Chapter 1

Introduction

Plant breeders have tried to improve yield quantities in crops, like wheat, from the earliest times. In recent years, breeding for resistance against disease proved to be as important for total wheat production, as breeding for an increase in yield potential (Byerlee & Moya, 1993). The use of single resistance genes against various pests and diseases plays a major role in resistance breeding of cultivated crops. The earliest single resistance gene was identified effective against yellow rust (Biffen, 1905). Numerous single genes for leaf rust resistance have since been identified, the 47th gene being recently discovered by Dubcovsky *et al.* (1998). The leaf rust resistance genes prevent crop losses due to *Puccinia recondita* Rob. Ex Desm. f. sp. *tritici* infections, which can range from 5-15 % depending on the stage of crop development (Chester, 1946).

Leaf rust resistance gene *Lr37* is an effective adult-plant resistance gene that increases resistance of plants against *P. recondita* f. sp. *Tritici* (UVPrt2 or UVPrt13) infections, especially when combined with genes *Lr13* and gene *Lr34* (Kloppers & Pretorius, 1997). *Lr37* originates from the French cultivar VPM1 (Dyck & Lukow, 1988). The line RL6081, developed in Canada for *Lr37* resistance, showed seedling and adult-plant resistance to leaf, yellow and stem rust (Dyck & Lukow, 1988). Crosses between the French cultivar VPM1 (Doussinault *et al.*, 1981) or the Canadian line RL6081 and South African cultivars will therefore introduce this gene into local germplasm. Not only will the gene be introduced, but the genetic variation of South African cultivars will also increase.

Molecular techniques have been used to estimate genetic distances among different wheat cultivars (Yu *et al.*, 1993; Nybom, 1994; Myburg *et al.*, 1997; Bohn *et al.*, 1999). With the genetic distances known predictions can be made for the best combinations concerning the two foreign genotypes carrying gene *Lr37*, VPM1 and RL6081 and local South African cultivars. This is especially important in wheat with its low genetic variation (Gale *et al.*, 1990; Devos & Gale, 1992). The gene will also be transferred with the least amount of backcrosses to cultivars genetically closest to each other, generating similar genetic offspring to the recurrent parent, but with gene *Lr37*. Genetic distances between near isogenic lines (NILs) for a particular gene will also give an indication of how many loci, amplified with molecular techniques, need to be compared in order to locate putative markers linked to the gene.

Molecular markers linked to resistant genes have numerous advantages over morphological markers. Genes have to be marked with molecular markers to enable breeders to combine the genes in gene pyramids (Kelly, 1995). Gene pyramids normally provides durable resistance and the pathogen has to mutate for virulence, simultaneously against more than one gene in order to be virulent (Schafer *et al.*, 1963). Important resistance genes can also be introduced quicker into new cultivars and the genes can be identified without any pathogenic tests and phenotypic confirmation (Winter & Kahl, 1995; Kelly, 1995), which will limit the spread of pathogens to nearby wheat fields.

Molecular techniques were used successfully to locate DNA markers for important leaf rust resistance genes with both the near isogenic line (NILs) (Weeden *et al.*, 1994) and bulk segregant analysis (Michelmore *et al.*, 1991) approaches. Markers were found for genes *Lr9* (Schachermayr *et al.*, 1994), *Lr24* (Schachermayr *et al.*, 1995; Dedryver *et al.*, 1996), *Lr25* (Procunier *et al.*, 1995), *Lr28* (Gupta *et al.*, 1998; Naik *et al.*, 1998), *Lr29* (Procunier *et al.*, 1995), *Lr32* (Autrique *et al.*, 1995), *Lr35* (Seyfarth *et al.*, 1999), *Lr37* (Troskie *et al.*, 1999), *Lr41* (Lottering *et al.*, 1999) and *Lr47* (Dubcovsky *et al.*, 1998). It will be an important advantage for plant breeders if gene *Lr37* can be detected with DNA markers, mainly because the gene can successfully be combined with other genes to increase resistance against leaf rust and therefore obtain durable resistance. The first aim of the study was to calculate genetic distances among foreign *Lr37* carrying genotypes and South African cultivars, which can be used to predict the best combining options among cultivars and to estimate the number of compared loci necessary to find markers for gene *Lr37*, when using NILs. The second and main objective of the study, was to identify and characterise DNA markers linked to gene *Lr37*, which can be used to identify the gene, especially where gene combinations, or pyramids are made up to obtain durable resistance against leaf rust.

Chapter 2

Molecular markers to facilitate leaf rust resistance, breeding programs in wheat

A literature review

2.1 Introduction

Wheat is one of the most important food crops in the world (Röder *et al.*, 1998). The global demand for wheat, which will increase 40 % by the year 2020 is, however, alarming (Pinstrup-Anderson & Pandya-Lorch, 1997). A way to meet these demands will be to increase wheat production in marginal producing areas, which consist of 45 % of the wheat producing area worldwide (Reeves *et al.*, 1999). Breeding for resistance against a disease like leaf rust is an economic and effective way to meet these demands (Beharav *et al.*, 1997).

Leaf rust in wheat is caused by the fungus *Puccinia recondita* Rob. Ex Desm. f. sp. *tritici* (Roelfs *et al.*, 1992). Yield reductions of 5-15 % can be expected, depending on the stage of crop development with initial infection (Chester, 1946). Control strategies for rust diseases in wheat include cultural methods, chemical control and breeding for resistance (Knott, 1989). Resistance breeding involves utilising resistant genes of which the first was identified when working with yellow rust (Biffen, 1905). Several resistance genes that are effective against the leaf, stem and yellow rust pathogens, have been designated. The gene of importance for this study is leaf rust resistance gene *Lr37*, which maps to chromosome 2AS (Bariana & McIntosh, 1993). The gene is effective against leaf rust pathotypes UVPrt2 and UVPrt13, especially when combined with genes such as *Lr13* and gene *Lr34* (Kloppers & Pretorius, 1997).

Resistance genes can be incorporated into susceptible wheat lines through backcrossing. The presence and segregation of the genes can be determined from infection types in F_2 or BCF_2 populations (Kolmer, 1996). Determining the presence of genes in lines where more than one gene is present is, however, not always possible. Marker-assisted selection (MAS) is the indirect selection of resistant genes, in selecting for a marker linked to the resistant gene of interest (Melchinger, 1990). Markers used to select for genes include morphological markers, like leaf tip necrosis, which is a symtom linked to leaf rust resistance gene *Lr34* (Singh, 1992). Morphological markers are limited in number (Paterson *et al.*, 1991 (b)), but both protein and DNA, or molecular markers exists for various resistance genes. DNA

markers are especially effective in selecting for genes where genes are combined (gene pyramiding) in breeding programs (Kelly, 1995).

One advantage of using molecular techniques to identify markers linked to genes includes the direct exposure of the genome, whereas traditional morphological markers interferes with each other and have undesirable pleïtropic effects (Weeden *et al.*, 1994). Molecular techniques that have been used to identify markers linked to important leaf rust resistance genes include RFLPs (Autrique *et al.*, 1995; Schachermayr *et al.*, 1995; Dubcovsky *et al.*, 1998), RAPDs (Schachermayr *et al.*, 1994; Procunier *et al.*, 1995; Dedryver *et al.*, 1996; Gupta *et al.*, 1998; Naik *et al.*, 1998; Troskie *et al.*, 1999), AFLPs (Zabeau & Vos, 1993; Lottering *et al.*, 1999) and STS markers (Gupta *et al.*, 1998; Naik *et al.*, 1998; Seyfarth *et al.*, 1999; Lottering *et al.*, 1999).

Approaches to locate molecular markers linked to leaf rust resistant genes consist of three integrated steps. First, genes are phenotypically identified through a genetic analysis of two to three plant generations inoculated with *P. recondita* f. sp. *tritici* (Kolmer, 1996). DNA analysis is subsequently performed with molecular techniques to identify unique banding patterns between resistant and susceptible plants. Polymorphisms obtained between resistant and susceptible plants with molecular techniques are then compared with the initial phenotypic analysis to identify molecular markers linked to the gene of interest. The distance of a marker linked to a specific gene can then be calculated in centimorgans (Kosambi, 1944).

2.2 *Wheat rusts*

2.2.1 *Wheat rust species*

Wheat rust pathogens belong to the genus *Puccinia*, the family Pucciniaceae and the order Uredinales of the class Basidomycetes (Littlefield, 1981). Leaf rust caused by *P. recondita* Rob. Ex Desm. f. sp. *tritici* is the most common of the three rust species that marginalize wheat production annually (Roelfs *et al.*, 1992). Two other species are stem rust (*P. graminis* Pers f. sp. *tritici*) and yellow or stripe rust (*P. striiformis*

Westend. P. sp. *tritici*) (Knott, 1989). The three species differ in life cycle, morphology and optimal growth conditions (Knott, 1989).

2.2.2 *Asexual stages of the leaf rust pathogen in the wheat plant*

The life cycle of leaf rust has both asexual and sexual stages. Wheat and related grasses serve as hosts during the asexual stage, while the sexual stage occurs in *Thalictrum* spp. (Knott, 1989). The asexual stage is the damaging stage from a wheat production perspective.

The asexual stage of the rust life cycle in the wheat plant is initiated when wind-borne urediospores make contact with the plant. Wind-borne urediospores is the main spreading method of the pathogen (Roelfs & Bushnell, 1985). These wind-borne urediospores can travel long distance. leaf rust urediospores were shown to travel from the autum shown areas in the south to the northern spring-shown areas of America (Roelfs *et al.*, 1992).

Urediospores that make contact with plant leaves initiate the fungal infection of the host plant. The urediospores carry a limited supply of triglyserides that function as a metabolic energy source for the formation of urediospore germ tubes (Bushnell & Roelfs, 1984). The germ tubes grow laterally on the surface of the leaf and penetrate the leaf through the stomata, where infection structures develop and eventually enter the mesophyll cells of the wheat plant (Bushnell & Roelfs, 1984). Uredia of leaf rust appear as small round yellow/brown pustules on the surface of leaves (Fig 2.1). When 37 % of the leaf surface is covered with leaf rust uredia, and the whole interior is colonized with mycelium, severity is taken as a 100 % (Knott, 1989). Yield loss due to leaf rust ranges between 5– 15 %, depending on the stage of development of the crop when initial infection occurs (Chester, 1946).

The mutation rate of leaf rust for virulence is high because of the large population of urediospores produced in leaf rust infested wheat fields. In a hectare field with 1 % infection of the leaf area by leaf rust, uredia could produce urediospores at a rate of 10^{11} spores per day. The more urediospores that are produced, the greater the

Fig 2.1 Uredia produced by *Puccinia recondita* f. sp. *tritici* on flag leaves of susceptible wheat plants.

chance of obtaining races with virulence against any existing resistance of their wheat hosts   (Parlevliet & Zadoks, 1977).   Newly mutant races with virulence against existing resistance will increase dramatically as a result of selection pressure being in favor of virulent races (Johnson, 1961).   The high possibility of new virulent pathogens increases, therefore the demand for new resistant genes and resistance gene combinations conferring durable resistance.

2.3 Control of rust diseases

2.3.1 Methods of rust control

Controlling methods of the three rust diseases include cultural methods where the life cycle of the fungus is interrupted, chemical control with expensive fungicides, and breeding for resistance, which is the most effective with minimum additional input costs (Knott, 1989). Breeders therefore aim to introduce effective resistance genes or combinations of resistance genes against wheat rusts in cultivars suitable for growth in environments where rusts commonly occur.

2.3.2 Genes for disease control

Rust resistance genes in wheat can be identified phenotypically through a genetic analysis. This procedure involves a cross between the cultivar carrying the gene with a susceptible parent to obtain a F_1 population. The F_1 population can be selfed to obtain a F_2 population, or backcrossed to the susceptible parent to obtain a BCF_1 population. F_2 populations and BCF_2 populations are then selfed to obtain F_3 or BCF_3 populations, respectively (Kolmer, 1996). Rust resistance genes are identified phenotypically with the inoculation of the different generations of plants with different rust races. Leaf rust resistance genes can therefore sometimes be identified with inoculations using different races of *P. recondita* f. sp. *tritici* where there is a host: pathotype interaction. Phenotypic classifications of plants are then performed on a 0-4 infection type scale to distinguish between resistance and susceptibility (Stakman *et al.*, 1962; Roelfs, 1988) (Table 2.1).

Table 2.1 Classification of different infection types of wheat infected with *Puccinia recondita* f. sp. *tritici* (Stakman *et al.*, 1962; Roelfs, 1988).

Class	Infection type	Description
Immune	0	No infection signs.
Very resistant	0;	No uredia – distinct flecks.
Resistant	1	Small uredia surrounded by yellow chlorotic or necrotic areas.
Moderately resistant	2	Small to medium size uredia typically in a dark green island surrounded by chlorotic area.
Mesothetic or heterogeneous	X	A range of infection types on a single leaf caused by a single isolate.
Moderately susceptible	3	Medium size uredia usually surrounded by a light green chlorosis.
Susceptible	4	Large uredia with a limited amount of chlorosis.

The first single resistance gene inherited in a Mendelian fashion against rust was identified in the yellow rust-wheat pathosystem (Biffen, 1905). The gene-for-gene concept was later introduced after the investigation of flax rust and its host. In the case of flax rust, where resistance is the result of a single gene in the host, virulence will be the result of a single virulent gene in the pathogen (Flor, 1942, 1946, 1947).

The close coevolution of host and pathogen is the reason why resistance genes are often found in the country of origin of both host and pathogen. Mapping of these resistance genes to chromosome arms is done with telocentrics (Sears, 1962, 1962a).

A number of single resistance genes have been designated for yellow rust, stem rust and leaf rust resistance. The number of resistance genes identified and mapped for leaf rust only, has grown to 47 (Dubcovsky *et al.*, 1998), with a molecular marker to identify this 47th gene (Helguera & Kahn, 2000) (Table 2.2).

The focus of the study is mainly on leaf rust resistance gene *Lr37*, because the gene confers effective adult-plant resistance to *P. recondita* f. sp. *tritici* (Dyck & Lukow, 1988; Pretorius, 1990). Leaf rust resistance gene *Lr37* originates from the French cultivar VPM1 (Doussinault *et al.*, 1981). VPM1 is a cultivar that showed resistance against eyespot and was derived from *Triticum ventricosa* (Dyck & Lukow, 1988). The gene also showed a dominant adult-plant response in the greenhouse after inoculations with pathotype UVPrt9 of *P. recondita* f. sp. *tritici* (Kloppers, 1994). F_1 seedlings obtained with crosses involving RL6081 and the South African cultivar Karee were susceptible to UVPrt9. The F_2 population segregated resistant to susceptible in a ratio of 1:3 for both UVPrt2 and UVPrt9 inoculations, indicating the gene to be recessive or partially dominant in wheat seedlings (Kloppers & Pretorius, 1994). Gene *Lr37* is further mapped together with the genes *Sr38*, *Yr17* and *Pm46* to chromosome 2AS (Bariana & McIntosh, 1993). An increase in resistance to leaf rust was also demonstrated when gene *Lr37* was combined with genes such as *Lr13* and gene *Lr34* (Kloppers & Pretorius, 1997).

Table 2.2 Designated leaf rust resistance genes, genome locations, linkage of the genes among each other, original gene sources, infection types, tester lines and molecular markers available with references and gene references.

Lr gene	Genome location	Linkage	Original source	Tester line	Molecular markers linked to the gene	Marker reference	Gene References
1	5DL		Malakof	RL6003	RFLP	Feuillet *et al.*, 1995.	Ausemus *et al.*, 1946.
2b	2DS		Carina	RL6019	RFLP	Parker *et al.*, 1998.	"
2c	2DS		Brevit	RL6047			"
3a	6BL		Democrat	RL6002			Browder, 1980.
3bg	6BL		Bage	RL6042			"
3ka	6BL		Klein Aniversario	RL6007			"
9	6BL		*Triticum umbellulatum*	RL6010	RAPD,STS (complete), RFLP(8 cM)	Schachermayr *et al.*, 1994.	Soliman *et al.*, 1962.
10	1AS		Lee	RL6004	RFLP, STS	Schachermayr *et al.*, 1997.	Choudhuri, 1958.
11	2A		Hussar	RL6053			Soliman *et al.*, 1962.
12	4BS		Exchange	RL6011			Dyck *et al.*, 1966.
13	2BS		Frontana	Manitou	RFLP	Seyfarth *et al.*, 1998.	Dyck *et al.*, 1966.
14a	7BL		Hope	RL6013			Dyck & Samborski, 1970.
14b	7BL		Bowie	RL6006			Dyck & Samborski, 1970.

Lr gene	Genome location	Linkage	Original source	Tester line	Molecular markers linked to the gene	Marker reference	Gene References
15	2DS		Kenya 1-12 E-19-J	RL6052			Luig & McIntosh, 1968.
16	2BS	to Sr23	Exchange	RL6005			Dyck & Samborski, 1968.
17	2AS		Klein Lucero	RL6008			"
18	5BL		T. timopheevi	RL6009			"
19	7DL	to Sr25	Agropyron elongatum	RL6040	RFLP	Autrique et al., 1995.	Sharma & Knott, 1966.
20	7AL		Thew	Thew	RFLP	Parker et al., 1998.	Browder, 1972.
21	1DL		T. tauschii	RL6043			Rowland & Kerber, 1974.
22a	2DS		Thatcher	RL6044			Rowland & Kerber, 1974.
22b	2DS		T. tauschii	Thatcher			Dyck, 1978.
23	2BS		Gabo	RL 6012	RFLP	Nelson et al., 1997.	McIntosh & Dyck, 1972.
24	3DL	to Sr24	A. elongatum	RL6064	RFLP (complete) and RAPD (complete)	Schachermayr et al., 1995; Dedryver et al., 1996	Browder, 1972.
25	4AB		Rosen rye	Transec	RAPD (tight)	Procunier et al., 1995.	Driscoll & Anderson, 1967.

Lr gene	Genome location	Linkage	Original source	Tester line	Molecular markers linked to the gene	Marker reference	Gene References
26	1BL	to *Sr31* & *Yr9*	Imperial rye	RL6078			Singh *et al.*, 1990.
27	3BS	to *Sr2*	Gatcher	Gatcher	RFLP	Nelson *et al.*, 1997.	Singh & McIntosh, 1984.
28	4AL		*T. speltoides*	RL6079	RAPD, STS	Gupta *et al.*, 1998; Naik *et al.*, 1998.	McIntosh *et al.*, 1982.
29	7DS		*A. elongatum*	RL6080	RAPD RFLP	Procunier *et al.*, 1995. Nelson *et al.*, 1997.	Sears, 1973.
30	4BL		Terenzio	RL6049			Dyck & Kerber, 1981.
31	4BS		Gatcher	Gatcher	RFLP	Nelson *et al.*, 1997.	Singh & McIntosh, 1984.
32	3D		*T. tauschii*	RL5497-1	RFLP RAPD	Autrique *et al.*, 1995. William *et al.*, 1997.	Kerber, 1987.
33	1BL	to *Lr44*	PI58458	RL6057			Dyck. 1987.
34	7D	to *Yr18*	Terenizo	RL6058		Nelson *et al.*, 1997.	Dyck, 1987.
35	2B		*T. speltoides*	RL5711	STS (complete)	Seyfarth *et al.*, 1999.	McIntosh *et al.*, 1983.

14

Lr gene	Genome location	Linkage	Original source	Tester line	Molecular markers linked to the gene	Marker reference	Gene References
36	6BS		*T. speltoides*	ER84018			Dvorak & Knott, 1990.
37	2AS	to *Sr38* & *Yr17*	*T. ventricosa*	RL6081		Troskie *et al.*, 1999.	Bariana & McIntosh, 1994.
38	2AL		*A. intermedium*	RL6097			Friebe *et al.*, 1993.
39	2DS		*T. tauschii*	KS86NGRC O2			Gill *et al.*, 1988.
40	1 D		*T. tauschii*	KS89WGRC O7			Gill *et al.*, 1991.
41	1 D		*T. tauschii*	KS90WGRC 10	AFLP, STS (22cM)	Lottering *et al.*, 1999.	Cox *et al.*, 1994.
42	1D			WGRC11			Cox *et al.*, 1994.
43	7D		*T. tauschii*	WGR16			Cox *et al.*, 1994.
44	1BL	to *Lr33*	*T. aestivum spelta* 7831	RL6147			Dyck & Sykes, 1993.
45	2AS		Rye	RL6144			McIntosh *et al.*, 1995.
46	1B		Pavon 76	Lalbahadur (*Lr1*)			Singh *et al.*, 1998.
47	7AS		*T. speltoides*	KS 90H450	RFLP	Helguera & Khan, 2000.	Dubcovsky *et al.*, 1998.

15

2.3.3. Gene pyramiding

Gene pyramiding is the best method of genetic control for leaf rust (Samborski & Dyck, 1982; Roelfs, 1988). The strategy refers to the combination of several resistance genes in a single genotype, mainly because virulent races of pathogens such as *P. recondita* f. sp. *tritici* can overcome single resistance genes. Gene pyramiding's effectiveness is, because a pathogen has to mutate against more than one resistance gene in order to obtain virulence (Schafer *et al.*, 1963). An example of the effectiveness of gene pyramiding was illustrated when a race of *P. recondita* f. sp. *tritici* virulent to *Lr13*, was introduced into Australia in the 1980s. This specific race of *P. recondita* was avirulent when *Lr13* was combined with genes *Lr1*, *Lr2*, *Lr3a*, *Lr23* or *Lr22b* (Luig, 1985). Increased resistance against yellow rust was also obtained, when yellow rust resistance genes were combined (Krupinsky & Sharp, 1978). The aim of gene pyramiding is therefore to obtain durable resistance in a crop like wheat, where durable resistance can be defined as resistance that remains effective over a wide spread area for a long time (Johnson & Law, 1975).

2.3.4 Success with resistance breeding

Advantages of implementing resistance genes in breeding programs can be seen worldwide in various crops. As a result of resistant cultivars, wheat farmers in the northern parts of Australia and Queensland last suffered severe damages due to leaf rust in 1960. In 1973, a major loss in production occurred in the south east of Australia due to the lack of resistant plants against leaf rust (McIntosh & Brown, 1997).

The financial benefits are also evident, when cultivars with resistance to leaf rust are used instead of susceptible cultivars. Savings of $217 million were estimated on the Canadian prairies since 1979, due to the introduction of cultivars with resistance to rust diseases (Green & Campbell, 1979). In north western Mexico the amount saved because of durable non-specific resistance against leaf rust was estimated to be $17 million on 120 000 ha of wheat from 1970-1990 (Smale *et al.*, 1998).

Breeding programs focus on different aspects of wheat improvement. A comparison thereof over the past few decades shows gains from breeding for resistance at least parallel to gains from breeding for an increased in yield potential (Byerlee & Moya, 1993). Advantages of implementing resistance genes are insurmountable when faced with an overwhelming increase in demand for food worldwide.

2.4 Wheat

2.4.1 General

The total area under wheat cultivation worldwide is 118.136 million ha of which 1.3 million ha is cultivated in Southern Africa (Reeves *et al.*, 1999). This large amount of cultivated land is, however, not enough to meet the needs of the explosive population increase globally. Developing countries are expected to increase their demand for cereals by 80 % between 1999 and 2020, as the global demand for wheat will increase by 40 % and that of maize by 47 % (Pinstrup-Anderson & Pandya-Lorch, 1997). A way to counter this problem is to increase production in regions that can accommodate production intensification (Borlaug & Downswell, 1997).

Marginal wheat producing areas totals 45 % globally and in the case of South Africa 91 % (Reeves *et al.*, 1999). A region is defined as marginal when production drops to 70 % of the optimum yield. These areas should be exploited more in order to increase wheat production worldwide. One way of increasing wheat production is to grow resistant plants in areas where various pests and diseases like leaf rust, do occur.

2.4.2 The Wheat genome

Wheat has a large genome of approximately 1.7×10^{10} bases per haploid nucleus (Devos & Gale, 1993). The genus consists out of three groups namely diploids, tetraploids and hexaploids (Knott, 1989). With a basic chromosome number of seven, these three groups have chromosome numbers of 14, 28 and 42 (Sakamura, 1918). Einkorn wheat is a diploid ($2n = 2x = 14$), durum and emmer wheats are tetraploids ($2n = 4x = 28$) and bread wheat is a hexaploid ($2n = 6x = 42$) (Knott, 1989). Meiotic

chromosome pairing studies have indicated that the three genomes of bread wheat are A, B & D (Knott, 1989). Hexaploid bread wheat is a result of the cross between wheat with a genome composition of AABB and the diploid *T. tauschii* DD (Kimber & Sears, 1987).

2.4.3 Genetic diversity of wheat plants

The earliest cereal crops existed as landraces, which included a mixture of diverse genotypes. Historically cereal crops often consisted of mixtures of wheat, oats and barley. Recent emphasis on high yielding crops grown in monoculture has led to genetic uniformity and subsequent susceptibility to diseases like leaf rust (Feldman & Sears, 1981; Knott, 1989). Natural resistance of wheat to diseases also declines with a decrease in genetic variability (Plucknett *et al.*, 1983).

Natural crosses and polyploidy played a major role in the evolution of wheat and inter generic crosses were therefore successful in the development of triticale (Knoblock 1968; Sharma & Gill, 1983). Genetic variation can be found in close relatives of wheat, which provide useful sources of resistance genes (Knott & Dvorak, 1976; Knott, 1978). *Aegilops tauschii* is an example of a close relative of wheat with a wide genetic diversity that can be exploited for resistance genes (Gill *et al.*, 1986; Cox *et al.*, 1992; Lange & Jochemsen, 1992; Innes & Kerber, 1994).

Breeders are continuously trying to increase genetic diversity of crop plants and in the process increase yield and natural resistance to various diseases. Factors that co-evolved with crop breeding include the need to distinguish different crop varieties from each other and the need to establish the purity of crops. Factors like these drive the need to introduce plant variety rights, seed certification and the protection of germplasm (Russell *et al.*, 1997). Molecular techniques can be used to calculate genetic distances among wheat cultivars and to identify wheat lines for the protection of plant breeder rights (Myburg *et al.*, 1997). Molecular techniques have also recently been used to estimate genetic diversity of German and Austrian winter wheat cultivars in order to increase prospects for crosses for line development (Bohn *et al.*, 1999).

2.4.4 Genetic distance analysis of wheat plants

The further apart wheat lines are genetically, the better the chances of increasing genetic variation which is desired in species like wheat where this is known to be low (Gale *et al.*, 1990; Devos & Gale, 1992). In order to maintain the same genetic composition of plants, donor species of resistance genes should be as closely related to cultivated species as possible (Bretó *et al.*, 1994). Genetic distances among different wheat lines can therefore be used to predict the viability of crosses to release new lines (Bohn *et al.*, 1999).

Wheat genotypes can be compared using two main fingerprinting strategies. Firstly, the most parsimonious tree can be calculated directly from an input data matrix using the program PAUP (Swofford, 1991). The trees are illustrated in the form of a cladogram, which is the most basic of trees and shows relative recency of common ancestry (Page & Holmes, 1998). The input matrix consists of a binary positive-negative score for different characters compared among the wheat genotypes. The same data matrix can also be used to calculate pairwise genetic distances among cultivars. Pairwise genetic distances are calculated with the program UPGMA (Swofford, 1991). The fraction of characters shared between any two individuals is firstly used to calculate the similarity coefficient (S) between any two individuals. The similarity coefficient is then converted to a genetic distance with the formula:

$$D = 1 - S \text{ or } D = -\ln(S).$$

The index of genetic similarity (F), although initially used for RFLP data, can be used to calculate the similarity coefficient between any two cultivars with the formula (Nei & Li, 1979):

$$F = 2n_{xy}/(n_x + n_y)$$

F is the ratio of shared bands between two individuals x and y; n_x and n_y are the number of bands from each individual, while n_{xy} is the number of shared bands between the compared individuals. The cluster analysis of the pairwise genetic

distances among different genotypes can then be used to generate a dendogram in order to visualize the pairwise genetic distances among the compared genotypes. Dendograms are additive trees where the tips of the trees are equidistant from the root of the tree and show not only relative recency of common ancestry, but can also be used to illustrate evolutionary time (Page & Holmes, 1998).

2.4.5 Mapping of the wheat genome

The first map of the human genome (Botstein *et al.*, 1980) inspired mapping of agronomically important genes in crop plants with molecular markers. Mapping of specific genes to chromosomes can be achieved with crosses involving monosomic lines and the location of genes on specific chromosome arms can be done with crosses involving telocentric lines (McIntosh, 1987).

Model species are used for the discovery of genes and their function (Sorrels, 1999). Comparative mapping is used to locate genes and provide a wider spectrum of resistance genes to improve cultivated species (Snape *et al.*, 1995; 1996). For wheat, restriction fragment length polimorphism (RFLP) and simple sequence repeat (SSR) maps are available (Devos & Gale, 1993; Xie *et al.*, 1993; Nelson *et al.*, 1995a, b, c; Van Deynze *et al.*, 1995; Marino *et al.*, 1996; Röder *et al.*, 1998). These markers can be tested with the segregation of populations to estimate possible linkage to important genes. Molecular markers linked to genes can then be used in plant breeding programs to facilitate selection and for determining the physical location of genes (Tanksley *et al.*, 1989; Martin & Harvey, 1994). Molecular techniques are therefore important tools for wheat breeding and are discussed in the following section.

2.5 Molecular techniques

2.5.1 General

Comparisons of plants using molecular techniques have the advantage of directly exposing sequence variation in the genomes. With traditional morphological markers this was often not possible, because the markers interfere with one another and have

undesirable pleiotropic effects. Protein markers on the other hand can access only a limited amount of genes (Weeden *et al.*, 1994).

In the early 1980s, RFLPs were the dominant marker type (Botstein *et al.*, 1980). Since then a range of molecular techniques have been developed. The techniques used in this study to search for molecular markers linked to *Lr37* leaf rust resistance gene include RAPDs, AFLPs and SSRs.

2.5.2 *RFLPs*

RFLPs is the first DNA based marker technique. When using the technique, polymorphisms between two compared genomes result, because of DNA fragment length variation between the genomes after the complete digestion of the compared genomes with restriction enzymes (Botstein *et al.*, 1980). The technique was originally used for human genome analysis (Botstein *et al.*, 1980), but agronomic applications were soon exploited in RFLP markers for plant breeding purposes (Tanksley *et al.*, 1989).

This DNA based technique provides more detectable markers than protein marker systems, is not sensitive to environmental factors and can be used at any developmental stage of the plant (Kelly, 1995). The main advantage of RFLPs is their co-dominant inheritage (Waugh & Powell, 1992; Rafalski & Tingey, 1993). Co-dominant markers can distinguish between heterozygots and homozygots for specific loci.

Disadvantages of the technique are cost, labour intensity and time (Myburg, 1996). The most time consuming part is the visualization of the digested DNA fragments. The DNA fragments are firstly, separated on agarose gels and then transferred to nylon membranes with the "Southern Blot" technique (Southern, 1975). The nylon membranes containing the DNA fragments are then probed with radioactively labeled oligomers and then exposed to X-ray film in order to visualize the DNA fragments.

2.5.3 *RAPDs*

RAPDs are an alternative markertype to RFLPs that requires less DNA, no Southern blotting and radioactive labeling, and is relatively quick to assay (Welsh & McClelland, 1990; Williams *et al.*, 1990). Several independent loci can be amplified (Williams *et al.*, 1993) and with the number of RAPD primers available, numerous loci can be amplified to compare among plants.

When RAPD amplification products of resistant and susceptible plants are compared between each other, polymorphisms result, because of different primer binding sites between the genomes. The different primer binding sites result from point mutations, or DNA sequence changes that result in different fragment sizes of amplified DNA (Demneke *et al.*, 1997). The efficiency of RAPDs in generating polymorphisms among different genomes is a reason why the technique is increasingly used to mark agronomically important genes (Martin *et al.*, 1991; Michelmore *et al.*, 1991; Haley *et al.*, 1994; Paran & Michelmore 1993; Penner *et al.*, 1993; Williamson *et al.*, 1994; Myburg *et al.*, 1997).

Disadvantages of RAPDs include the fact that the technique is very sensitive to several reaction parameters (Nair *et al.*, 1995), which results in poor reproducibility of RAPD amplification products. This is mainly due to impurities in the DNA template (Micheli *et al.*, 1994). The dominant nature of RAPDs also restricts the identification of homozygous plants from heterozygous plants (Kleinhofs *et al.*, 1993). Another disadvantage of RAPDs in a species like wheat with 80 % repetitive DNA (Bennet & Smith, 1976), is that only one in 10 RAPD amplified products does not contain repetitive DNA sequences (Devos & Gale, 1992). One out of ten RAPD amplification products will therefore contain useful DNA sequence information.

2.5.4 *AFLPs*

AFLPs are a more recently developed technique than RAPDs, based on the restriction of the genome with a set of restriction enzymes and the PCR amplification of the subset of restricted fragments (Zabeau & Vos, 1993). AFLPs provide a large number

of amplified loci without any prior sequence information of the genome and can be visualized on a single gel (Vos *et al.*, 1995). The large amount of amplified DNA fragments obtained when using this technique increases the probability of finding polymorphisms and thus enhances the potential for fingerprinting and mapping (Zabeau & Vos, 1993; Thomas *et al.*, 1995; Vos *et al.*, 1995). Thomas *et al.* (1995) were among the first to apply this technique for identifying markers linked to disease resistance in agronomic crops. A number of applications for AFLPs have been found since, and a marker linked to leaf rust resistance gene *Lr41* was recently found (Lottering *et al.*, 1999).

AFLPs are useful for fingerprinting and can be used to assay diversity within species and pedigrees (Mackill *et al.*, 1996; Paul *et al.*, 1997; Barret *et al.*, 1998a, b; Law *et al.*, 1998; Yee *et al.*, 1999). AFLPs can also, be used for mapping (Mackill *et al.*, 1996). With more amplified DNA fragments, more polymorphism is expected. AFLPs therefore showed the highest diversity index in barley, when compared with techniques such as RFLPs, SSRs and RAPDs (Thomas *et al.*, 1995).

A disadvantage of AFLPs, as for RAPDs, is the dominant nature of the techniques (Mackill *et al.*, 1996). With a dominant marker type, heterozygotes for an allele can usually not be distinguished from homozygotes, which is important in breeding programs. AFLPs are also more time consuming and expensive than RAPDs, but the number of loci that can be scored is greater than any other technique (Mackill *et al.*, 1996).

2.5.5 *SSRs*

All eucaryotic genomes have a tendency to contain a class of sequences termed microsatellites (Litt & Luty, 1989), or simple sequence repeats (SSRs) (Tautz *et al.*, 1986). These repeats consist of repeated oligonucleotides such as $(GA)_n$ or $(GT)_n$ (Röder *et al.*, 1998). The length variation of SSRs is often the result of slipped strand mispairing during DNA replication (Levinson & Gutman, 1987). Slipped strand mispairings are forced with the repeats of the small uniform DNA sequences and the technique is therefore highly polymorphic.

The high polymorphic rate of the technique was demonstrated in barley where only four SSR primer pairs were sufficient to discriminate between two closely related barley cultivars (Russell *et al.*, 1997). A few SSR primer pairs also provided sufficient markers to discriminate among closely related wheat breeding lines (Röder *et al.*, 1995; Plaschke *et al.*, 1995; Bryan *et al.*, 1997). SSRs are also widely used in wheat for genome mapping (Röder *et al.*, 1998). Further applications of SSRs in wheat breeding include markers for rust resistance genes, including a marker linked to yellow rust resistance gene *Yrvav* (Bariana *et al.*, 1999), and genetic diversity studies (Plaschke *et al.*, 1995). SSRs are also multiallelic and have therefore great application in evolutionary studies (Schloetterer *et al.*, 1991, Buchanan *et al.*, 1994).

The advantages of SSRs are the simplicity of this PCR based technique and the repeatability (Russell *et al.*, 1997). SSRs are also codominant (Weissenbach *et al.*, 1992) and describe a large number of alleles (Lee *et al.*, 1995). Another advantage is the fact that a much higher level of polymorphism has been obtained with SSRs in wheat than with other marker techniques (Röder *et al.*, 1995; Plaschke *et al.*, 1995; Bryan *et al.*, 1997).

2.5.6 *SCARs*

Reproducibility can be a problem when using a marker technique like RAPDs, mainly because of the low temperatures (36-40 °C) used for the annealing of the short oligonucleotide primers (Kelly, 1995). The specificity of markers can be increased with longer more specific primers and higher annealing temperatures (50-65 °C) (Kelly, 1995). The amplification products of the longer primers also show a less complex banding pattern than RAPDs and AFLPs (Procunier *et al.*, 1997). The amplification products of these longer more specific primers known as ASAPs (Weeden *et al.*, 1992), SCARs (Kesseli & Michelmore, 1992; Paran & Michelmore, 1993), or STSs (Olson *et al.*, 1989) can be particularly useful for incorporating various resistance genes in different genetic backgrounds (Paran & Michelmore, 1993).

The application of molecular markers in breeding programs depends on the speed and reliability of the markers to identify genes (Penner *et al.*, 1996). The single amplification product of the SCAR marker used to detect gene *Lr24*, enables researchers to detect the gene without any electrophoretic separations, but with merely the addition of ethidium bromide to the PCR amplification product and the exposure thereof, to UV light (Dedryver *et al.*, 1996). The simplicity with which gene *Lr24* is detected with the SCAR marker makes the technique highly desirable for gene identifications.

2.6 Optimization of PCR based technologies

When using PCR based techniques to locate molecular markers, several of the components concentrations need to be optimized in order to obtain reproducible results. The reaction component concentrations that need to be optimized include DNA template, primer, $MgCl_2$, enzyme and dNTP concentrations (Bassam *et al.*, 1991). A modification of the method proposed by Taguchi, to optimize various reaction component concentrations with the least amount of reactions were proposed by Cobb & Clarkson (1994).

2.6.1 Template DNA

DNA isolation plays an important role in the quality of the DNA used as template in PCR reactions. DNA isolations without hazardous chemicals (phenol or chloroform) have been performed with good PCR results (Edwards *et al.*, 1991). A slight modification, where the ethanol precipitated DNA is scooped with a glass rod produced more reproducible RAPD amplification products than precipitation of DNA by centrifugation (Micheli *et al.*, 1994). An excess of DNA template further causes a smear with undistinguishable banding patterns after RAPD amplifications (Yu *et al.*, 1993). The optimum template concentration to obtain reproducible RAPD banding patterns with wheat DNA is estimated to be in the region of 0.2-0.4 ng/μl (Devos & Gale, 1992). The intensity of PCR banding patterns was also shown to be affected by DNA template concentration. It is therefore important to standardize all DNA

template concentrations where different samples are compared (Hosaka & Hanneman, 1994).

2.6.2 $MgCl_2$

$MgCl_2$ concentration plays an important role in PCR reactions. Mg^{2+} concentration depends on the concentrations of dNTPs, PPi and EDTA. It is therefore practical to optimize $MgCl_2$ and dNTP concentrations last (Rolfs et al., 1992). The total Mg^{2+} concentration should further exceed the dNTP concentration by 0.5-2.5 mM (Schienert, 1995), but the stringency of annealing decreases with an increase in $MgCl_2$ concentration (Myburg, 1996). The proposed concentration of $MgCl_2$ to produce reproducible RAPD results is 1.5 mM per reaction for wheat (Devos & Gale, 1992).

2.6.3 dNTPs

The dNTP concentration of PCR reactions depends on the length of amplification products, $MgCl_2$ concentrations and reaction stringency (Myburg, 1996). Unbalanced dNTP concentrations decreases Taq DNA polymerase fidelity (Rolfs et al., 1992) and low dNTP concentration also decreases the number of bands obtained in RAPD reactions (Yu et al., 1993).

2.6.4 Primers

Primer concentrations have an effect on the number of bands obtained for RAPD reactions. Where primer concentrations were decreased the number of bands amplified decreased (Yu et al., 1993). DNA template concentrations have to be optimized for different primers to prevent background smears (Devos & Gale, 1992). The GC content of primers further influences the number of loci amplified with RAPDs. The average number of loci amplified with primers with a GC content of 40 % were increased from 2.45 to 8.14, when the GC content was increased to 70 % (Hosaka & Hanneman, 1994) due to priming stability (Fukouka et al., 1992).

2.6.5 *Enzyme*

Enzyme concentrations affect the specificity of PCR reactions; the higher the enzyme concentration the lower the PCR specificity (Rolfs *et al.*, 1992). Enzyme activity is also dependent on the temperature of the reaction. The synthesis rate of the enzyme is 0.25 nucleotides per second at 22 °C and 150 nucleotides per second at 75-80 °C (Gelfand & White, 1990). The half-life of *Taq* DNA polymerase is reduced from 130 min at 92.5 °C to 6 min at 97.5 °C (Gelfand & White, 1990).

2.6.6 *Reaction volume*

The reaction volume can also affect the accuracy of PCR amplifications. Ramp and hold temperatures decrease with smaller reaction volumes and the shorter the ramp and hold temperatures, the higher the amplification specificity. Half reactions (25 µL) were therefore suggested for PCRs (Rolfs *et al.*, 1992).

2.6.7 *Number and duration of PCR cycles*

A basic three step PCR was suggested for RAPD reactions with a denaturization step of 1 min at 94 °C, a primer annealing step of 1 min at 36 °C and an elongation step of 2 min at 72 °C (Williams *et al.*, 1994). It was further suggested to initiate the PCR reaction with a 3-5 min denaturation step in order to denature complex genomic DNA completely. DNA will not renature completely with steps following this initial denaturing step and denaturing times can therefore be decreased in steps following the initial denaturization (Myburg, 1996). The total reaction time will therefore be decreased with the advantage of increasing the life span of *Taq* DNA polymerase throughout all the PCR cycles.

Annealing temperature of different primers is the crucial part in PCR reactions. The specificity with which primers anneal to template DNA depends on the temperature, the time, the number of single stranded targets and the concentration of the primers (Rolfs *et al.*, 1992). The GC content of the primers also plays a role in annealing time. The annealing time required for primers with a GC content of 70 % or 80 % is 5

s, while 30 s are required for primers with GC contents of 50 % to 60 % (Yu & Pauls, 1992).

A linear relationship exists between extension time and the maximum size of the fragments amplified during RAPD reactions. With an extension time of 5 s a 0.6 kb fragment can be amplified, while 60 s is necessary to amplify a fragment of 3 kb (Yu & Pauls, 1992). It is also preferable to end a PCR reaction with a final extension time of 5-15 min at 72 °C, to avoid wrongly re-annealed PCR products, which occur if reactions are terminated at 94 °C (Rolfs *et al.*, 1992). The total number of PCR cycles further increases PCR product concentration. However, no significant differences in product concentrations have been obtained among PCR reactions compared at 35, 55 and 75 cycles because PCR reagents become depleted after a certain number of cycles (Yu & Pauls, 1992). The maximum PCR product concentration will thus be obtained with a critical number of cycles. Thirty-five cycles were sufficient to identify RAPD polymorphisms between wheat plants carrying Russian wheat aphid resistant genes and susceptible plants (Myburg *et al.*, 1998).

2.6.8 *Gel electrophoresis*

Electrophoresis plays an essential role in the interpretation of the end product of PCR based techniques. Agarose gels are sufficient to separate RAPD amplification products and single PCR products (Sambrook *et al.*, 1989). The resolution of agarose gels is, however, lower than that of acrylamide gels (Sambrook *et al.*, 1989) and a more complex banding pattern can be observed when RAPD amplification products are separated on acrylamide gels (Sobral & Honeycutt, 1994). The latter is therefore preferred when AFLP amplification products are separated, because an AFLP primer pair can amplify 20-100 fragments per amplification reaction (Perkin-Elmer, 1996). The number of fragments amplified with AFLP reactions causes the products to co-migrate in a smear on agarose gels. Acrylamide separations are also preferred in the case of SSRs, because the high resolution of the gels make it possible to distinguish between fragments with the minimum changes in base pairs, which is often the case for SSRs where small DNA repeats force slipped strand mispairings (Levinson & Gutman, 1987).

2.7 Marker assisted selection

2.7.1 General

Marker assisted selection (MAS) is the indirect selection of genes that confer resistance to diseases or pests with the aid of a marker linked to the gene (Melchinger, 1990). Markers used for the identification of genes include morphological and molecular markers. Morphological markers can be used to identify genes such as leaf rust resistance gene *Lr34*, which is associated with leaf tip necrosis symptoms (Singh, 1992). Marker assisted selection with the aid of morphological markers is, however, limited, because of the low number of markers available and difficulty of ditecting these markers in segregating populations (Paterson *et al.*, 1991). The alternative for morphological markers is molecular markers, which can be used to accelerate the introgression of single resistance genes into plants and for the accumulation of major and minor resistance genes as well as to improve quantitatively inherited traits (Winter & Kahl, 1995). Molecular markers are, however, not a replacement for traditional resistance breeding methods, but can assist in the combination of resistance genes in a single plant (gene pyramiding) (Kelly, 1995), which is almost impossible with conventional breeding methods (Winter & Kahl, 1995).

2.7.2 Molecular markers

Molecular techniques have been used to identify markers linked to leaf rust resistance genes. Markers that identify leaf rust resistance genes for breeding purposes, include markers linked to genes *Lr9* (Schachermayr *et al.*, 1994), *Lr24* (Schachermayr *et al.*, 1995; Dedryver *et al.*, 1996), *Lr25* (Procunier *et al.*, 1995), *Lr28* (Gupta *et al.*, 1998; Naik *et al.*, 1998), *Lr29* (Procunier *et al.*, 1995), *Lr32* (Autrique *et al.*, 1995), *Lr35* (Seyfarth *et al.*, 1999), *Lr37* (Troskie *et al.*, 1999), *Lr41* (Lottering *et al.*, 1999) and *Lr47* (Helguera & Khan, 2000) (Table 2.1).

Molecular markers in a breeding program are most effective when fast detected (Penner *et al.*, 1996), with a linkage distance of 5 centimorgans (cM) or closer to the gene of interest (Kelly, 1995). Linkage distances are measured in centimorgans and

one centimorgan is defined as one recombination event in a hundred meioses (Winter & Kahl, 1995). The distance of markers from genes of interest can be calculated using the formulas proposed by Kosambi (1944). The recombination fraction of the marker is used to calculate the distance of the marker to the gene in centimorgans using the formula:

$$X = ((1/4) \ln [(1+2c)/(1-2c)]) \times 100$$

X is the distance of the marker to the gene in centimorgans and c is the recombination fraction of the marker. The recombination fraction (c) is the number of recombinations in a segregating population divided by the total number of individuals screened for the marker. Markers tightly linked to an important gene can be used to facilitate selection in breeding programs and as a starting point in the physical location of such genes (Tanksley et al., 1989; Martin & Harvey 1994).

Molecular markers can be either linked to the gene in the coupling phase, or in repulsion phase (Haley et al., 1994). Coupling phase markers refers to markers linked to a gene of interest in a segregating population, while repulsion phase markers identifies susceptible individuals in a segregating population, or individuals without a specific gene. Repulsion phase markers are important where it is necessary to differentiate between homozygotes and heterozygotes for a specific gene (Haley et al., 1994). The advantage of repulsion phase markers is the high selection efficiency. Selecting against a repulsion phase marker linked at 10 centimorgan is more efficient than selecting for a coupling phase marker with a linkage distance of 1 centimorgan, because selection will be limited only for homozygous resistant plants (Kelly et al., 1994).

The main advantage of using molecular markers in plant breeding is the speed at which new plant varieties can be released. This occurs because markers connect phenotypic characters directly with the genes responsible (Winter & Kahl, 1995). Gene introgression is also reduced to three generations if DNA marker technology is used to select for offspring with the least amount of donor plant DNA (Tanksley et al., 1989). One of the best examples where time is saved due to MAS is with Gall midge resistance breeding in wheat. Gall midge resistance can only be tested in the presence of the insect, which occurs only in a two to four month period each year (Nair et al., 1995). With the aid of molecular markers breeding programs can

continue throughout the year. Another time saving application of marker assisted selection with molecular markers is the fact that seed level tests can reveal the phenotype of seeds, while the remainder of a seed can be retained for germination (Chunwose *et al.*, 1993). Molecular markers have another advantage in selecting for genes without using any pathogenic tests (Kelly, 1995). Selection for resistance genes can therefore be carried out in areas where pathogens don't occur naturally. The most important application of molecular markers is, however, to facilitate gene pyramiding. When more resistant genes are added to already resistant plants, molecular markers are essential, because it is difficult to detect gene combinations with traditional breeding methods (McIntosh *et al.*, 1995).

2.7.3 *Bulk segregant analysis to identify molecular markers*

Bulk segregation analysis is one method of identifying molecular markers linked to single resistant genes. The procedure involves the combination of ± 10 resistant and susceptible F_2 individuals in two different DNA samples, a susceptible and resistant bulk respectively, from a population that segregates for a monogenic trait. Polymorphisms obtained with molecular techniques between the resistant and susceptible bulks can be tested as markers linked to the gene of interest (Michelmore *et al.*, 1991).

2.7.4 *Near isogenic lines to identify molecular markers*

The identification of molecular markers linked to resistant genes follows phenotypic classification of plants. Markers can be identified with molecular techniques performed on near isogenic lines (NILs). NILs are theoretically identical. The only difference is the gene of interest and a small piece of DNA surrounding the gene that is introgressed with the gene (Weeden *et al.*, 1994). Polymorphic DNA markers between the NILs can be tested for linkage to resistance genes (Weeden *et al.*, 1994). A set of Thatcher near isogenic lines with single genes for leaf rust resistance genes is used throughout the world for marker identification (Samborski, 1985).

2.7.5 General approaches to locate molecular markers linked to genes

Phenotypic classifications have to be performed on a segregating population for a monogenic trait. DNA isolations have to be done prior to the inoculation of the population with *P. recondita* f. sp. *tritici* in order to obtain pure wheat DNA for linkage analysis, because the same individuals have to be used for both phenotypic tests and linkage analysis (Winter & Kahl, 1995). The DNA obtained from the population prior to inoculations can either be bulked (Michelmore *et al.*, 1991), or molecular techniques can be performed on NILs (Weeden *et al.*, 1994). Polymorphisms obtained between resistant and susceptible plants using either the near isogenic lines or the bulk segregant analysis approach, can then be tested for markers linked to the gene. A disadvantage is that resistant plants due to dominant genes consist of homozygotes and heterozygotes throughout the different plant generations and the difference cannot be estimated with either NILs nor DNA bulks. The genotype of individuals when working with dominant genes can therefore not be revealed with dominant natured molecular techniques or with phenotypic results of a segregation analysis (Winter & Kahl, 1995).

Chapter 3

A comparative genetic analysis using RAPDs, AFLPs and SSRs, to predict the best combinations of different genotypes, with the emphasis on leaf rust resistance gene *Lr37*

3.1 *Introduction*

The domestication of wheat has caused a reduction in genetic variation (Feldman & Sears, 1981). This was demonstrated by the low level of polymorphism detected among different wheat cultivars (Gale *et al.*, 1990; Devos & Gale, 1992). The narrow genetic variation in wheat has the disadvantage of increasing susceptibility to disease and climatic changes (Plucknett *et al.*, 1983). Wheat breeders are therefore, continuously looking for ways to broaden the genetic base of wheat. Accurate identification of germplasm is, however, important to protect plant breeder's rights and to accelerate wheat breeding programs (Myburg *et al.*, 1997).

DNA fingerprinting techniques such as RAPDs (Welsh & McClelland, 1990; Williams *et al.*, 1990; Myburg *et al.*, 1997), AFLPs (Zabeau & Vos, 1993) and SSRs (Litt & Luty, 1989) can be used for cultivar identification and genetic distance analysis between different cultivars (Yu *et al.*, 1993; Nybom, 1994; Myburg *et al.*, 1997). RAPDs have the advantage of requiring small amounts of DNA, no radioactive labeling and require only a quick assay (Welsh & McClelland, 1990; Williams *et al.*, 1990). AFLPs on the other hand, provide a large number of loci without any prior sequence information that can be visualized on a single gel (Vos *et al.*, 1995). SSRs have a multi allelic nature, an advantage in evolutionary studies (Schloetterer *et al.*, 1991; Buchanan *et al.*, 1994).

RAPDs have been used to differentiate between Southern African wheat cultivars and to estimate genetic distances between cultivars in order to identify cultivars most suitable for NIL development in Russian wheat aphid resistance breeding programs (Myburg *et al.*, 1997). AFLPs, SSRs and RFLPs have recently been used to estimate genetic similarities between German and Austrian winter wheat cultivars in order to predict the viability of cultivar crosses and of these, AFLPs were proven to be the most effective (Bohn *et al.*, 1999). These three molecular techniques; RAPDs, AFLPs and SSRs can therefore be used to estimate genetic distances among wheat cultivars.

Like the Russian wheat aphid program, leaf rust breeding programs are essential in Southern Africa. Recent yield losses due to leaf rust outbreaks in Southern Africa

includes an outbreak on Senga wheat in Zimbabwe and an estimated 40 % yield loss in the western Cape (Pretorius *et al.*, 2000). The aim of this study was to predict combining abilities of the wheat genotypes VPM1 and RL6081, which both contain the adult-plant leaf rust resistance gene *Lr37* (Kolmer, 1996), with local South African cultivars. The estimation of combining abilities will be based on genetic distances among cultivars calculated from molecular data obtained using the techniques RAPDs, AFLPs and SSRs. The genetic distance calculated with molecular techniques between the near isogenic genotypes Thatcher and RL6081, will also be an indication of how many amplified loci, need to be compared between the lines before polymorphisms and possible markers for gene *Lr37* can be obtained, when using the near isogenic line approach.

3.2 *Materials and methods*

3.2.1 *Plant material*

Seed from different wheat (*Triticum aestivum*) genotypes, rye (*Secale cereale*) and Kiewiet (*x Triticosecale*), were obtained from the Small Grain Institute Bethlehem and from the Department of Plant Pathology of the University of the Orange Free State, Bloemfontein, South Africa (Table 3.1). Genotypes included in the study were rye and the triticosecale cultivar Kiewiet used as outgroups, South African wheat genotypes resistant and susceptible to the Russian wheat aphid, as well as sources of resistance from the southern USSR and the Middle East. Also, South African genotypes resistant and susceptible to leaf rust, as well as sources of resistance to leaf rust resistance genes from France (VPM1) and Canada (RL6081 (Thatcher*8/VPM1)), together with the susceptible near isogenic genotype Thatcher, initially used as recurrent parent to develop line RL6081 (Dyck, 1992).

The seeds were surface sterilized in 70 % ethanol, placed in Petri dishes between wet filter paper sheets and allowed to germinate at 25 °C. After a week the seedlings were removed, from the still attached remainders of the seed, for DNA isolations. The seed remains were removed, because the endosperm is not representitive of the 'new'

individual and to prevent high levels of carbohydrates in the DNA sample, which can interfere with further DNA analysis.

Table 3.1 The 21 different genotypes used to conduct a molecular genetic analysis in the study.

Genotype	Species	Genome	Origin	Resistance gene
Palmiet	*Triticum aestivum*	AABBDD	SA	*Lr24*
Tugela	*T. aestivum*	AABBDD	SA	*Lr34*
Molopo	*T. aestivum*	AABBDD	SA	
Tugela*5/SA1684	*T. aestivum*	AABBDD	SA	*Dn1*
Gamtoos*4/SA463	*T. aestivum*	AABBDD	SA	*Dn5*
Betta	*T. aestivum*	AABBDD	SA	
Palmiet*4/SA463	*T. aestivum*	AABBDD	SA	*Dn5*
SA2199/Letaba	*T. aestivum*	AABBDD	SA	*Dn2*
Molopo*4/SA1684	*T. aestivum*	AABBDD	SA	*Dn1*
Kiewiet	*x Triticosecale*	AABBRR	SA	
Rye	*Secale cereale*	RR		
SST65	*T. aestivum*	AABBDD	SA	
SST86	*T. aestivum*	AABBDD	SA	
SST825	*T. aestivum*	AABBDD	SA	
Inia	*T. aestivum*	AABBDD	Mexico	*Dn4*
Corwa	*T. aestivum*	AABBDD	USA	*Dn4*
SA1684	*T. aestivum*	AABBDD	Iran	*Dn1*
SA2199	*T. aestivum*	AABBDD	Former USSR	*Dn2*
Thatcher	*T. aestivum*	AABBDD	USA	*Lr22a, b*
RL6081 (Thatcher*8/VPM1)	*T. aestivum*	AABBDD	Canada	*Lr37*
VPM1	*T. aestivum*	AABBDD	France	*Lr37*

3.2.2 *DNA isolation*

DNA isolations were performed according to a modified method of Myburg *et al.* (1998). One gram of leaf tissue was ground to a fine powder in liquid nitrogen, before 500 µL preheated (65 °C) isolation buffer (0.5 M NaCl, 0.1M Tris-HCl pH 8, 0.04 M EDTA, 0.043 M SDS, 0.02 M NaBisulfite) was added. The samples were then vortexed and incubated at 65 °C for 30 minutes. After incubation, 500 µL chloroform: isoamylalcohol (24:1) was added before each sample was vortexed and centrifuged at 12 000 x *g* for 15 minutes. Two volumes of 100 % ethanol were then

added to the upper phase of the samples in order to precipitate the dissolved DNA. The precipitated DNA was scooped from the ethanol solution with a sterile pipette tip, washed with 70 % ethanol, air dried and dissolved in 200 µl sterile water. The DNA concentrations were determined spectrofotometrically and diluted to concentrations of 3.75 ng/µl for RAPDs, 100 ng/µl for AFLPs and 10 ng/µl for SSRs.

3.2.3 *RAPD-PCR analysis*

RAPD-PCR amplifications were performed according to Myburg *et al.* (1998). A 25 µl reaction mixture was used containing: 7.5 ng template DNA, 1 x PCR buffer (20 mM $(NH_4)_2SO_4$, 75 mM Tris-HCL pH 9.0, 0.1 % (v/v) Tween), 269 µM dNTPs, 0.0004% (m/v) gelatin, 3 mM $MgCl_2$, 0.3 pmol 10-mer primer [OPA19, OPB13 or OPG13 (Operon Technologies, Almeda, California, USA)] and 0.8 U of *Taq* DNA polymerase (Promega). Primers OPA19, OPB13 and OPG13 were previously shown to amplify wheat DNA, with OPA19 the most polymorphic and also included in the selection key to distinguish among different South African wheat cultivars (Myburg *et al.*, 1997). Samples were overlaid with oil and amplified using a thermal cycler (Hybaid$_{TM}$ Omnigene model TR3CM220). The following temperature profile was used: 2.3 min at 94 °C, followed by 35 cycles of 45 s at 35 °C, 90 s at 72 °C, 30 s at 94 °C and a final step of 7 min at 72 °C. Amplification products were separated on 2 % agarose (Seakem, USA) gels containing ethidium bromide (0.5 µg/ml) at 3.5 V/cm in 0.5 x TBE buffer (22.3 mM Tris, pH 8.0, 22.2 mM Boric acid, 0.5 mM EDTA) and visualized under UV light.

3.2.4 *AFLP analysis*

Genomic DNA was digested with restriction enzymes *Eco*RI and *Mse*I as initial DNA template preparation. Digestion reactions with a total volume of 20 µl consisted of 1 µg of genomic DNA, 1 x NE buffer (50 mM NaCl, 10 mM Tris-HCl pH 7.9, 10 mM $MgCl_2$, 1 mM DTT), 20 ng BSA, 1.5 U *Mse*I, 1 x buffer H (50 mM Tris-HCl pH 7.5, 100 mM NaCl, 10mM $MgCl_2$ and 1 mM DTT) and 5 U of *Eco*RI. Digestions were performed at 37 °C for 4 h. *Eco*RI and *Mse*I adapters were then ligated to the set of restricted fragments in a 20 µl reaction volume consisting of 14.5 µl of digested DNA,

0.5 x T4 DNA ligase buffer (Boehringer Mannheim, Germany), 0.025 M MgCl, 1 U T4 DNA ligase, 50 ng BSA and an excess of *Eco*RI and *Mse*I adapters (Perkin-Elmer, USA). Ligation reactions were incubated at 25 °C overnight.

Pre-amplification was performed in a 25 µl reaction volume containing: 1 x PCR buffer (20 mM $(NH_4)_2SO_4$, 75 mM Tris-HCL pH 9.0, 0.1 % (v/v) Tween), 269 µM dNTPs, 0.0004% (m/v) gelatin, 3 mM $MgCl_2$, 0.3 pmol pre-selective amplification primer pair (Perkin-Elmer, USA), 0.8 U *Taq* DNA polymerase (Promega, USA) and 2 µl diluted restriction-ligation template DNA. The pre-selective amplification products were diluted 10 fold, before a selective amplification was performed.

The selective amplification was performed in a 25 µl reaction volume containing: 1 x PCR buffer (20 mM $(NH_4)_2SO_4$, 75 mM Tris-HCL pH 9.0, 0.1 % (v/v) Tween), 269 µM dNTPs, 0.0004% (m/v) gelatin, 3 mM $MgCl_2$, 0.15 pmol of each selective amplification primer pair (E-AAG: M-CAG, E-ACA: M-CTC and E-AGC: MCTC) (Perkin-Elmer, USA), 0.8 U *Taq* DNA polymerase (Promega, USA) and 2 µl diluted pre-selective amplification products as template. Selective amplification products were separated and analyzed on an ABI PRISM™ 377 DNA sequencer, using GeneScans (Perkin- Elmer, USA).

3.2.5 *SSR analysis*

SSR amplifications were performed according to the technique described by Lee *et al.* (1995). Five primers that amplify regions from chromosome 2AS were selected randomly (Röder *et al.*, 1998.). A 25 µl reaction contained 20 pM of the primer pair [XGWM122 (forward 5' GGG TGG GAG AAA GGA GATG 3' and reverse 5' AAA CCA TCC TCC ATC CTG G 3') or XGWM337 (forward 5' CCT CTT CCT CCC TCA CTT AGC 3' and reverse 5' TGC TAA CTG GCC TTT GCC 3')], 20 ng DNA, 800 µM dNTPs, 1 U *Taq* polymerase, 1 x *Taq* buffer (Promega, USA) and 1.5 mM $MgCl_2$. The following temperature profile was used: 40 cycles at 94 °C for 1 min, 65 °C for 1 min and 72 °C for 2 min and a final cycle at 72 °C for 10 min. SSR amplification products were separated on six percent nondenaturing polyacrylamide gels (Sambrook *et al.*, 1989).

The visualization of SSR amplification products on the gels was performed, using a modified silver staining method (Bassam *et al.*, 1991). Silver staining of acrylamide gels started with fixation for 30 minutes in a 10 % acetic acid: 30 % ethanol solution. Two wash steps of 10 minutes each in 10 % ethanol and two rinse steps of 20 minutes each with UHQ water followed the fixation step. An impregnation step of 30 minutes in a 5.917 mM $AgNO_3$ solution, containing 7.5 ml of 37 % HCOH per liter UHQ water followed the rinse steps. A quick rinse of 20 s in UHQ water followed the impregnation step. The final development of the gels were performed for 2-5 minutes in a solution containing 285.714 mM Na_2CO_3, 1.5 ml 37 % HCOH for each 200 ml solution and 2.848 mM $Na_2S_2O_3.5H_2O$ in UHQ water. Development was consequently stopped with a solution containing 10 % acetic acid in UHQ water (Bassam *et al.*, 1991).

3.2.6 *Genetic distance analysis among wheat genotypes*

Amplification products of three selected RAPD primers, three selected AFLP primer pairs and two selected SSR primer pairs were considered independent loci. A data matrix consisting of a binary data set (1 = present and 0 = absent), of these independent amplified loci were compiled for all the amplified loci, of the genotypes used in the study (Appendix 3.1). Pairwise genetic distances among the different genotypes were further calculated, using the formula $F = 2n_{xy}/(n_x + n_y)$ (Nei & Li, 1979). F is the index of genetic similarity and therefore the ratio of shared loci between individuals x and y. $2n_{xy}$ equals the number of shared loci, and n_x and n_y are the number of loci for each individual (Wang & Tanksley, 1989). Distance values were then calculated with the formula $D = 1 - F$.

Unweighted pair-group mean arithmetic analysis (UPGMA) was further used to construct a cladogram from the binary data matrix. This represents the different groups of the genotypes used in the study. UPGMA analysis was also used to group the pairwise genetic distances, in order to generate a dendogram that represents the genetic distances among different genotypes used in the study. All calculations and the construction of the dendograms were performed using UPGMA analysis with the aid of PAUP4 (Swofford, 1991).

3.3 Results

3.3.1 RAPD analysis

DNA isolations from the seedlings resulted in genomic DNA concentrations of 0.5-1 $\mu g/\mu L$. RAPD-PCR amplifications performed with a dilution of the genomic DNA to concentrations of 7.5 ng/μl resulted in 27 scorable loci when using primer OPA19 (Fig. 3.1), 25 scorable loci with primer OPB13 (Fig. 3.2) and 16 scorable loci with OPG13 (Fig. 3.3). The total loci amplified with the three RAPD primers were therefore 68, with 90 % of the loci polymorphic among the 21 genotypes tested and only 10 % monomorphic. The average number of loci amplified when using the three primers was 19 (Appendix 3.1). Unique loci amplified for rye include: $OPB13_{235}$, $OPB13_{580}$ and $OPG13_{350}$. $OPA19_{214}$ was uniquely amplified in Kiewiet and one of the wheat cultivars, Molopo could be distinguished from the rest of the genotypes with amplified loci $OPB13_{250}$ and $OPB13_{200}$. VPM1 and RL6081 containing leaf rust resistance gene *Lr37* and the susceptible cultivar Thatcher could be distinguished from the other genotypes with $OPA19_{3904}$, $OPG13_{2778}$ and $OPG13_{2027}$. RAPD amplification products proved therefore to be highly polymorphic among the 21 genotypes compared in the study.
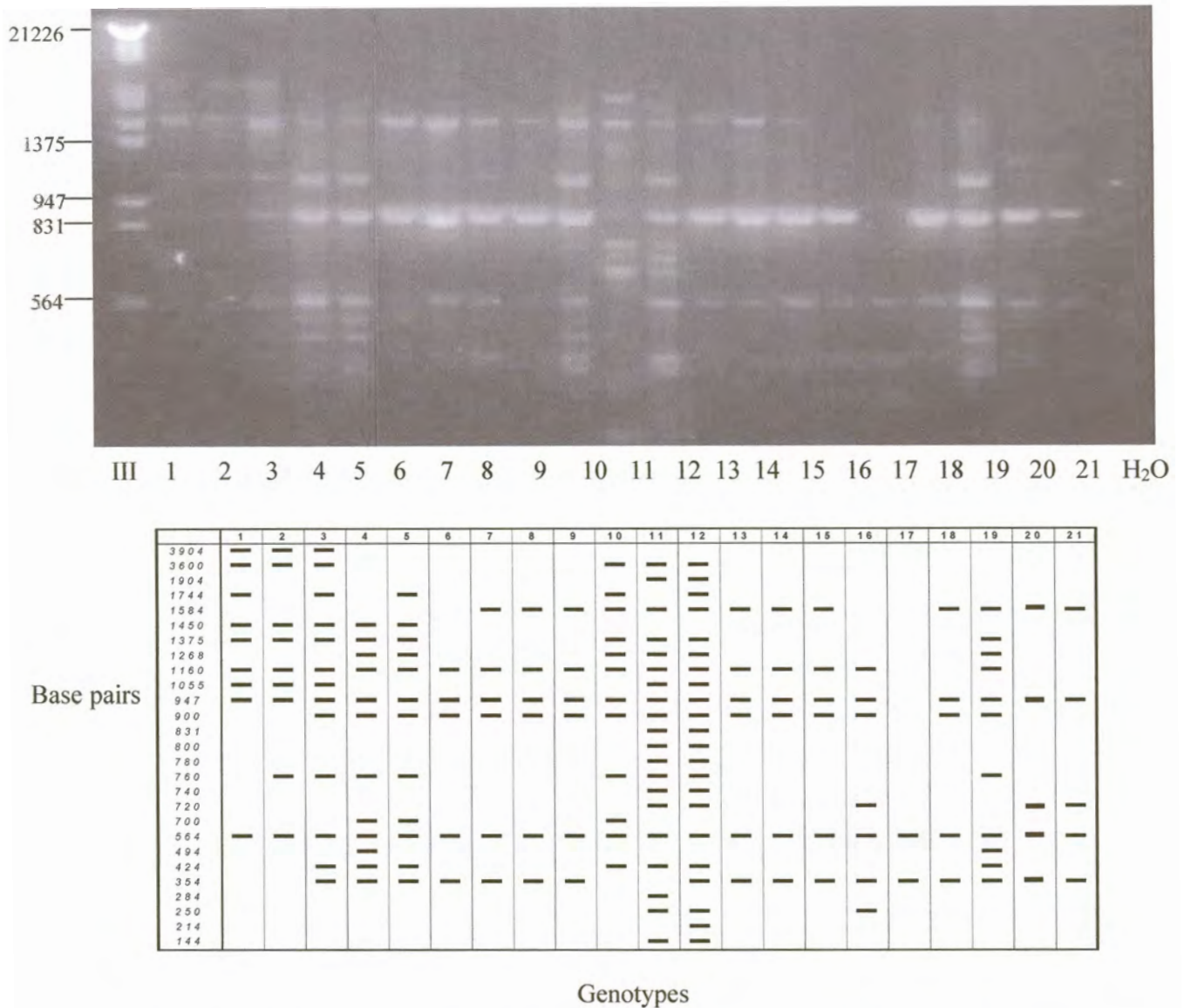
Fig 3.1 RAPD amplification profiles obtained for the 21 genotypes with primer OPA19 after separation on 2 % agarose gels. Lane 1 = VPM1, Lane 2 = RL6081, Lane 3 = Thatcher, Lane 4 = SA2199, Lane 5 = SA1684, Lane 6 = Corwa, Lane 7 = Inia, Lane 8 = SST825, Lane 9 = SST86, Lane 10 = SST65, Lane 11 = Rye, Lane 12 = Kiewiet, Lane 13 = Molopo*4/SA1684, Lane 14 = SA2199/Letaba, Lane 15 = Palmiet*4/SA463, Lane 16 = Betta, Lane 17 = Gamtoos*4/SA463, Lane 18 = Tugela*5/SA1684, Lane 19 = Molopo, Lane 20 = Tugela, Lane 21 = Palmiet and Lane 22 = control. The first lane contains a molecular size marker (λDNA digested with *EcoRI* and *Hind*111). The last lane contains a H_2O control.

41

21226 —

1375 —

947 —
831 —

564 —

III 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 H₂O
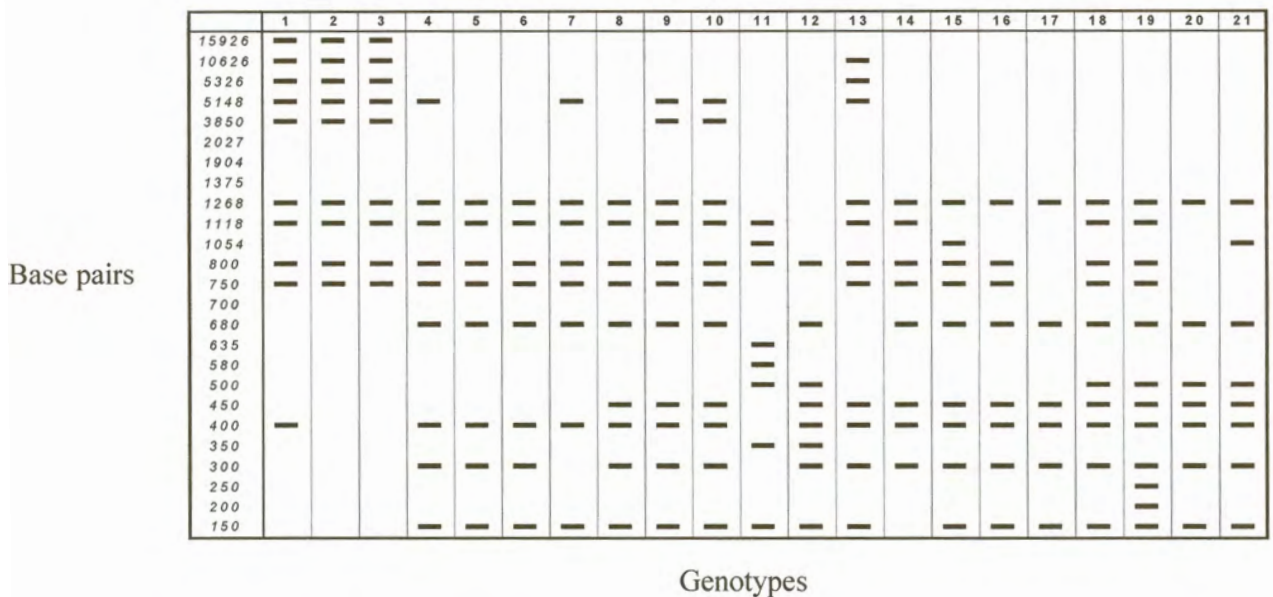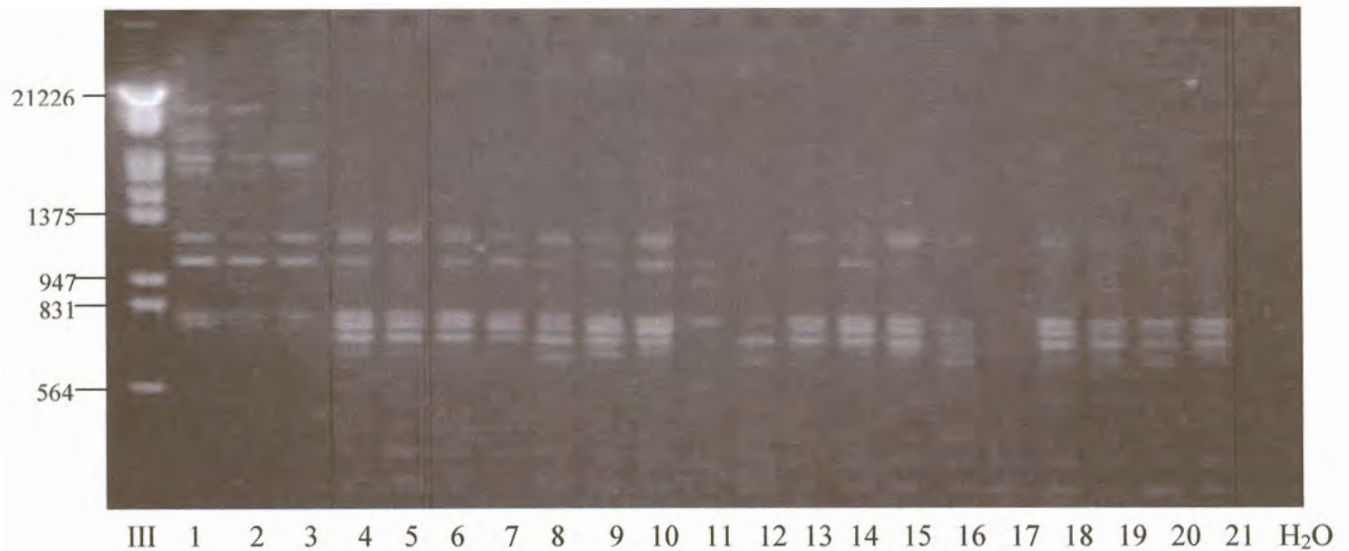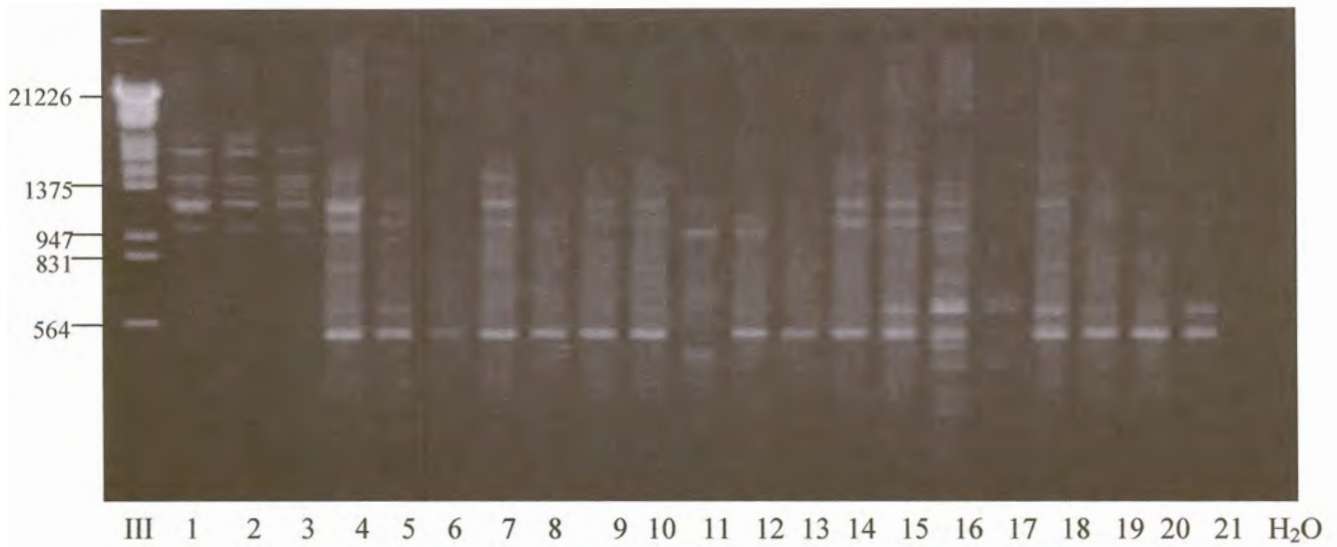
Base pairs

Genotypes

Fig 3.2 RAPD amplification profiles obtained for the 21 genotypes with primer OPB13 after separation on 2 % agarose gels. Lane 1 = VPM1, Lane 2 = RL6081, Lane 3 =Thatcher, Lane 4 = SA2199, Lane 5 = SA1684, Lane 6 = Corwa, Lane 7 = Inia, Lane 8 = SST825, Lane 9 = SST86, Lane 10 = SST65, Lane 11 = Rye, Lane 12 = Kiewiet, Lane 13 = Molopo*4/SA1684, Lane 14 = SA2199/Letaba, Lane 15 = Palmiet*4/SA463, Lane 16 = Betta, Lane 17 = Gamtoos*4/SA463, Lane 18 = Tugela*5/SA1684, Lane 19 = Molopo, Lane 20 = Tugela, Lane 21 = Palmiet and Lane 22 = control. The first lane contains a molecular size marker (λDNA digested with *EcoRI* and *Hind*111). The last lane contains a H₂O control.

42

Molecular size marker (lane III): 21226, 1375, 947, 831, 564

Gel lanes: III 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 H₂O

Base pairs schematic (Genotypes 1–21):

Base pairs	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
2778	▬	▬	▬																		
2020	▬	▬	▬																		
1800	▬	▬	▬	▬			▬	▬	▬	▬				▬	▬	▬		▬			
1584	▬	▬	▬	▬			▬	▬	▬	▬				▬	▬	▬		▬	▬		
1375									▬	▬											▬
1200	▬	▬	▬		▬		▬	▬	▬		▬	▬	▬	▬	▬	▬				▬	
1050	▬	▬	▬		▬		▬	▬	▬		▬	▬	▬	▬	▬	▬		▬		▬	▬
900	▬	▬	▬	▬	▬	▬	▬	▬	▬		▬	▬	▬	▬	▬	▬	▬	▬	▬	▬	▬
831	▬		▬													▬					
700				▬	▬		▬	▬	▬					▬	▬			▬	▬		
650				▬	▬										▬						
600						▬	▬				▬	▬	▬	▬			▬	▬			
540	▬	▬	▬																		
350				▬	▬	▬	▬	▬	▬	▬		▬	▬	▬	▬	▬		▬	▬	▬	▬
250											▬						▬				
150				▬			▬	▬	▬							▬	▬		▬	▬	▬

Genotypes

Fig 3.3 RAPD amplification profiles obtained for the 21 genotypes with primer OPG13 after separation on 2 % agarose gels. Lane 1 = VPM1, Lane 2 = RL6081, Lane 3 = Thatcher, Lane 4 = SA2199, Lane 5 = SA1684, Lane 6 = Corwa, Lane 7 = Inia, Lane 8 = SST825, Lane 9 = SST86, Lane 10 = SST65, Lane 11 = Rye, Lane 12 = Kiewiet, Lane 13 = Molopo*4/SA1684, Lane 14 = SA2199/Letaba, Lane 15 = Palmiet*4/SA463, Lane 16 = Betta, Lane 17 = Gamtoos*4/SA463, Lane 18 = Tugela*5/SA1684, Lane 19 = Molopo, Lane 20 = Tugela and Lane 21 = Palmiet. The first lane contains a molecular size marker (λDNA digested with *EcoRI* and *Hind*111). The last lane contains a H₂O control.

3.3.2 AFLP analysis

The AFLP primer pair *Eco*RI-AAG and *Mse*I-CAG resulted in 18 amplified loci, with electropherogram amplitudes higher than 25 units, when using GeneScans (Perkin-Elmer, 1996). The primer pair *Eco*R1-AAG and *Mse*I-CAG amplified 20 loci and the primer pair *Eco*RI-ACA and *Mse*I-CTC amplified 30 loci, with electropherogram amplitudes higher than 25 units. The average amount of loci amplified with the three primer pairs, with amplitude higher than 25 units were, thus 23 (Appendix 3.1), less than half the amount of all the loci amplified for each primer. A total of 68 loci were therefore scored with the three AFLP primers. Seventy eight percent of the loci scored were polymorphic and 22.1 % were monomorphic among the 21 compared genotypes. Rye could be distinguished from the wheat genotypes and Kiewiet with loci E-AAG: M-CAG_{87}, E-AAG: M-CAG_{279}, E-ACA: M-CTC_{102}, E-ACA: M-CTC_{102}, E-ACA: M-CTC_{105}, E-ACA: M-CTC_{129} and E-ACA: M-CTC_{170}. Kiewiet could be distinguished from the other genotypes with E-AAG: M-CAG_{295}. Palmiet*4/SA463 could be distinguished with E-AAG: M-CAG_{89}. The cultivar Thatcher could be distinguished from the other genotypes with loci E-AAG: M-CAG_{101}. SST65 could be distinguished from the other genotypes with amplified locus E-AGC: M-CTC_{106}. The cultivar Molopo could be distinguished from the other genotypes with amplified locus E-AGC: M-CTC_{257} and SA1684 could also be distinguished from the other genotypes with amplified locus E-ACA: M-CTC_{168}. AFLP amplification products also proved to be highly polymorphic, although less polymorphic than the loci amplified with the three RAPD primers, mainly because highly polymorphic RAPD primers were selected from a study previously conducted, involving the Russian wheat aphid (Myburg *et al.*, 1997), while AFLP primers were selected randomly.

3.3.3 SSR analysis

SSR primer pairs, XGWM122 (Fig 3.4) and XGWM337 amplified six scorable loci. A total of 12 amplified loci with the two primers resulted in an average of six loci per primer pair (Appendix 3.1). All 12 loci amplified with SSRs showed polymorphism among the 21 compared genotypes. Rye could be distinguished from the other

44

genotypes with amplified loci $XGWM377_{150}$. SSR amplified loci proved to be highly polymorphic with all the loci polymorphic among the 21 different genotypes (Appendix 3.1).
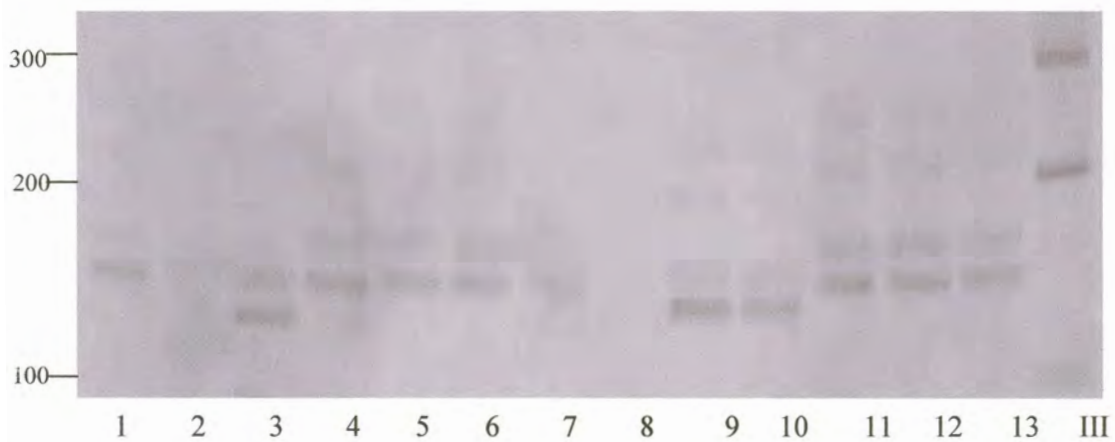

Fig 3.4 A silver stained, 6 % polyacrylamide gel showing SSR amplification products of different genotypes, with primer pair XGWM122. Lane 1 = Tugela*5/SA1684, Lane 2 = Gamtoos*4/SA463, Lane 3 = Betta, Lane 4 = Palmiet*4/SA463, Lane 5 = SA2199/Letaba, Lane 6 = Molopo*4/SA1684, Lane 7 = Kiewiet, Lane 8 = rye, Lane 9 = SST65, Lane 10 = SST86, Lane 11 = SST825, Lane 12 = Inia, Lane 13 = Corwa. The last lane contains a 100 base pair molecular size marker.

3.3.4 Dendogram

A dendogram based on pair wise genetic distances (Fig 3.5), resulted in rye being genetically distinct from the wheat genotypes, with an average distance converted to a percentage of 11.11 % (Nei & Li, 1979) (Table 3.2). Rye and Kiewiet are 5.4 % genetically distinct (Table 3.2), half the average genetic distance when rye is compared with wheat in this study. The average pairwise genetic distance between an entry grouped in the middle of the wheat group (Corwa) and the rest of the wheat genotypes is 3.1 % (Table 3.2). The *Lr37* line, RL6081, is further, on average 5.4 % genetically different from the South African cultivars Palmiet, Tugela, Molopo, SST65, SST825 and SST86 (Table 3.2). The two *Lr37* carrying genotypes RL6081 and VPM1 are closest to the South African cultivar Betta, with pairwise genetic distances converted to a percentage of 3.6 % and 3.7 %, respectively (Table 3.2). RL6081 and VPM1 are furthest apart from the cultivar Molopo with a distance of 7.4 % and 6.5 %, respectively (Table 3.2).

3.3.5 Cladogram

The cladogram based on the 148 independent loci, amplified with RAPDs, AFLPs and SSRs confirmed the two main groups obtained with the dendogram (Fig. 3.6). Firstly, Kiewiet and rye grouped out with a bootstrap value of 87. The wheat genotypes formed the second main clade, with a bootstrap value of 95. The genotypes Thatcher, VPM1 and RL6081 formed a distinguishable secondary clade within the wheat clade, with a bootstrap value of 99. The genotypes Thatcher and RL6081 grouped together with a bootstrap value of 66 in this secondary clade. A second secondary clade within the wheat clade consisted of South African, former USSR and Iran genotypes, with a bootstrap value of 72. The cultivars SA2199 (USSR) and SA1684 (Iran) grouped together within the second secondary clade with a bootstrap value of 56. SA463/ *4Gamtoos and Corwa also grouped together in the second secondary clade with a bootstrap value of 64. The cladogram complemented the dendogram overall and shows an additional clade, within the main wheat clade, with the two Russian wheat aphid resistant genotypes SA1684 and SA2199 from Iran and the former USSR, respectively (Fig 3.6).

Table 3.2 Pairwise genetic distances among the 21 genotypes were calculated from the binary data set according to Nei & Li (1979) and converted to a percentage,.

|	Palmiet	Tugela	Molopo	Tugela*5/SA1684	Gamtoos*4/SA463	Betta	Palmiet*4/SA463	SA2199/Letaba	Molopo*4/SA1684	Kiewiet	Rye	SST65	SST86	SST825	Inia	Corwa	SA1684	SA2199	Thatcher
Palmiet	-																		
Tugela	2.9	-																	
Molopo	4.1	3.7	-																
Tugela*5/SA1684	2.9	2.8	2.7	-															
SA 463/*4Gamtoos	2.6	2.2	4.4	2.9	-														
Betta	3.4	3.0	4.0	2.2	1.7	-													
Palmiet*4/SA463	2.8	2.3	3.7	1.9	2.4	1.7	-												
SA2199/Letaba	2.3	2.7	3.4	2.0	2.3	1.6	1.4	-											
Molopo*4/SA1684	3.6	2.6	3.7	2.4	2.4	1.9	1.7	1.6	-										
Kiewiet	6.3	4.7	8.5	6.8	5.8	5.3	5.8	6.1	5.6	-									
Rye	10.6	10.1	15.1	12.8	11.0	10.5	10.6	10.0	10.8	5.4	-								
SST65	4.9	3.4	4.4	4.0	5.0	4.0	3.4	3.7	3.7	6.3	12.5	-							
SST86	3.4	2.8	3.7	2.4	3.4	2.4	2.1	2.3	2.1	6.1	12.1	2.3	-						
SST825	4.1	3.4	3.6	2.0	3.3	2.0	2.0	2.2	2.5	5.5	11.4	3.4	2.0	-					
Inia	4.1	3.4	4.5	2.8	3.3	2.3	2.2	3.0	2.0	5.0	10.1	4.2	2.0	1.8	-				
Corwa	3.7	3.0	4.4	2.6	1.6	2.1	2.5	2.2	1.8	5.2	10.6	4.3	2.6	2.7	2.7	-			
SA1684	4.4	3.2	5.8	4.1	3.6	3.8	3.5	3.0	3.8	5.6	9.8	3.9	4.3	3.4	3.7	3.6	-		
SA2199	4.7	3.5	6.1	3.6	4.2	3.1	3.0	2.5	3.1	5.0	9.1	3.7	3.3	3.0	1.8	3.6	2.0	-	
Thatcher	5.7	4.4	6.1	4.5	5.2	3.8	4.2	3.7	3.3	6.4	10.5	4.4	4.2	4.6	4.3	4.6	4.2	4.2	-
RL6081	6.2	5.6	7.3	4.7	5.4	3.6	4.6	4.3	3.3	7.0	10.8	5.1	3.8	4.8	4.5	4.2	5.2	4.6	1.3
VPM1	6.0	5.2	6.5	4.5	5.2	3.7	3.9	4.1	3.4	7.1	12.2	4.1	3.7	3.8	4.6	4.6	4.7	4.7	1.8

UPGMA
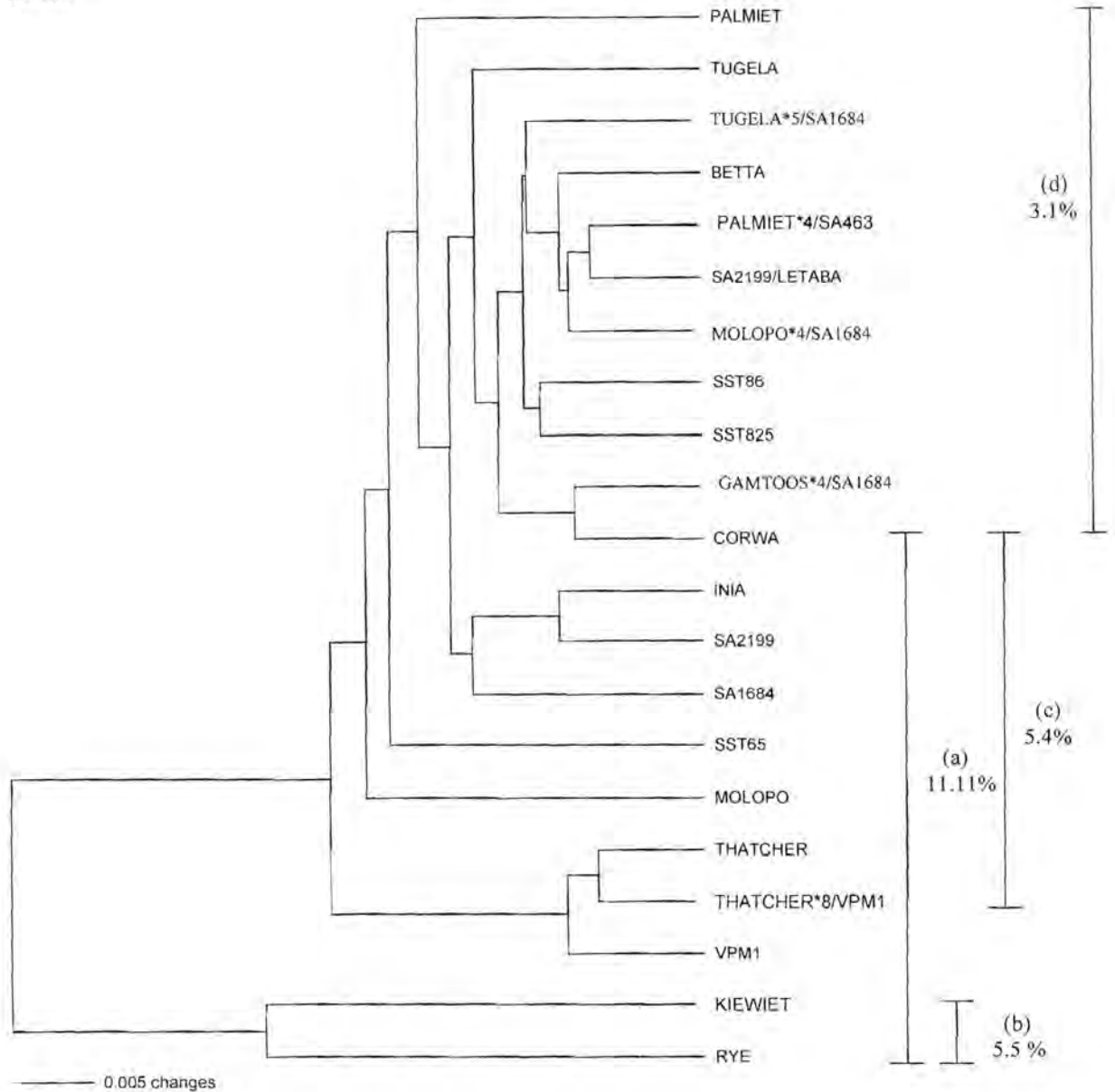
Fig 3.5 A dendogram based on pairwise genetic distances (Nei & Li, 1979), among the different genotypes used in the study. Pairwise genetic distances are calculated using a matrix based on loci amplified with RAPDs, AFLPs and SSRs. (a) The distance between rye and wheat. (b) The distance between rye and Kiewiet. (c) The distance between RL6081 and the SA genotypes. (d) The average distance among wheat cultivars.
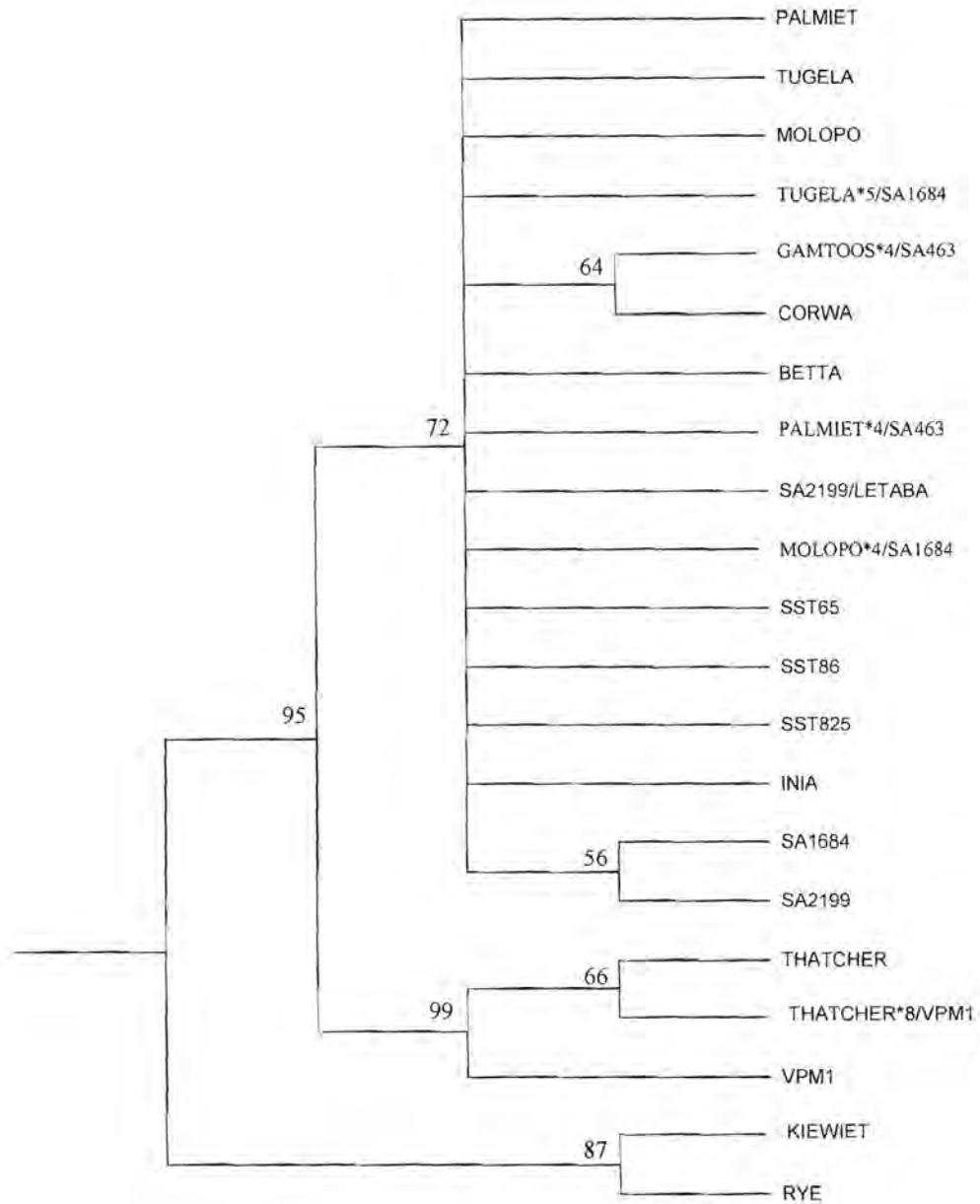
48

Fig 3.6 A cladogram of the different genotypes used in the study based on RAPDs, AFLPs and SSRs. Bootstrap values are indicated as percentages on the different branches. All calculations were performed with the aid of PAUP 4.0 (Swofford, 1991).

3.4 *Discussion*

The three RAPD primers used in the study amplified an average of 19 independent loci. Genotypes that could be distinguished with uniquely amplified RAPD loci include rye, Kiewiet and Molopo. The genotypes contained no known resistance genes of direct interest to the Russian wheat aphid, or leaf rust. The unique loci amplified for Molopo, $OPB13_{200}$ and $OPB13_{250}$, can be used for cultivar identification. Rye and Kiewiet, grouped as outgroups as expected from the other genotypes with uniquely amplified loci, because hexaploid bread wheat has a chromosome composition of AABBDD (Knott, 1989) and rye has a chromosome composition of RR, while the hybrid known as Kiewiet has a chromosome composition of AABBRR. Although the three RAPD primers could distinguish between different genotypes and also the different genome compositions of rye, wheat and Kiewiet, no markers showed any linkage to the resistant genes of interest and more primers should therefore be used to reveal polymorphisms that can be tested as markers linked to these specific resistance genes (Myburg, 1996).

Three AFLP primers amplified an average of 23 loci with amplitudes higher than 25 units, when using Genescans. The loci amplified with amplitudes less than 25 units were discarded to eliminate all possible backgroud amplicons. The 68 loci amplified with the three AFLP primers include highly repeatable loci, with strong Genescan signals. Rye, Kiewiet, Palmiet*4/SA463, Thatcher, SST65, Molopo and SA1684 could be uniquely identified. Polymorphic loci were again amplified for rye and Kiewiet. Palmiet*4/SA463, Thatcher, SST65, Molopo and SA1684 could also be distinguished with polymorphic loci. Palmiet*4/SA463 differs from Palmiet because of the *Dn5* resistance gene. The amplified polymorphic locus in Palmiet*4/SA463 can therefore be submitted to a linkage test against gene *Dn5*, using a F_2 population from a Palmiet and Palmiet*4/SA463 cross. A amplified polymorphic locus also occurs in the cultivar Thatcher, a near isogenic line to RL6081 (Thatcher*8/VPM1). This polymorphic locus in the Thatcher cultivar can thus, be tested for repulsion phase linkage to gene *Lr37*. The three AFLP primers tested proved useful for genotype identification and provided two putative markers linked to gene *Dn5* in the coupling phase and linked in repulsion to gene *Lr37*, respectively.

Two SSR primer combinations both provided six scorable loci. This low number of amplified loci, in comparison with RAPDs and AFLPs are mainly due to the high specificity of the 22 base pair oligomer primers used for SSR amplifications, which makes the markers allele specific and therefore codominant. Four SSR markers were sufficient to discriminate between two closely related barley cultivars (Russell *et al.*, 1997). The high level of polymorphism obtained between the different genotypes, used in the study, with the two primer pairs was therefore not surprising. Rye could be distinguished from the other genotypes, because primer pair $XGWM122_{150}$, amplified a 150 base pair fragment in all the plants except for rye (Fig. 3.4). The 150 base pair fragment is probably amplified from the A or B genomes, which occurs in all the plants except for rye.

The dendogram based on pairwise genetic distances (Nei & Li, 1979), resulted in rye being the furthest from the other genotypes. Rye was on average 11.11 % genetically different from the wheat genotypes and 5.4 % from Kiewiet (Table 3.2). With the wheat genome being AABBDD (Knott, 1989) and the rye genome being only RR, a larger pairwise genetic distance was expected between rye and the average wheat genotype. The pairwise genetic distance between rye and Kiewiet is however, 5.4 %, which is halfway that of rye to the average wheat genotype. The fact that Kiewiet grouped between wheat and rye is unexpected, because Kiewiet contains a genome composition, 66.6 % similar to wheat and only 33 % similar to rye. It is therefore expected for Kiewiet to be two thirds different from rye and only a third different from wheat. The unexpected shift in genetic distance of the Kiewiet towards the middle of rye and wheat is, because not all the loci amplified in Kiewiet were present in the parents (neither rye nor wheat). If all the loci amplified in Kiewiet were present in wheat or rye, the expected genetic similarity of Kiewiet to wheat (33.3 %) would have been obtained. The polymorphic loci amplified in Kiewiet are an indication that the hybrid evolved after the original cross between rye and wheat. These polymorphic loci amplified in Kiewiet cause the genetic distance between Kiewiet and wheat to increase and is therefore the main reason why the genetic distance of Kiewiet and wheat is in the region of 50 % and not 33.3 %. The average pairwise genetic distance between a cultivar grouped in the middle of the wheat group (Corwa) and the rest of the wheat genotypes is 3.1 %, which indicates a low variation among the wheat genotypes.

The genetic distance between RL6081, carrying gene *Lr37*, is 5.4 % from the average South African wheat genotype. If a cross between two genotype halfs the genetic distance between two cultivars: four backcrosses would be required to transfer the resistance gene from RL6081 to a South African genotype in order to obtain a genotype with 98.1363 % genetic similarity to the South African genotype, but with the *Lr37* resistance gene. The lines RL6081 and VPM1, carrying gene *Lr37* are genetically closest to the South African cultivar Betta, with genetic distances of 3.6 % and 3.7 %, respectively. The gene can therefore be theoretically transferred with the least amount of backcrosses to the South African cultivar Betta in order to obtain a NIL with resistance to leaf rust. RL6081 and VPM1 are furthest apart from the cultivar Molopo with a distance of 7.4 % and 6.5 %, respectively. Crosses of Molopo with the *Lr37* carrier cultivars will therefore, result in the most genetically diverse offspring, which is most desirable in wheat with its generally low genetic diversity (Gale *et al.*, 1990; Devos & Gale, 1992). The pairwise genetic distance between the near isogenic lines Thatcher and RL6081 is further 1.28 %. A comparison of a 100 molecular amplification products between the two lines will result in 1.28 of the compared products being polymorphic. This implies that a polymorphic fragment between the two near isogenic lines, that can be tested as a marker linked to gene *Lr37*, will result after 78.125 amplified loci are compared between the two genotypes.

The three molecular techniques used in the study provided valuable information regarding combining ability aspects among *Lr37* carrying genotypes and local South African genotypes. Putative markers linked to genes *Dn5* and *Lr37* were also identified with AFLPs. The genetic distance based on molecular data, between the near isogenic lines Thatcher and RL6081, further reveals exciting prospects due to the fact that every 78th amplification product being compared between the NILs, can be a possible marker linked to gene *Lr37*.

Chapter 4

Identification and characterisation of DNA markers linked to leaf rust resistance gene *Lr37*

4.1 *Introduction*

Growing crops with resistance to disease started as early as 1880, in Australia (McIntosh & Brown, 1997) and is the most effective method of controlling the three rust diseases of wheat (Knott, 1989). A very serious disease of wheat is caused by the leaf rust fungus, *P. recondita* f. sp. *tritici* (Samborski, 1985; Roelfs, 1988). Leaf rust often reaches epidemic proportions in South Africa (Pretorius *et al.*, 1987; Pretorius & Le Roux, 1988; Pretorius, 1990) and control through genetic resistance is therefore essential. Forty-seven genes conferring resistance to leaf rust have been designated and pyramiding of these genes (Schafer *et al.*, 1963) often results in durable resistance (Johnson & Law, 1975). Markers that are linked to the different genes are necessary for indirect selection of the genes (Melchinger, 1990), especially were genes are combined that are effective against the same pathotypes.

At first, mainly morphological characteristics (markers) were used to identify resistant plants. Morphological markers are, however, limited in number and multiple morphological markers are often ineffective in a single segregating population (Paterson *et al.*, 1991). Molecular markers can therefore be used, not to replace classical methods, but to assist breeders with gene pyramiding (Kelly, 1995). Molecular markers may speed up the process of developing new resistant cultivars (Winter & Kahl, 1995), because pathogenic tests are not required when molecular markers are available (Kelly, 1995).

Two methods can successfully be used to identify molecular markers linked to a single gene. Near isogenic lines (NILs) can be compared on DNA level. Molecular differences between NILs which only differ in the gene of interest and small pieces of DNA surrounding the gene, can be used as markers linked to the gene (Weeden *et al.*, 1994). Bulk segregant analysis of an F_2 population can also be used for marker identification (Michelmore *et al.*, 1991). Resistant F_2 plant DNA ($\pm$ 10 plants) and susceptible F_2 plant DNA ($\pm$ 10 plants) are bulked into two separate bulks and molecular differences between the bulks can be tested for markers linked to the gene of interest.

Molecular techniques such as RAPDs, AFLPs and SSRs can be used to identify molecular markers linked to important genes. Leaf rust resistance gene *Lr24* is an example of a gene tagged with a marker resulting from a polymorphism when using RAPDs (Dedryver *et al.*, 1996). Gene *Lr41* was recently tagged at the University of Pretoria, biotechnology laboratory, with a specific marker resulting from a polymorphism when using AFLPs (Lottering *et al.*, 1999) and SSRs were also used to identify three important genes in waxy wheat breeding programs (Shariflou & Sharp, 1999). The aim of this study was to identify molecular markers linked to leaf rust resistance gene *Lr37*, using three techniques RAPDs, AFLPs and SSRs. This gene is an important adult-plant resistance gene that can also be combined effectively with genes such as *Lr13* and gene *Lr34* to increase resistance against leaf rust (Kloppers & Pretorius, 1997).

4.2 *Materials and methods*

4.2.1 *Plant material*

The resistance gene to leaf rust from the wheat cultivar VPM1 was designated *Lr37* and mapped to chromosome 2AS (Bariana & McIntosh, 1993, 1994). Backcrosses between a susceptible Thatcher line and the line VPM1, resulted in a *Lr37* testline, RL6081 (Thatcher*8/VPM1) with resistance against leaf rust (Dyck & Lukow, 1988). A BC_2F_2 progeny from a cross involving the line RL6081 and the susceptible South African cultivar Karee was used to test the segregation of gene *Lr37* phenotypically in the study. These crosses were obtained from the department of Plant Pathology at the University of the Orange Free State, South Africa.

4.2.2 *Inoculations*

Innoculations of the cultivar Karee and the line RL6081 as well as the BC_2F_2 population with the pathotype UVPrt9 of *P. recondita* f. sp. *tritici* were performed according to Kloppers *et al.* (1994), 48 days after planting. Freshly harvested urediospores, suspended in odourless kerosene (McSherry & Harris, Wedmore, Somerset, UK) (1mg spores/ml oil) were sprayed onto the plants. The plants were

dried at room temperature and placed in a dark dew-simulation chamber at 18-20 °C. The plants were removed after 19 h, allowed to dry at room temperature and placed in a greenhouse compartment at 20 °C. A 14h photoperiod using light of 200 $\mu Em^{-2}s^{-1}$ was applied. The leaf rust reactions were scored 14 days post inoculation, according to Roelfs (1988).

4.2.3 DNA isolation

DNA was extracted as described in Chapter 3, except for using 1g of leaf tissue (prior to inoculations) instead of the seedlings.

4.2.4 RAPD-PCR analysis

The same procedures for RAPD analysis were used as described in Chapter 3.

4.2.5 AFLP analysis

The same procedure for AFLP analysis were used as described in Chapter 3, except for the visualization that was performed with silver staining of nondenaturing polyacrylamide gels as described for SSRs product visualization in Chapter 3.

4.2.6 SSR analysis

SSR amplifications were also performed according to the technique described by Lee *et al.*, (1995). Five primers that amplify regions from chromosome 2AS, where gene *Lr37* is located, were selected randomly (Röder *et al.*, 1998). A 25 μL reaction contained 20 pM of the primer pair XGWM497 (5' GTA GTG AAG ACA AGG GCA TT 3' & CCG AAA GTT GGG TGA TAT AC 3'), XGWM359 (5' CTA ATT GCA ACA GGT GAT GGG 3' & 5'TAC TTG TGT TCT GGG ACA ATG G3'), XGWM 636 (5' CGG TAG TTT TTA GCA AAG AG 3' & 5'CCT TAC AGT TCT TGG GAG AA 3'), XGWM71.1 (5' GGC AGA GCA GCG AGA CTC 3' & 5'CAA GTG GAG CAT TAG GTA CAC G 3') or XGWM95 (5'GAT CAA ACA CAC ACC CCT CC3' & 5'AAT GCA AAG TGA AAA ACC CG3'), 20 ng DNA, 800 μM

dNTPs, 1 U Taq polymerase (Promega), 1 x PCR buffer (20 mM $(NH_4)_2SO_4$, 75 mM Tris-HCL pH 9.0, 0.1 % (v/v) Tween), and 1.5 mM $MgCl_2$. The following temperature profile was used: 40 cycles at 94 °C for 1 min, 65 °C for 1 min and 72 °C for 2 min and a final cycle at 72 °C for 10 min. SSR amplification products were separated on six percent nondenaturing polyacrylamide gels (Sambrook *et al.*, 1989). The visualization of SSR amplification products on the gels were performed, also using the silver staining method as described in Chapter 3.

4.2.7 *Linkage analysis*

A bulk segregant analysis approach was followed as described by Michelmore *et al.* (1991). DNA from resistant and susceptible BC_2F_2 plants, resulting from a cross between the South African cultivar Karee and the *Lr37* line, RL6081, was bulked. DNA bulks were prepared from equal concentrations of DNA of five susceptible and five resistant individual plants, respectively. RAPD amplified fragments that showed polymorphisms when the resistant and susceptible DNA bulks were compared, were further analysed with individual plants to determine linkage within the BC_2F_2 population. All the relevant bulks and parental lines were also included in the linkage tests, to discriminate against markers that are not repeatable in different genetic backgrounds.

AFLP polymorphisms between resistant and susceptible BC_2F_2 bulks were tested for linkage to gene *Lr37*, using Dot Blots. SSR amplification products were first tested for polymorphisms between the susceptible and resistant parental lines, Thatcher and line RL6081, as described for tests with near isogenic lines by Weeden *et al.* (1994). Markers that were polymorphic between the bulks and parent lines that also segregated in the BC_2F_2 population were regarded as linked to the gene. Recombination fractions of the markers in the BC_2F_2 population were then transformed to linkage distances in centimorgans (cM) according to Kosambi (1944).

4.2.8 Cloning of marker fragments

RAPD marker $OPQ7_{1160}$ was isolated from the 2 % agarose gel using a GENECLEAN^{R}111 kit (BIO101) according to manufacturer's instructions. The purification of AFLP fragments R_{470} and S_{800}, amplified with selective primers EcoRI-ACA and MseI-CTG from the 6 % acrylamide gel involved a PCR reaction using the same AFLP selective amplification primers. A piece of polyacrylamide containing the fragment was excised from the gel and suspended in two volumes of distilled water for 24 hours. A 0.5 µl volume of the resuspended polyacrylamide piece was then used as template to reamplify the fragment in a PCR reaction similar to selective amplification reactions.

Purified RAPD marker fragment $OPQ7_{1160}$ and AFLP marker fragments R_{470} and S_{800} were cloned within the α-peptide coding region of the enzyme β-galactosidase of pGEM^R-T easy cloning vector systems (Promega). The vectors with the cloned inserts were then transformed into JM109 high efficiency competent (>10^8 cfu/µg DNA) $E.$ $coli$ cells according to manufacturer's instructions (Promega). The $E.$ $coli$ cells were plated on a LB (3 µg/µl BACTO^R-Tryptone, 1.5 µg/µl Bacto^R-Yeast extract and 1.5 µg/µl NaCl) selection medium that contains IPTG (0.5 mM), X-Gal (0.08 µg/µl) and Ampicillin (0.1 µg/µl) and incubated at 37 °C for 24 hours.

Colony PCR amplification of recombinant $E.$ $coli$ colonies were used to determine if the correct DNA fragment was cloned for each of the markers $OPQ7_{1160}$, EcoACA-MseCTG$_{470}$ and EcoACA-MseCTG$_{800}$. The PCR reaction was similar to RAPD amplification reactions except for the use of 0.3 pmol SP6 and T7 primer pair (Promega) and at least one recombinant $E.$ $coli$ cell as template. After electrophoresis, the sizes of the individual colony PCR products were compared with the sizes of the corresponding markers. The correctness of this assay was confirmed with dot blots, especially where colony PCR products were to close to each other to distinguish the appropriate fragment on size.

4.2.9 Dot blot hybridization of marker fragments

Dot blot hybridization of the marker fragments (*Eco*ACA-*Mse*CTG$_{470}$ and *Eco*ACA-*Mse*CTG$_{800}$) were performed according to the technique described by Sambrook *et al.* (1989). AFLP selective amplification products of wheat genotypes Thatcher, RL6081, VPM1, the susceptible and the resistant BC$_2$F$_2$ bulks, were denatured with the addition of 0.1 volume of 3 M NaOH followed by incubation at 60 °C for one hour. The sample was subsequently cooled and 1 volume of 5 x SSC buffer (0.937 M NaCl, 0.0937 M C$_6$H$_5$Na$_3$O$_7$.2H$_2$O, pH 7) was added. A positively charged nylon (Boehringer Mannheim) membrane was wetted with UHQ water and then soaked in 5 x SSC solution. Of each sample 2 µl was applied to the membrane and the applied sample was immobilized by baking the membrane at 120 °C for 20 minutes.

Colony PCR products were labelled radioactively with a nick translation system (Promega). A 25 µl nick translation mixture contained: 3 uM dNTP mix (dCTP, dGTP, dTTP), 1 x Nick translation buffer (50 mM Tris-HCL pH 7.2, 10 mM MgSO$_4$ and 1 mM DTT), 1 µg sample DNA, 1 µM [α-^{32}P] dATP (70 µCi at 400 Ci/mmole and 10 mCi/ml) and 2.5 µl DNA polymerase I/ DNAse I mix (1 U/µl DNA polymerase I and 0.2 ng/µl DNase I in a solution containing 50 % glycerol, 50 mM Tris-HCL pH 7.2, 10 mM MgSO$_4$, 0.1 mM DTT and 0.5 mg/ml nuclease-free BSA.). The Nick translation reaction was incubated at 15 °C for 1.5 hours. This was followed by purification of the probes from unincorporated radioactive dNTPs using a Sephadex G-50 column calibrated with a 1 x TEN buffer (0.1 M NaCl, 10 mM TrisHCl pH 8.0, 1 mM EDTA).

The membrane was pre-hybridized in a solution (5 x SSC, 1 % blocking Reagent (Roche)) for 2 h at 60 °C. After incubation the pre-hybridization solution was removed and a hybridization solution (5 x SSC, 0.5 % SDS) was added to the membrane. The probe was denatured at 95 °C for five minutes and placed on ice, before the denatured probe was added to the hybridization solution. The hybridization step was performed for 16 h at 52 °C. Post hybridization procedures involved two washing steps of 30 minutes in 2 x SSC and 0.5 % SDS at 52 °C and a final wash step

at 52 °C in 2 x SSC. Results were visualised using autoradiography (Sambrook *et al.*, 1989).

4.2.10 *Sequencing*

The plasmid containing the cloned fragment $OPQ7_{1160}$ was purified from *E. coli* cells according to the technique described by Sambrook *et al.* (1989). The isolated plasmids were treated with RNaseA for 2 h at 37 °C and further purified using a 13% PEG 6000, precipitation step. Sequencing was performed using an ABI PRISM^{TM} Dye Terminator Cycle Sequencing Ready Reaction Kit with ampli*Taq* DNA polymerase (Perkin Elmer). A 5 µl total reaction mixture contained: 2.68 µl of template DNA, 2 µl Terminator Ready Reaction mixture and 3.2 pmol primer ($SP6_{forward}$ and $T7_{reverse}$ primers (Promega)). The internal region of fragment $OPQ7_{1160}$ was sequenced with nested primers, OPQ7LI (5' TGC TGT GGA TCT GAC ATA GTT T $3'_{forward}$) and primer OPQ7RI (5' GAT TTA AAC CAT ATA ACT CAC $3'_{reverse}$). The sequencing reaction products were separated and analyzed on an ABI PRISM^{TM} 377 DNA sequencer (Perkin Elmer).

4.2.11 *SCAR analysis*

RAPD marker $OPQ7_{1160}$ was selectively amplified from genomic wheat DNA with a combination of SCAR primers (OPQ7L 5' CCC CGA TGG TCC AAA GAT GGA A 3' and OPQ7R 5'CCC CGA TGG TAC ACC CAA ATA C 3'). The SCAR primers were also tested in combination with the nested primers used for sequencing the internal region of marker $OPQ7_{1160}$. The PCR products resulting from primers OPQ7R and OPQ7LI were also digested with a variety of restriction enzymes, in order to locate differences between PCR products from resistant and susceptible wheat lines.

4.3 *Results*

4.3.1 *Segregation analysis*

The parental plants Karee (susceptible) and RL6081 (resistant) displayed the expected susceptible (3^{++}) and resistant (2^{cn}) phenotypes after inoculation with the pathotype UVPrt9 of *P. recondita* f. sp. *tritici*. The BC_2F_2 population was tested for a single recessive Mendelian segregation of 1: 3 resistant to susceptible, respectively. Five of the 35 BC_2F_2 individuals produced a resistant infection type $(2 - 2^{cn})$, while 30 individuals were susceptible and produced a susceptible (3^{++}) infection type according to the $0 - 4$ scale (Roelfs, 1988) (Appendix 4.1). The X^2-test, revealed a X^2 value of which is less than the maximum value (3.84) for $X^2_{0.05}(1)$. The hypothesis that the segregation of *Lr37* is recessive, is therefore confirmed. (Table 4.1).

Table 4.1 A segregation analysis of 35 BC_2F_2 plants from a cross between a leaf rust susceptible South African cultivar Karee and a leaf rust resistant line RL6081.

Class	Obtained value	Expected value	Deviation	(Deviation)^2/ Expected
Resistant F_2 plants	5	8.75	3.75	1.607
Susceptible F_2 plants	30	26.25	-3.75	0.534
Total (X^2)				2.141

4.3.2 *RAPD analysis*

A total of 1074 loci were amplified with 132 RAPD primers (Appendix 4.2). The average number of RAPD loci amplified per primer was 8.136. Nine of the loci were polymorphic between resistant and susceptible BC_2F_2 DNA bulks and the polymorphism rate with RAPDs was therefore 0.84 %. Polymorphisms associated with the coupling phase to gene *Lr37*, included fragments amplified with primers OPH17, OPL11 and OPQ7 (Table 4.2). Also, putative markers associated in repulsion to gene *Lr37* were amplified with OPF14, OPH4, OPK9, OPN5, OPO19 and OPR7 (Table 4.2). Only one fragment amplified with primer OPQ7 with a size of

1160 base pairs was linked to gene *Lr37*. The fragment amplified with primer OPQ7 (OPQ7$_{1160}$) was present in the resistant parent, RL6081 and the original donor parent of gene *Lr37*, line VPM1, but absent in the susceptible Thatcher parental line and absent in 4 out of 5 susceptible BC$_2$F$_2$ individuals from a RL6081 and Karee cross. OPQ7$_{1160}$ was further present in 3 out of 4 resistant BC$_2$F$_2$ individuals (Fig 4.1). Marker OPQ7$_{1160}$ is therefore linked to gene *Lr37* with a distance of 14cM (Kosambi, 1944) (Fig 4.1).

4.3.3 *Cloning and sequencing of OPQ7$_{1160}$*

After RAPD marker OPQ7$_{1160}$ were cloned successfully, recombinant *E. coli* cells remained white, while non-recombinant *E. coli* cells turned blue after twenty-four hours (Fig 4.2). A colony PCR of seven recombinant colonies, with the possibility of having marker OPQ7$_{1160}$ incorporated in their genome revealed two colonies (five and seven) with inserted fragments of the correct size (Fig 4.3). The two colonies showed similar inserts of 1160 base pairs when the purified plasmids were digested with the enzyme *Eco*RI (Fig 4.4).

4.3.4 *SCAR analysis of OPQ7$_{1160}$*

Three hundred and fifty bases were sequenced from both ends of marker OPQ7$_{1160}$, with primers SP6 and T7 (Promega) and the internal region of marker OPQ7$_{1160}$ were furthermore sequenced, with primers OPQ7LI and OPQ7RI (Fig. 4.5). When using a combination of the SCAR primers, they were not able to distinguish between the resistant and susceptible plants (not shown). PCR amplification products of OPQ7R and nested primer OPQ7LI were digested with various enzymes, but no differences between resistant and susceptible plants could be found either (Fig 4.6).

Table 4.2 Putative markers linked to leaf rust resistance gene *Lr37* obtained after screening resistant and susceptible plants with RAPDs, AFLPs and SSRs.

	Primers screened	Putative markers		Linkage test type		Markers linked
		c	r	A	B	
RAPDs	132		$OPF14_{831}$	-		
		$OPH17_{450}$		-		
			$OPH4_{764}$	-		
			$OPK9_{600}$	-		
		$OPL11_{650}$		-		
			$OPN5_{400}$	-		
			$OPO19_{630}$	-		
		$OPQ7_{1160}$		+		14 cM
			$OPR7_{947}$	-		
AFLPs	33	EcoAAC-MseCAT$_{630}$		-	-	
		EcoACC-MseCTT$_{697}$		-	-	
			EcoACC-MseCTT$_{889}$	-		
			EcoACA-MseCTG$_{800}$	+	-	
		EcoACA-MseCAC$_{1500}$		-		
		EcoACA-MseCTG$_{470}$		+	-	
SSRs	5	XGWM 359_{220}		-		

c = Coupling phase polymorphisms.
r = Repulsion phase polymorphisms.
A = The linkage test were done with a small F_2 population, or parents and bulks for AFLPs.
B = The linkage test was done with dot blots.
+ = Marker is linked.
- = Marker is not linked.
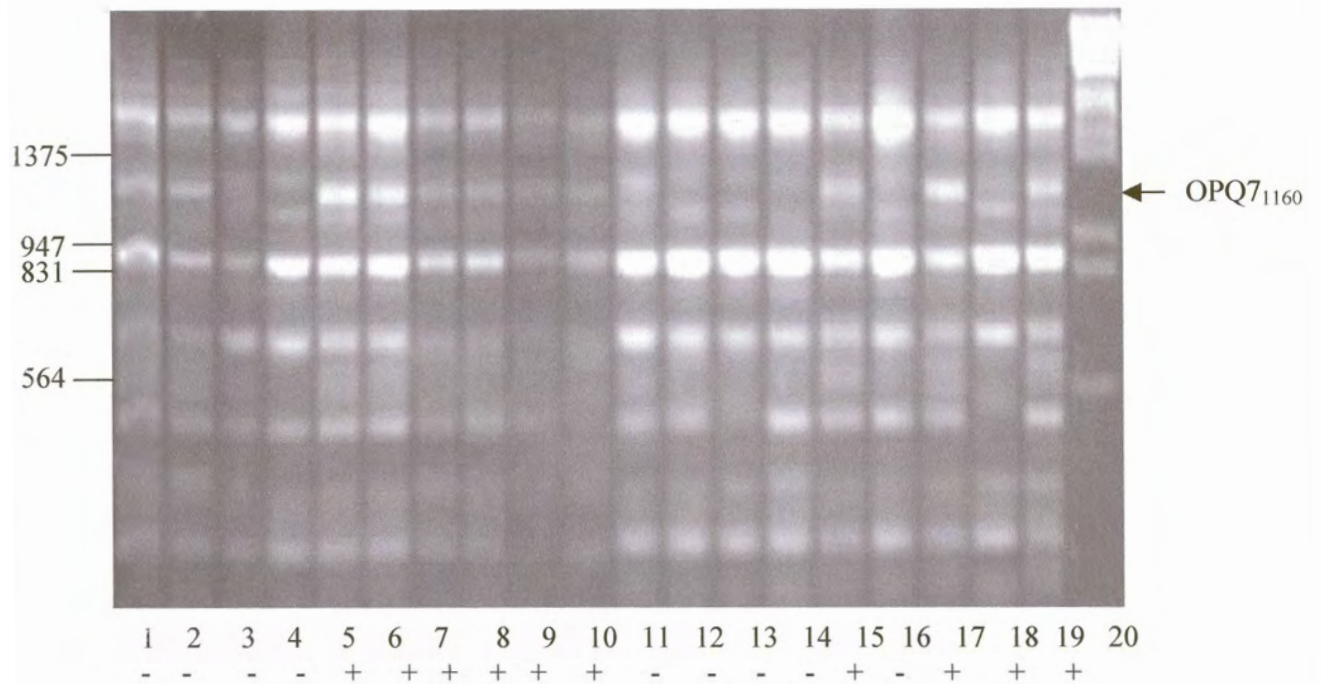
Fig. 4.1 A 2 % agarose gel stained with Ethidium Bromide, showing RAPD amplification products of primer OPQ7 when exposed to UV light. Lanes 1 and 2 = Thatcher, lanes 3 and 4 = susceptible BC_2F_2 bulk, lanes 5 and 6 = resistant BC_2F_2 bulk, lanes 7 and 8 = RL6081, lanes 9 and 10 = VPM1, lanes 11-15 = Individual susceptible BC_2F_2 plants and lanes 16-19 = individual resistant BC_2F_2 plants. Lane 20 = *Eco*RI and *Hind*III digested λ DNA marker.
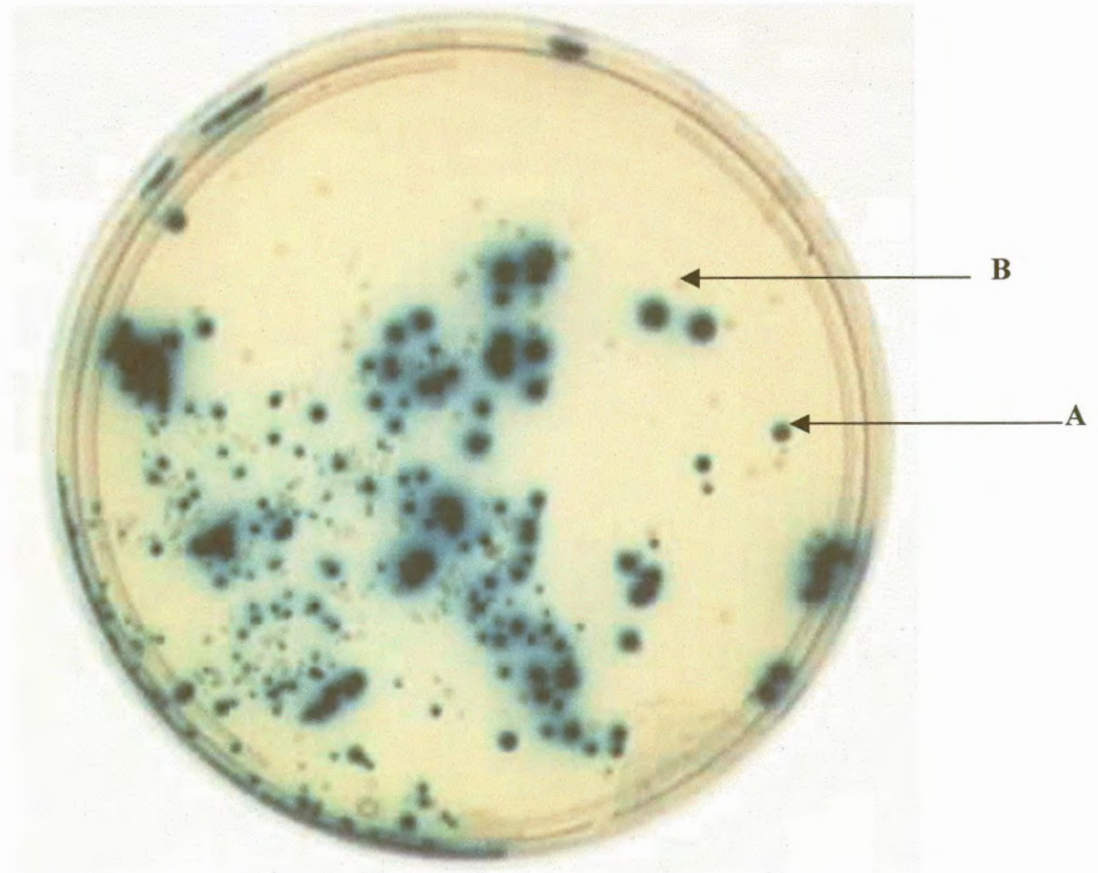
Fig 4.2 *E. coli* cells after transformation with a pGEM^R-T vector. The cells were grown on LB media that contains IPTG, X-Gal and Ampicillin. All the cells that produced a functional β-galactosidase protein formed blue colonies (A) in the presence of X-Gal, while recombinant cells that produced non-functional β-galactosidase proteins produced white colonies (B).
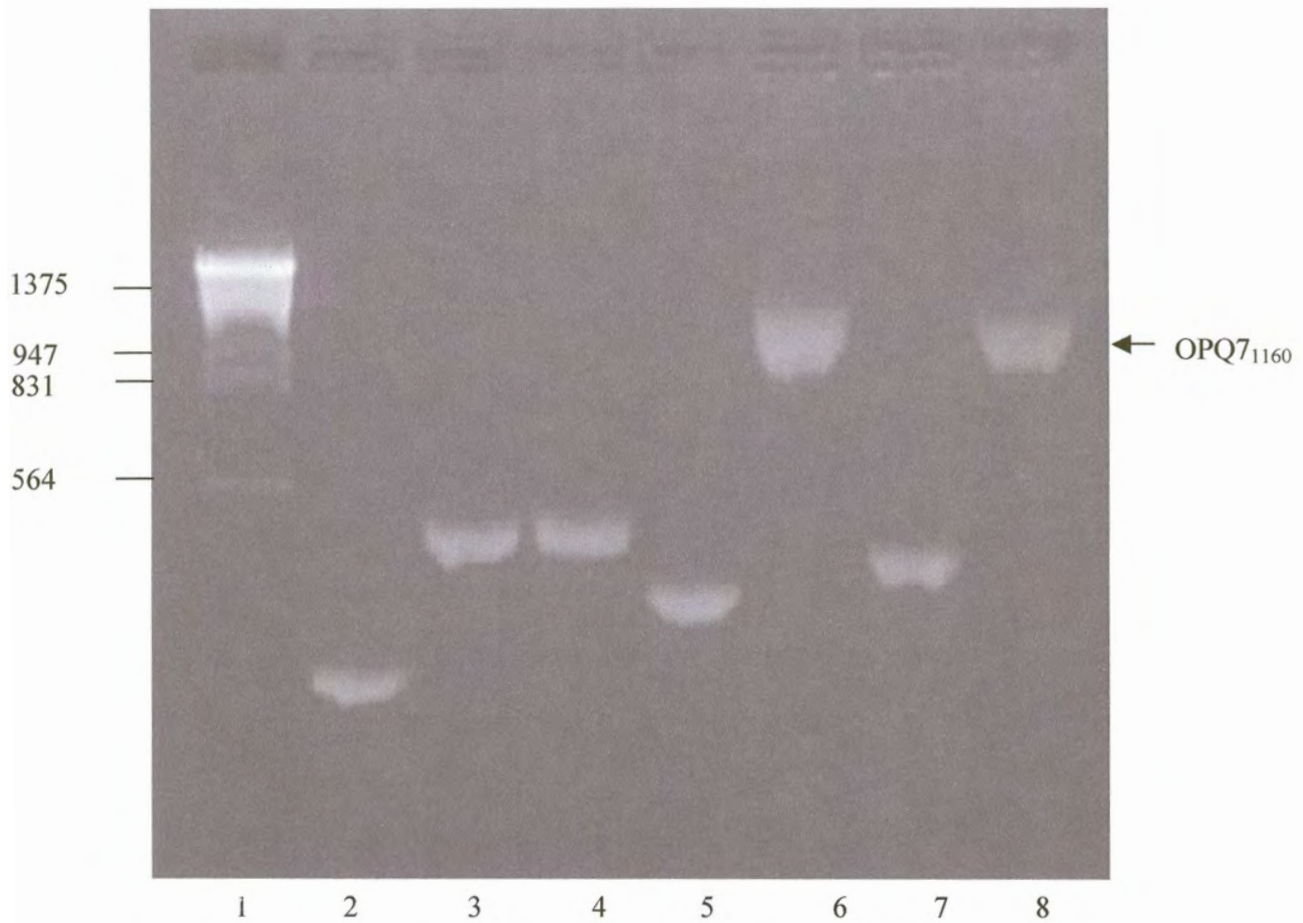
Fig 4.3 An Ethiduim Bromide stained, 2 % agarose gel showing colony PCR amplification products when exposed to UV light. Lane 1 = *Eco*RI and *Hind*III digested λ DNA marker. Lanes 2-8 = seven randomly selected recombinant *E. coli* colonies that were used in the PCR amplification with specially designed primers [SP6 & T7 (Promega, USA)] to amplify the insert region of the pGem^R-T vector. Colonies 5 and 7 both contain inserts of the correct size, which is slightly bigger than 1160 base pairs, because of the added vector sequences, amplified with the insert.
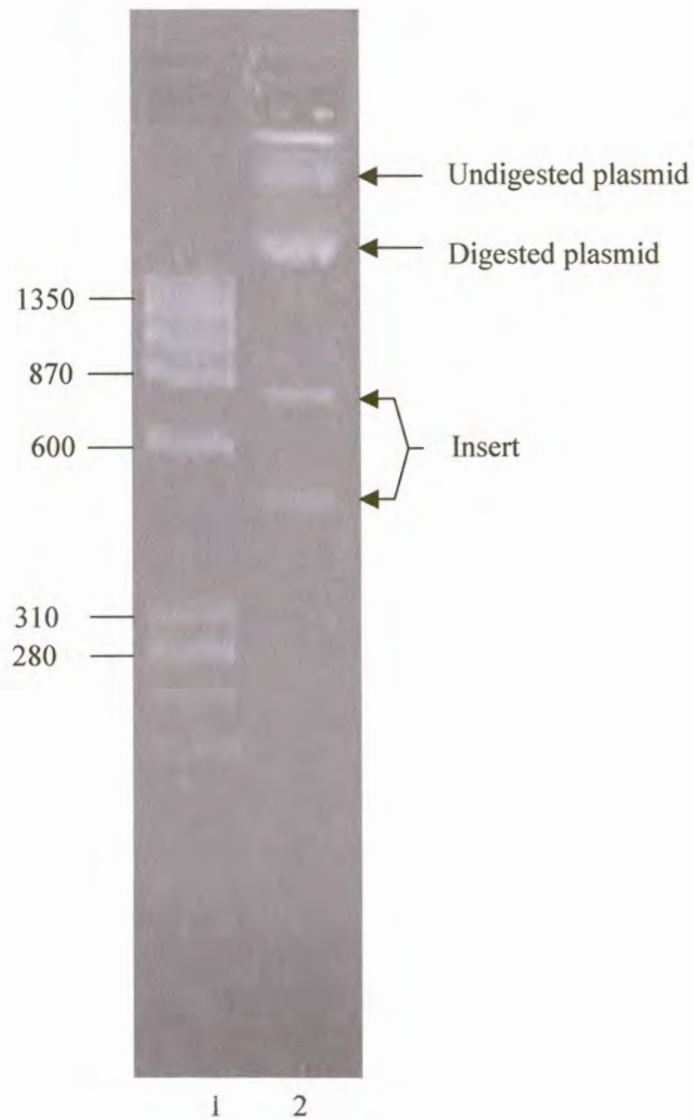
Fig 4.4 A 2 % agarose gel showing the result of an *Eco*RI digest of plasmid pGem^R-T containing the inserted RAPD amplification product OPQ7$_{1160}$, when exposed to UV light. Lane 1 contains a DNA marker and lane two contains the digested plasmid. The DNA fragment contains an *Eco*RI digestion site and therefore the two fragments with a total size of 1160 base pairs.

5'**CCCCGATGGTCCAAAGATGGAA**AATTTGAGCCCCTCCCCGAGGGGGNATGGTGATCGGTTGGACGACCANTGGCAGAA
CTCCGCGATGATAGGCCCTGGGGCATGTGTCTCTAGGAGTGCCCTTTGTTGCTCCCCTTCATCATGTGCAACACGTTTACTG
TTTTGACTTCTGGCGGGAATTTCTTCTGTTCCCTAGTGCTTTGCTAGCAAGGCTCGTCCTCGTCTTCGCTTGGTGTATCCTGC
CCCTTGTGTTCGGCATTGAGTTTGCCAGACTGCCTGAAAACCCAACATTTNTATGGGTATGATTTGCANGTTTATCGGGGGG
TGCTGTGGATCTGACATAGTTTGCCAGAATCTTGTTCAGGCTGGATAGCTCGTCTCTGTTGCCTTTGAAGGCAGCTTTTGC
TGATTTGGCCAAGAGCTTCTGAATTCAGCGTTTACCGTCGTGCTCTTCTTTATTCCGGCGCTTGTTTTTGCTGTGTCGTGGCT
TCCCGTTTCCGTCCCTAACTTCGGATGTACTTGGGTCGCTGGTGCTGCATCGAGCTAACCAGCTGTCCAGCCCCGTGCAAAA
GCGGGTCATGAGGCTTGTTAATGCTGCCATTGTTCCGGCTTTTGTTGGCCGAGGTGTATGGCAAGCCATTCGTCTCGAACGT
TGTGCTTGAAAGCTGCTAAGGCTTCAGCGTCCGGACAGTCGACTATTTGGTTCGTTTTAGTGAGAGACCTATTCAAAGATTC
GGCTAACTCTCCGGGTGGTGAGTATATGGGTTAAATCGGTCTGCATCCGATGGTCAGACGTAGGTCCCTTGAAAACCCGCC
CGAAAAGCATCCTCGAGCTCTTCCCAGCTTCCAATGGAGTTTTCGGGGAGGCTTTTAAGCCAGTGCCGAGCTGGCCCTTTGA
GCTTGAGGGTAAGTACTTGATGGCGTGGAGATCGTCTCCTCGAGTTATGCAGATATGGAGGATGTAAGTCCTCAATCCAGA
CCCCAGGGTCCGTCGTTCCGTCGTAGCGCTCTATGTTTACGGGTTTGAATCCCTCTGGTAATTCATGGTCCAGCACCTCATC
AGCGAAACATAGGGGGTGTGTGACGCCCTTGTATTTGGGTGTACCATCGGGG 3'

3'
GGGGCTACCAGGTTTCTACCTTTTAAACTCGGGGAGGGGCTCCCCCNTACCACTAGCCAAACCTGCTGGTNACCGTCTTGA
GGCGCTACTATCCGGGACCCCGTACACAGAGATCCTCACGGCGAAACAACGAGGGGAAGTAGTACACGTTGTGCAAATGA
CAAAACTGAAGACCGCCCTTAAAGAAGACAAGGGATCACGAAACGATCGTTCCGAGCAGGAGCAGAAGCGAACCACATA
GGACGGGGAACACAAGCCGTAACTCAAACGGTCTGACGGACTTTTGGGTTGTAAANATACCCATACTAAACGTNCAAATA
GCCCCCCACGACACCTAGACTGTATCAAACGGTCTTAGAACAAGTCCGACCTATCGAGCAGAGACAACGGAAACTTCCGTC
GAAAACGACTAAACCGGTTCTCGAAGACTTAAGTCGCAAATGGCAGCACGAGAAGAAATAAGGCCGCGAACAAAAACGA
CACAGCACCGAAGGGCAAAGGCAGGGATTGAAGCCTACATGAACCCAGCGACCACGACGTAGCTCGATTGGTCGACAGGT
CGGGGCACGTTTTCGCCCAGTACTCCGAACAATTACGACGGTAACAAGGCCGAAAACAACCGGCTCCACATACCGTTCGGT
AAGCAGAGCTTGCAACACGAACTTTCGACGATTCCGAAGTCGCAGGCCTGTCAGCTGATAAACCAAGCAAAATCACTCTCT
GGATAAGTTTCTAAGCCGATTGAGAGGCC**CACCACTCATATACCAATTTAG**CCAGACGTAGGCTACCAGTCTGCATCCAG
GGAACTTTTGGGCGGGCTTTTCGTAGGAGCTCGAGAAGGGTCGAAGGTTACCTCAAAAGCCCCTCCGAAAATTCGGTCACG
GCTCGACCGGGAAACTCGAACTCCCATTCATGAACTACCGCACCTCTAGCAGAGGAGCTCAATACGTCTATACCTCCTACA
TTCAGGAGTTAGGTCTGGGGTCCCAGGCAGCAAGGCAGCATCGCGAGATACAAATGCCCAAACTTAGGGAGACCATTAAG
TACCAGGTCGTGGAGTAGTCGCTTTGTATCCCCCACACACTGCGGGAA**CATAAACCCACATGGTAGCCCC** 5'

Fig 4.5 Sequence information obtained after sequencing marker fragment OPQ7$_{1160}$ with primers T7 and SP6 (Promega) from both ends. The internal region was later sequenced with nested primers OPQ7LI and OPQ7RI (blue). The two SCAR primers designed to amplify specifically the marker fragment are indicated in green.
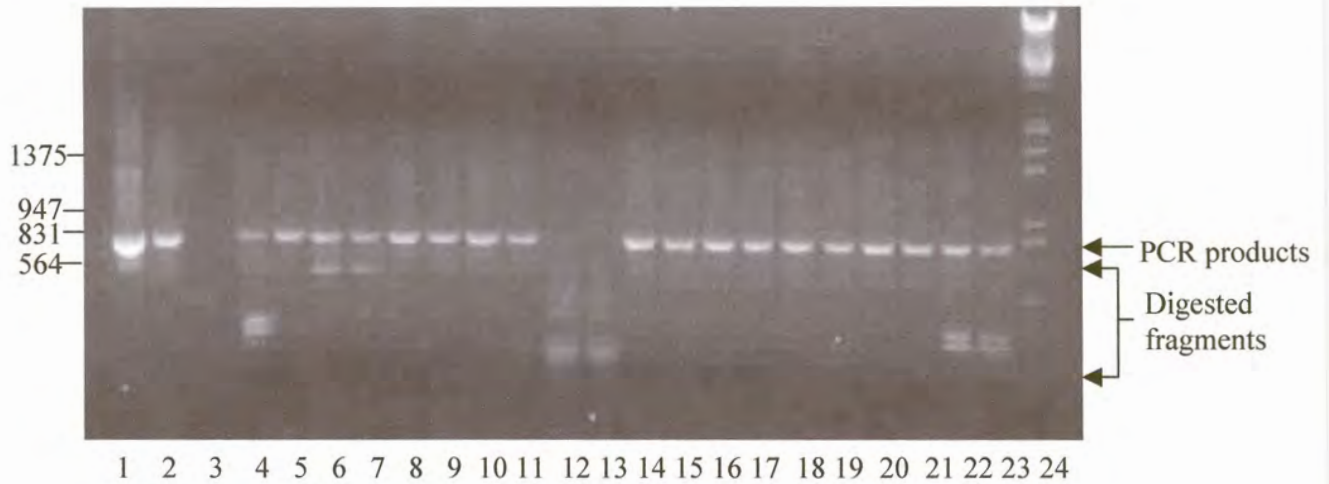
Fig 4.6 A 2 % agarose gel showing the PCR result of a combination of primers OPQ7LI and OPQ7R when using genomic wheat DNA of the cultivar Karee and line RL6081 as template. The PCR products were also digested with various enzymes. Lanes 1 & 2 = Karee and RL6081. Lane 3 = a negative control. Lanes 4 & 5 = digestions with *Hind*III. Lanes 6 & 7 = digestions with *Nde*I. Lanes 8 & 9 = digestions with *Alu*I. Lanes 10 & 11 = digestions with *Ava*I. Lanes 12 & 13 = digestions with *Mse*I. Lanes 14 & 15 = digestions with *Pst*I. Lanes 16 and 17 = digestions with *Bam*HI. Lanes 18 and 19 = digestions with *Eco*RI. Lanes 20 and 21 = digestions with *Kpn*I. Lanes 22 and 23 = digestions with *Sal*I. Lane 24 = *Eco*RI and *Hind*III digested λ DNA marker.

4.3.5 *AFLP analysis*

A total of 1761 fragments were amplified with 33 AFLP primer pairs (Appendix 4.3). An average of 53.363 loci was thus amplified per primer pair. Six of the fragments were polymorphic between the susceptible and resistant BC_2F_2 DNA bulks. Of the six polymorphic fragments, EcoAAC-MseCAT$_{630}$, EcoACC-MseCTT$_{697}$, EcoACA-MseCAC$_{1500}$ and EcoACA-MseCTG$_{470}$ (Fig 4.7) were polymorphic in the coupling phase, while fragments EcoACC-MseCTT$_{889}$ and EcoACA-MseCTG$_{800}$ were polymorphic in repulsion to gene *Lr37* (Table 4.2). A polymorphic rate of 0.34 % was therefore obtained between resistant and susceptible DNA bulks. The polymorphisms of AFLP amplified fragments EcoACA-MseCTG$_{470}$ and EcoACA-MseCTG$_{800}$ showed putative linkage to gene *Lr37* when amplification products of parental and BC_2F_2 bulks are compared. The fragments were successfully cloned in order to test the putative markers with dot blots for linkage to gene *Lr37*. Colony PCR products of the two cloned fragments EcoACA-MseCTG$_{470}$ and EcoACA-MseCTG$_{800}$, resulted in products, 172 bases larger than the original marker fragment excised from the gel. The 172 bases includes multiple cloning site sequences of the vector being amplified with the inserted fragments, when using SP6 and T7 primers to amplify inserted fragments from pGEM^R-T easy cloning vector systems (Promega, USA).

Seven selected recombinant *E. coli* colonies, containing the polymorphic fragment EcoACA-MseCTG$_{470}$, revealed colony PCR products to close to each other to distinguish on size (Fig 4.10). PCR products of the inserted regions from the recombinant colonies were therefore used to hybridise against AFLP amplification products of a subset of the segregating population. Six of the fragments selected for the dot blot test, hybridised to both susceptible and resistant parents, with 4 out of six X-Ray films revealing the polymorphism between the resistant and susceptible BC_2F_2 bulks (Fig 4.9). A similar dot blot linkage test was performed for positive *E. coli* colonies containing marker EcoACA-MseCTG$_{800}$ and the hybridization results showed no repulsion phase linkage to gene *Lr37* (results not shown).
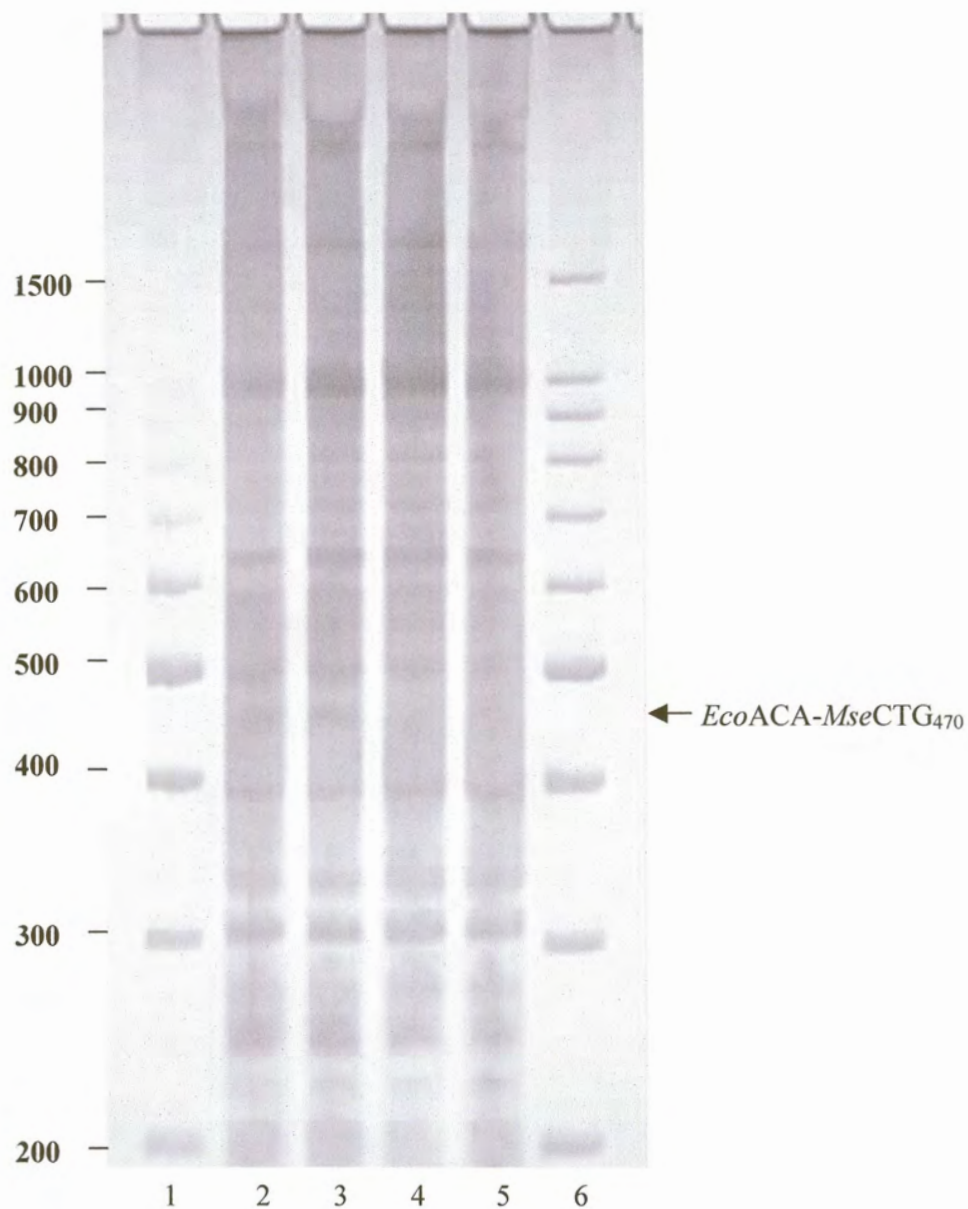
$\leftarrow$ *Eco*ACA-*Mse*CTG$_{470}$

Fig 4.7 A 6 % nondenaturing polyacrylamide gel showing the results visualized with silver staining after an AFLP selective amplification. The primer combination used in this amplification was *Eco*RI-ACA & *Mse*I-CTG. Lane 1 = 100 base pair DNA marker. Lanes 2 & 3 = resistant BC$_2$F$_2$ bulk DNA, Lanes 4 & 5 = susceptible BC$_2$F$_2$ bulk DNA and Lane 6 = 100 base pair DNA marker. A polymorphism between resistant and susceptible bulks is indicated with the arrow at 470 base pairs.
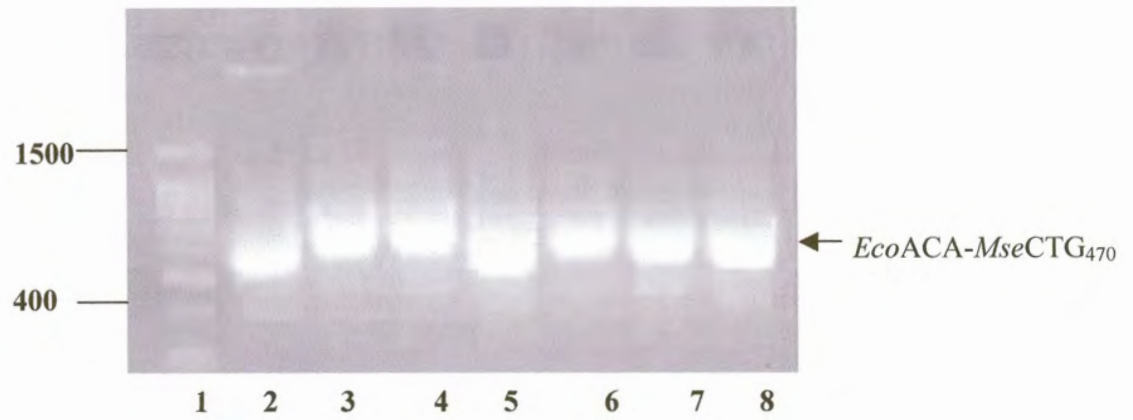
Fig 4.8 An Ethidium Bromide stained, 2 % agarose gel showing colony PCR amplification products when exposed to UV light. Lanes 2-8 = Seven randomly selected recombinant *E. coli* colonies used in the PCR amplification, with specially designed primers [SP6 & T7 (Promega, USA)].
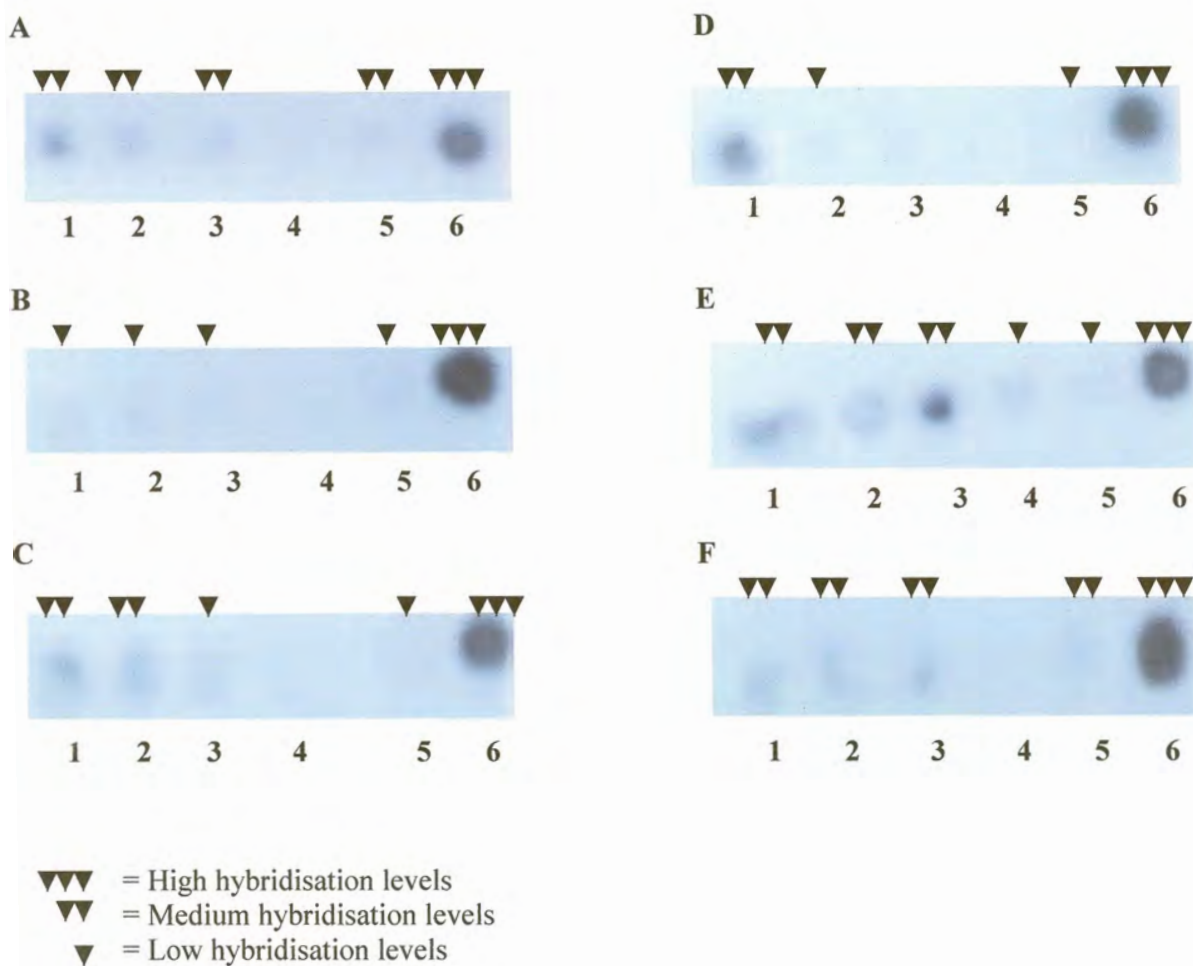
▼▼▼ = High hybridisation levels
▼▼ = Medium hybridisation levels
▼ = Low hybridisation levels

Fig 4.9 Hybridization results obtained after ^{32}P [dATP] labeling of six randomly chosen clones (A-F) of AFLP fragment EcoACA-MseCTG$_{470}$. The fragment was hybridized against AFLP amplification products, using primers EcoRI-ACA and MseI-CTG and template DNA as indicated. Lane 1 = Thatcher, Lane 2 = RL6081, Lane 3 = VPM1, Lane 4 = susceptible BC$_2$F$_2$ bulk, Lane 5 = resistant BC$_2$F$_2$ bulk and Lane 6 = positive control.

4.3.6 *SSR analysis*

One primer pair (XGWM359) amplified a 220 base pair polymorphic fragment between the NILs, Thatcher and the resistant line RL6081. The 220 base pair fragment was present in the resistant line RL6081 and absent in the susceptible cultivar Thatcher. The putative marker was tested for linkage to gene *Lr37* with a subset of the segregating population. However, no linkage to gene *Lr37* could be confirmed, when samples from the original donor parent of gene *Lr37* (VPM1), the susceptible South African cultivar Karee, and both susceptible and resistant BC_2F_2 DNA bulks from a cross between the cultivar Karee and line RL6081, were included in the study (Fig 4.10).
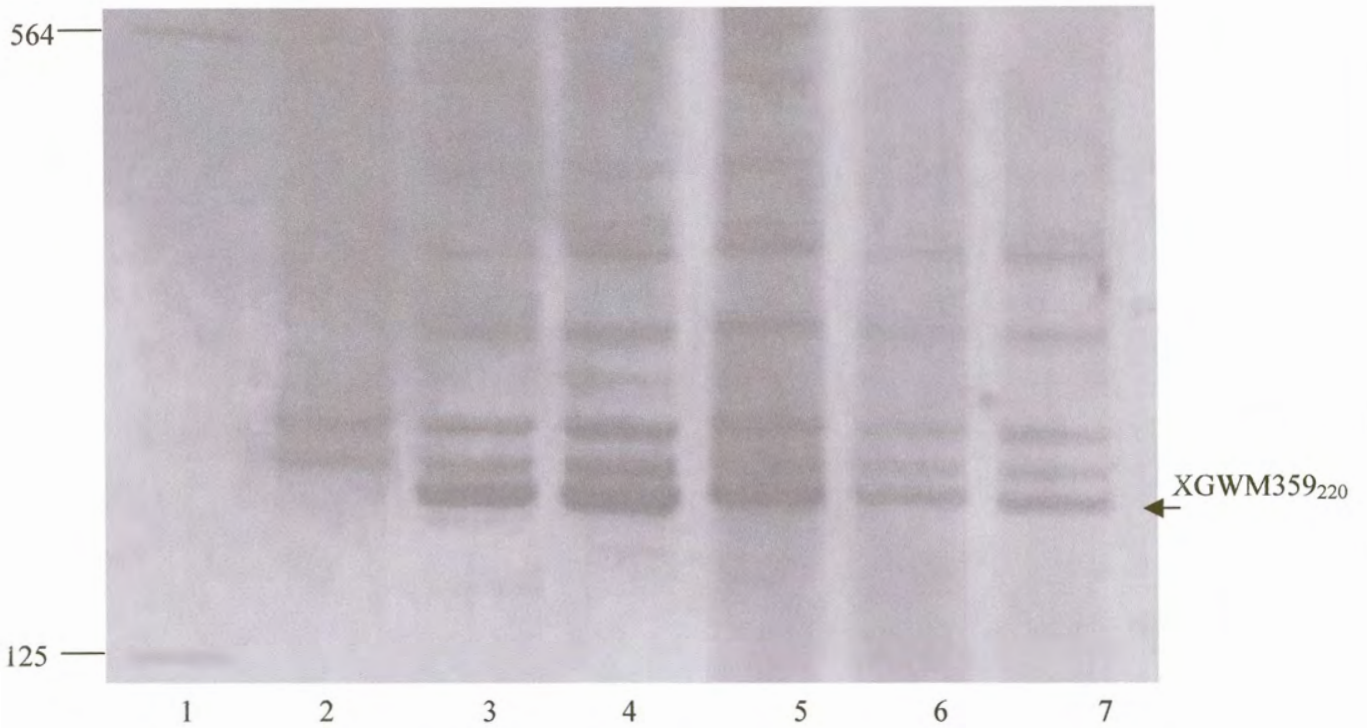
Fig 4.10 A 6 % nondenaturing polyacrylamide gel showing the results visualized with silverstaining after genomic wheat DNA is amplified with SSR marker XGWM359. Lane 1 = *EcoRI* and *Hind*III digested λ DNA marker, Lane 2 = Thatcher, Lane 3 = RL6081, Lane 4 = VPM1, Lane 5 = Karee, Lane 6 = susceptible BC_2F_2 bulk DNA and Lane 7 = resistant BC_2F_2 bulk DNA. Marker $XGWM359_{220}$ shows the polymorphism obtained between lanes 2 and 3, which contain amplified products from the near isogenic lines Thatcher and RL6081, respectively.

4.4 Discussion

Parental lines responded to *P. recondita* f. sp. *tritici* as expected, with line RL6081 resistant (2^{cn}) and the local Karee line susceptible (3^{++}). A segregation analysis of the BC_2F_2 population revealed resistant plants to susceptible plants in a ratio of 5 to 30, respectively. The Chi square value of 2.141 is lower than 3.84, which is the critical value to expect for a recessive Mendelian segregation on a 95 % level of confidence. The hypothesis (H_0) that gene *Lr37* is inherited in a recessive manner was therefore confirmed. Previous studies have also shown that the gene (*Lr37*) known at the time as *LrV* produces an excess of susceptible plants in a segregating F_2 generation (Dyck & Lukow, 1988). The recessive nature of the *Lr*37 gene was also illustrated in a Karee background (Kloppers & Pretorius, 1994). Both resistant and susceptible plants were obtained in the BC_2F_2 population, in the expected 1:3 ratio for a normal Mendelian segregation. The presence of resistant and susceptible plants in the BC_2F_2 population was sufficient to compare a phenotypic segregation with the segregation of molecular markers linked to the gene.

The average of 8.136 loci per primer compares well to the average of 9 bands per primer obtained in a study conducted to locate molecular markers linked to *Dn1*, *Dn2* and *Dn5* resistance genes (Myburg, 1996). Nine of the fragments were polymorphic between susceptible and resistant BC_2F_2 DNA bulks. RAPD Polymorphism $OPQ7_{1160}$ could be repeated with a linkage test involving both resistant and susceptible parents and a subset (both resistant and susceptible BC_2F_2 plants) of the population. Molecular markers linked to a gene with a distance of five centimorgans or less can be effectively employed in breeding programs (Tanksley & Orton, 1983). The distance of marker $OPQ7_{1160}$ from the gene was calculated to be 14 cM (Kosambi, 1944). Plants can therefore be classified as resistant or susceptible with a success rate of 86 % when using marker $OPQ7_{1160}$. Marker $OPQ7_{1160}$ can therefore be used in combination with other markers linked to gene *Lr37* to increase the selection efficiency for resistant plants.

SCAR markers show a less complex banding pattern than RAPD profiles and the PCR reactions are more robust than in the case of RAPDs (Procunier *et al.*, 1997). Most

RAPD markers are therefore converted to SCAR markers in order to have a more robust marker system. SCAR primers designed with the sequence information obtained from the cloned marker fragment $OPQ7_{1160}$ revealed similar amplification products in both resistant and susceptible plants. The SCAR primers were also used in combination with the nested primers, but amplification products were still present in both resistant and susceptible plants. This indicates a very small sequence difference between the resistant and susceptible plants that caused the initial RAPD fragment ($OPQ7_{1160}$) to be polymorphic.

Enzymes were successfully used to determine differences between SCAR amplification products of resistant and susceptible plants involving the *Dn5* resistant gene against the Russian wheat aphid (Venter & Botha, 2000). This approach was also tested to determine differences between OPQ7R and OPQ7LI PCR products of resistant and susceptible plants involving gene *Lr37*. The 10 different enzymes targeted random DNA sequences and did not show any differences between susceptible and resistant plants.

The average of 53.363 loci compared well with the amplification range of 20-100 loci suggested by Perkin Elmer, (1996). Of the six fragments that showed polymorphisms between resistant and susceptible BC_2F_2 DNA bulks, *Eco*ACA-*Mse*CTG$_{470}$ and *Eco*ACA-*Mse*CTG$_{800}$, showed putative linkage to gene *Lr37* when amplification products of the parents were included in the study. The fragments cloned for *Eco*ACA-*Mse*CTG$_{470}$ and *Eco*ACA-*Mse*CTG$_{800}$ were tested for linkage to gene *Lr37* with dot blots, but the tested fragments hybridized to both susceptible and resistant parents. This implies that the two AFLP markers are not linked to the gene, but were only polymorphic between resistant and susceptible BC_2F_2 DNA bulks. Four X-ray films confirm the polymorphism between resistant and susceptible DNA bulks in the case of *Eco*ACA-*Mse*CTG$_{470}$.

Four SSR markers have been used successfully to distinguish between closely related barley cultivars (Russell *et al.*, 1997). More primers are expected therefore to distinguish between NILs, because the NILs are closer related than the different barley cultivars. The wheat cultivar Thatcher and the test line RL6081 are near isogenic, except for gene *Lr37* in RL6081. Five SSR markers mapped to chromosome

2AS (Röder *et al.*, 1998), where gene *Lr37* is located (Bariana & McIntosh, 1993), proved enough to distinguish between the near isogenic genotypes Thatcher and RL6081. The polymorphism obtained between the near isogenic genotypes Thatcher and the test line RL6081 with primer combination XGWM359, revealed a putative marker linked to the gene. The marker showed, however, no linkage to the gene when using the BC_2F_2 population obtained from crosses between Karee and RL6081. This implies that the mutation that causes the marker to be absent in the susceptible cultivar Thatcher does not occur in the local susceptible cultivar Karee. The marker can therefore still be tested for linkage in a Thatcher background and the distance of the marker to the gene can be calculated using a BC_2F_2 population, resulting from a cross between lines Thatcher and RL6081.

All three of the molecular techniques used in the study provided putative markers linked to gene *Lr37*. AFLPs is the best method to identify markers linked to the gene, because of the numerous loci amplified with a single primer pair. The number of loci can, however, also be a problem because of the comigration of fragments on the gels when there are too many loci amplified. The one marker amplified with RAPD primer OPQ7 proved to be linked to the gene. The phenotypic segregation of the gene showed the gene to be recessive. Heterozygotes for recessive genes cannot be identified phenotypically, or with dominant marker techniques. The other putative markers could therefore not distinguish between resistant and susceptible plants, because heterozygotes will test susceptible in a phenotypic analysis, although gene *Lr37* will be present in one allele. Dominant markers linked in the coupling phase will incorrectly identify heterozygotes in this case as resistant.

Chapter 5

Conclusions

Wheat production can be increased when plants are better equipped to deal with diseases such as leaf rust caused by *P. recondita* f. sp. *tritici*. Forty-seven leaf rust resistance genes are available to increase natural resistance of wheat plants against leaf rust. The incorporation of these genes in plants can be assisted with molecular markers, especially where the genes are combined to obtain durable resistance (Kelly, 1995). Resistance genes alone are, however, not the only means of increasing resistance to diseases. Genetic diversity of crops is important to maintain natural resistance to various pests and diseases. A major disadvantage of cultivated crops is the genetic uniformity, which makes the plants vulnerable to diseases, but increase production in disease free areas (Knott, 1989).

Molecular techniques used to estimate genetic distances among South African cultivars, the French cultivar VPM1 (with gene *Lr37*), the Canadian developed line, RL6081 and the Thatcher recurrent parent used to develop line RL6081, showed very low levels of genetic variation. Gene *Lr37* is effective against leaf rust and South African wheat genotypes have to be crossed either with the original source VPM1 or the test line RL6081 in order to obtain the resistant gene. Crosses of local South African genotypes with *Lr37* carrying genotypes will also increase the genetic diversity. The most diverse offspring will be obtained when the *Lr37* carrying genotypes are crossed with the South African cultivar Molopo. The progeny of such a cross are therefore expected to have the most diverse genome that will enable the plants to deal with a bigger variety of environmental conditions. The *Lr37* carrying genotypes are also genetically closest to the South African cultivar Betta and the gene can therefore be transferred to the cultivar with the least amount of backcrosses in order to obtain a genotype similar to Betta, but with the resistance gene.

A leaf rust resistance gene such as *Lr37* can be tagged with molecular markers when a phenotypic segregation analysis of a segregating population is compared with polymorphisms obtained between resistant and susceptible plants using molecular techniques. The segregation analysis of a BC_2F_2 population from a Karee and RL6081 cross, resulted in an expected five resistant plants from 35 plants tested. The gene showed thus a recessive behavior. Recessive genes will be difficult to mark with dominant marker techniques, because heterozygots will be wrongly identified as false positives for the gene with dominant marker techniques. Another disadvantage of the

segregation analysis is that all statistics were calculated with only five resistant F_2 plants available to test the markers on.

The two near isogenic lines Thatcher and RL6081 proved to be genetically closest to each other. One way to locate molecular markers is to compare molecular amplification products of near isogenic lines with each other (Weeden *et al.*, 1994). In this case where the near isogenic lines are 1.28 % apart, every 78th amplified fragment should be polymorphic and can therefore be tested for linkage to the gene. RAPDs amplified an average of 8.136 loci per primer pair and 9.5 primers should therefore be sufficient to provide a polymorphism that can be tested for linkage to the gene. Only nine polymorphisms of an expected 13.89 was obtained from the total of 1074 loci amplified with 132 RAPD primers. This lower polymorphism rate is, because a different approach was followed to identify polymorphisms in the form of bulk segregant analysis (Michelmore, 1991). The recessive nature of the gene also causes polymorphisms between resistant and susceptible plants to be hidden in heterozygots. One of the polymorphisms, $OPQ7_{1160}$, was linked to the gene with a distance of 14 cM. The desired distance of markers that can be used in breeding program are less than 5 centimorgans (Tanksley, 1983), but the marker obtained can be used in combination with other markers to increase the selection efficiency for gene *Lr37*. Attempts to convert this RAPD marker into a SCAR marker failed. The SCAR markers that were designed with the original RAPD primer plus 12 bases for specificity amplified a product from both susceptible and resistant plants. The products amplified with the SCAR primers are too close to each other to distinguish differences between susceptible and resistant genotypes with a quick test on agarose gels even though the amplification products have been digested with 10 different restriction enzymes.

AFLPs have the advantage over RAPDs and SSRs in that up to 100 amplification products can be obtained with a single primer pair (Perkin-Elmer, 1996). More loci can therefore be amplified with the same selective amplification primer pair, than with the same primer in the case of RAPDs. Markers linked to genes can therefore be identified quicker with the AFLP technique. A disadvantage of the technique is that too many amplicons tends to comigrate on the gels and polymorphisms between resistant and susceptible plants can therefore be overlooked. The 33 primer pairs

amplified 1761 loci, of which six were polymorphic between resistant and susceptible bulks. The expected number of polymorphic loci was 22.6 and the same reasons can be applied as for the low RAPD polymorphism rate, as well as the fact that polymorphisms can be hidden where too many overlapping amplicons were obtained with a single primer pair.

SSRs proved to be highly polymorphic. The markers amplify specific regions of the genome, because of the longer more specific primers used with this technique. The screening process for markers can be simplified when using primers specific for the chromosome where the gene is located. The near isogenic lines could be distinguished from each other with one of the primer pairs. The primer pair XGWM359 amplified a 220 base pair fragment polymorphic between the susceptible cultivar Thatcher and the resistant line RL6081. The marker showed no linkage to the gene when it was tested in a segregating population with Karee as background, but the distance of the marker to gene *Lr37* can be tested with a BC_2F_2 population resulting from a cross between Thatcher and RL6081.

The three molecular techniques provided the expected structure for the dendograms in that rye and Kiewiet, which is a hybrid between rye and wheat, grouped out from the rest of the wheat cultivars with both the UPGMA and PAUP techniques. The three molecular techniques proved also capable of providing polymorphisms between resistant and susceptible plants. One marker was also found with RAPDs that was linked to gene *Lr37*. This marker can be used in breeding programs to select for plants containing gene *Lr37* with a success rate of 86 %. It is therefore, necessary to use the marker in combination with other markers to improve the selection efficiency for gene *Lr37*. Markers linked in repulsion to the gene can be used to identify heterozygotes, when used in combination with coupling phase markers. The use of the marker together with a marker linked in repulsion to the gene is therefore preferred, because of the recessive nature of the gene. Selecting against a repulsion phase marker and selecting for a coupling phase marker simultaniously will enable a breeder to identify plants that are homozygous for the gene.

Chapter 6

Summary / Opsomming

Leaf rust, caused by *Puccinia recondita* f. sp. *tritici* is one of the most important wheat diseases that causes crop losses of up to 15 %. The total production of wheat can furthermore be increased with plants, resistant to leaf rust. Wheat further more has low levels of genetic variation, which increases the vulnerability of the crop to disease. Resistance genes are therefore introduced from relatives, with the added advantage of increasing genetic variation. Forty-seven leaf rust resistant genes are known, which can be used to increase the resistance of plants against leaf rust.

The molecular techniques RAPDs, AFLPs and SSRs have been used in the study to estimate the genetic distances among foreign *Lr37* carrying cultivar, VPM1 (France) and line RL6081 (Canada) and local South African genotypes. The genetic distances have been used to predict which South African genotype will combine best with the two *Lr37* carrying genotypes. The South African cultivar Molopo proved to be genetically the furthest apart from the *Lr37* carrying genotypes. Crosses with Molopo will therefore have the most genetically diverse off spring and the best chances of natural resistance against a wider spectrum of diseases and climate changes. The cultivar Betta proved to be genetically closest to the *Lr37* genotypes and the gene can therefore be transferred with the least amount of backcrosses in order to obtain a Betta like genotype, with gene *Lr37*.

Leaf rust resistance gene *Lr37* have been shown to increase resistance against leaf rust when combined with genes such as *Lr13* and *Lr34* and thus needs to be tagged with molecular markers. Three molecular techniques have been used in this study to identify markers linked to gene *Lr37*. RAPDs, AFLPs and SSRs were used to compare DNA from resistant and susceptible parental plants and a segregating BC_2F_2 population. A marker ($OPQ7_{1160}$) linked in the coupling phase to gene *Lr37* with a distance of 14 cM was identified with RAPDs. An attempt was made to convert the RAPD marker to a more robust SCAR marker, but the polymorphism obtained with RAPDs could not be repeated with the SCAR primers, even though the amplicons of resistant and susceptible plants were digested with 10 different restriction enzymes.

The 132 RAPD primers amplified 1074 loci with nine of the loci polymorphic between resistant and susceptible BC_2F_2 DNA bulks. The 33 AFLP primers amplified 1761 loci with six of the loci polymorphic between the bulks. The five SSR primer

pairs amplified a polymorphic locus of 220 base pairs between the NILs for gene *Lr37*, Thatcher and RL6081. The polymorphisms obtained with RAPDs, as well as the putative marker obtained with SSRs were tested for linkage to gene *Lr37* with a subset of the F_2 population and the polymorphisms obtained with AFLPs were tested for linkage to the gene with dot blots.

Only the one polymorphism showed linkage to gene *Lr37*. The marker identified with RAPD primer OPQ7 can be used to identify resistant plants confirming *Lr37* with a success rate of 86 %. The marker can therefore be used in combination with other markers in order to increase the selection efficiency for gene *Lr37*.

Blaarroes is een van die belangrikste koringsiektes en word veroorsaak deur die fungus *Puccinia recondita* f. sp. *tritici*. Verliese van tot 15 % kan veroorsaak word, afhangend van die groeistadium van die plant wanneer infeksie plaasvind. Die totale koringproduksie kan egter verhoog word met plante wat weerstand teen blaarroes bied. Omdat koring 'n lae genetiese variasie besit, is dit nodig om weerstandsgene vanuit naverwante spesies oor te dra na koring met behulp van 'n kruisingsprogram. Sewe-en-veertig blaarroesweerstandsgene is al geïdentifiseer.

Molekulêre tegnieke is in die studie gebruik om die genetiese afstande tussen genotipes met geen-*Lr37*, komende uit Frankryk (VPM1) en Kanada (RL6081) en plaaslike Suid Afrikaanse kultivars te bepaal. Genetiese afstande tussen genotipes kan gebruik word om die beste kombinasies vir kruisings uit te soek. Die kultivar Molopo was geneties die verste van die *Lr37* draende genotipes en 'n kruising met hierdie kultivar sal die grootste genetiese variasie tot gevolg hê, wat voordelig is aangesien natuurlike weerstand teen siektes en klimaatsveranderinge verbeter namate genetiese variasie toeneem. Betta is geneties die naaste aan die buitelandse genotipes en die geen kan dus met die minste terugkruisings oorgedra word om 'n Betta tipe kultivar te kry, bevattende met die weerstandsgeen, *Lr37*.

Blaarroesweerstandsgeen-*Lr37* verhoog die weerstand van plante teen blaarroes wanneer dit met gene soos *Lr13* en *Lr34* gekombineer word. Molekulêre merkers vergemaklik die seleksie vir geen-kombinasies en drie molekulêre tegnieke is in hierdie studie gebruik om merkers vir geen-*Lr37* te identifiseer. RAPDs, AFLPs en SSRs is gebruik om DNA van ouerplante, asook segregerende weerstandbiedende en vatbare TK$_2$F$_2$ plante met mekaar te vergelyk. 'n Merker wat met 'n afstand van 14 cM aan geen-*Lr37* gekoppel is, is gevind met RAPDs. Daar is ook gepoog om die RAPD merker na 'n meer robuuste SCAR-merker om te skakel, maar dieselfde polimorfisme wat met die RAPD-inleiers gekry was, kon nie herhaal word met die SCAR-inleiers nie, ten spyte vir die feit dat die SCAR-produkte met 10 verskillende ensieme verteer is.

Die 132 RAPD-inleiers het 'n totaal van 1074 DNA-lokusse geamplifiseer. Nege polimorfies was tussen weerstandbiedende en vatbare TK$_2$F$_2$ DNA-kombinasies, bestaande uit vyf weerstandbiedende en vyf vatbare plante elk geïdentifiseer. AFLPs

het met 33 inleierpare, 'n totaal van 1761 lokusse geamplifiseer, waarvan ses polimorfies tussen die vatbare en weerstandbiedende TK_2F_2-kombinasies geïdentifiseer is. Die vyf SSR-inleier pare het 'n lokus van 220 basispare groot polimorfies tussen die na-isogenies lyne, Thatcher en RL6081, geamplifiseer. Polimorfismes wat met RAPDs en die tentatiewe merker wat met SSR-inleierpare gevind is, is met 'n deel van die TK_2F_2 populasie getoets, om vas te stel of die polimorfismes aan die geen gekoppel is. Die polimorfismes wat met AFLPs gekry is, is met "dot blots" getoets vir hierdie doeleindes.

Slegs die een merker wat met RAPD-inleier OPQ7 geamplifiseer is, het koppeling aan geen-*Lr37* getoon. Die merker kan weerstandsplante met *Lr37*, 86 % suksesvol identifiseer. Die merker kan dus tesame met ander merkers gebruik word om sodoende die seleksie-effektiwiteit vir geen-*Lr37* te verhoog.