --GGGGGGGTGGGG--AGAGC--GGCC--GACTTGTGGATTTATGATGTAGTGTGTGT
P4-2 --GGGGGGGTGGGG--AGAGC--GGCC--GACTTGTGGATTTATGATGTAGTGTGTGT
T93-16 --GGGGGGGTGGGG--AGAGC--GGCC--GACTTGTGGATTTATGATGTAGTGTGTGT
P2-1 --GGGGGGGTGGGG--AGAGC--GGCC--GACTTGTGGATTTATGATGTAGTGTGTGT

[illegible]

CURRICULUM VITAE

NAME: ERDOGAN E., HAKKI
DEPARTMENT: BIOTECHNOLOGY

Nationality: Turkish
Foreign languages: English and Greek
Ph.D. Thesis subject: A novel method for developing microsatellite markers and sequence analysis of wheat disease resistance gene analogs¹
M.S. Thesis subject: Partial isolation of the gene coding for lactate dehydrogenase from *Rhizopus oryzae* using RT-PCR²
Supervisors: Dr. Mahinur S. Akkaya, METU Chemistry Dpt^{1,2}
Dr. Aykut Kence, METU Biology Dpt¹
Dr. Ufuk Bakir, METU Food Eng. Dpt²
Address: Middle East Technical University
Department of Chemistry / 06531 / Ankara /TURKEY
Tel: (+90) 312 210 31 96-7 **Fax:** (+90) 312 210 12 80
E-mail: eehakki@metu.edu.tr / eehakki@yahoo.com

Education:		
Program	Years	CUM GPA
Biotechnology Graduate Prog.-Ph.D.	1997 - 2000	3.64/4.00
Biotechnology Graduate Prog.-M.S.	1994 - 1997	3.54/4.00
Ankara Univ.Fac.Sc.Biology Dept.B.S	1990 - 1994	72.49/100

Scholarship and Supports:

- Ministry of Education Student Support, 1988-89
- Middle East Technical University, Graduate School of Natural and Applied Sciences 1996-97
- Turkish Scientific and Research Council Support:
"The 8th European Congress on Biotechnology, Budapest, Hungary, 1997"
- Turkish Scientific and Research Council, NATO-A2:
"The University of Michigan Biology Department, 1999-2000"
- GAMA Scholarship, 2000

Project Supports:

- TÜBİTAK "Çiftleştirilmiş (Doubled) Haploid Buğdayı Çeşitlerinin DNA Moleküller Belirleyicileri Kullanılarak Genotip Karakterizasyonu (TBAG-1720, 1998- 1999)
- DPT "Türkiye Buğday Genotiplerinin DNA Profillerinin oluşturulması ve mikrosatelit bağlantı haritalarının başlatılması" (AFP-01- 03-DPT-97K12200, 1997-1998 ve AFP-06 02-DPT-98K12280, 1998-1999)
- AFP Ortak proje "Çiftleştirilmiş Haploid Türk Durum Buğdaylarının Mikrosatelitlerle Karakterizasyonu" (AFP-98-06-02-06, 1998-1999)
- AFP "Buğdayda Sari Pas Hastalığına Dirençlilik Gösteren Çeşitlerden Dirençlilik Geni Analoglarının Arastırılması (AFP-2000-07-02-00-11, 2000-2001)

Laboratory Experience :

- Growth of fungi, yeast, bacteria and plants
- cDNA synthesis and RT-PCR
- DNA and RNA isolation techniques
- Subcloning into TA Vector, pBS, pBR322
- Southern and Northern blotting
- Transformation of bacteria
- DNA sequencing
- Construction of DNA libraries
- Development of novel SSR markers
- Application of SSR markers
- Application of AFLP markers

Publications:

- E. E. Hakki and M. S. Akkaya "Microsatellite isolation using amplified fragment length polymorphism markers: no cloning, no screening" Accepted for publication, *Molecular Ecology*, 9 (12) (2000) 2152-2154.
- Erdogan E. Hakki and Mahinur S. Akkaya "RT-PCR amplification of a *Rhizopus oryzae* lactate dehydrogenase gene fragment" Accepted to be published in *Enzyme and Microbial Technology*, 28 (2000) 000-000.
- Erdogan E. Hakki, Cigdem Savaskan & Mahinur S. Akkaya "Genotyping of Anatolian doubled-haploid durum lines with SSR markers" Accepted to be published *Euphytica*,
- E. E. Hakki and M. S. Akkaya "Cloning and characterization of resistance gene analogs (RGAs), targeting NBS-LRR class, in wheat". Manuscript in preparation.
- SS Hakki, EE Hakki, MS Akkaya, and R Nohutcu "The effect of basic-Fibroblast Growth Factor (FGF) and Dexamethasone on the Periodontal Ligament Cells, *in vitro*. International and American Association for Dental Research, April 5-8, Washington DC, USA, 78th General Session of IADR, *Journal of Dental Research*, 79 (IADR Abstracts), Special Issue, 2000, P2065

Meeting Abstracts:

- Mahinur S. Akkaya, Erdogan E. Hakki, Ufuk Bakir, "Isolation of a fragment of lactate dehydrogenase gene from *Rhizopus oryzae* via RT-PCR" *Turkish Biochemical Society, XIII. International Biochemistry Congress*, Antalya, Turkey, C-109, 26-30 March, 1996.
- Erdogan E. Hakki, Ufuk Bakir, Mahinur S. Akkaya, "RT-PCR amplification of a fragment of Lactate Dehydrogenase gene of *Rhizopus. oryzae*" *NATO-ASI Molecular Microbiology Meeting*, Birmingham UK. April, 1997
- Erdogan E. Hakki, Mahinur S. Akkaya, Ufuk Bakir, "Amplification of a *ldh* gene fragment of *R. oryzae* via RT-PCR" *8th European Congress on Biotechnology*, Budapest, Hungary 17-21 August, TU1114, 1997
- Erdogan E. Hakki, "Amplification and isolation of a fragment of lactate dehydrogenase gene from *Rhizopus oryzae* using RT-PCR" Thesis submitted to the Graduate School of Natural and Applied Sciences of the Middle East Technical University, Department of Biotechnology, September, 1997
- Erdogan E. Hakki, C. Savaskan, M. S. Akkaya "Analysis of Doubled-Haploid Turkish Durum Wheat Varieties Using Simple sequence Repeat (SSR) DNA Markers" *25th Silver Jubilee FEBS Meeting*, Copenhagen, Denmark, July 5-10, 1998, P27-12.
- Mahinur S. Akkaya, Nurdan Dograr, Aytul Incirli, Erdogan E. Hakki, ve Hatice Bilgic "Türk Durum Bugday Cesitlerinin Genetik Iliski Calismalarinda Kullanilan Molekuler Belirleyiciler" *Tarimda Biyoteknoloji*, Bornova-Izmir, 12-16 Nisan, 1999.
- Mahinur S. Akkaya, Nurdan Dograr, Aytul Incirli, Erdogan E. Hakki, E. Banu Buyukunal and Hatice Bilgic "Molecular Markers in the Genetic Relationship Studies of Turkish Durum Wheat Varieties" *The First International Biosciences Days*, April 20-24, Antalya, 1999, L10.
- Erdogan E. Hakki, Nurdan Dograr, Cigdem Savaskan and Mahinur S. Akkaya "Genotype comparisons of Doubled Haploid Durum Wheat Lines and the Originator Varieties" *The First International Biosciences Days*, April 20-24, Antalya, 1999, P222.

- Mahinur S. Akkaya, Nurdan Dograr, Aytul Incirli, **Erdogan E. Hakki**, E. Banu Buyukunal ve Hatice Bilgic "Turkiye Durum Bugday Kultivarlarında DNA Belirleyicileri Kullanılarak Genetik Cesitliliğin Çalisilmasi" *Orta Anadolu Hububat Tariminin Sorunlari ve Çozum Yollari Sempozyumu*, 8-11 Haziran, Konya, 1999, P26.
- **Erdogan E. Hakki**, Mahinur S. Akkaya "Yeni Basit Dizilim Tekrar (BDT) DNA Belirleyicilerinin Uretilmesi" *Orta Anadolu Hububat Tariminin Sorunlari ve Çozum Yollari Sempozyumu*, 8-11 Haziran, Konya, 1999, P135.
- **Erdogan E. Hakki**, Mahinur S. Akkaya. Buğdayda hastalığa dirençlilik genlerinin araştırılması. XVth National Biology Congress (with international attendance), Page:30, September 5-9, 2000, Ankara, Turkey.
- **Erdogan E. Hakki** ve Mahinur S. Akkaya. Yeni moleküler belirleyicilerinin geliştirilmesi ve Triticum çeşitlerinin analizinde kullanımı. XVth National Biology Congress (with international attendance), Page:25, September 5-9, 2000, Ankara, Turkey.

