CHAPTER 1

LITERATURE REVIEW

Chapter 1

Literature Review

1. Introduction

Scientific findings suggest that wheat, botanically speaking from the genus *Triticum* L and the grass family Gramineae, was first cultivated domestically about 7 500 years BC. These findings have been confirmed by further evidence found in prehistoric Turkish settlements, as well as in the area known as the "fertile crescent". This area ran through Mesopotamia, now Iraq and Syria (Jones & Clifford, 1983). Wheat was cultivated in the United Kingdom about 2000 BC, but only appeared in North America in the 15th Century (Jones & Clifford, 1983).

Today, wheat is an annual grass that can be grown in areas at sea level to altitudes over 3000 m. It prefers a habitat with well-drained, clay-loam soils and with a temperate, arid or semi-arid environment (Wiese, 1977). Most plants grow up to about 1 meter in length and have more than two-thirds of their fibrous roots within 20 cm of the soil surface. However, certain species may reach up to two meters in length (Wiese, 1977).

Wheat ranks first in the world's grain production and accounts for more than 20% of the total food calories consumed by man (Wiese, 1977). This crop can be grown throughout temperate, Mediterranean-type and sub-tropical regions of the world. Wheat is the main staple of traditional farming communities throughout the Atlantic coast of Europe to the Northern parts of the Indian subcontinent and from Scandinavia and Russia to Egypt (Perrino *et al.*, 1995).

2. The origin of wheat

It is commonly agreed that the evolution of bread wheat involved four different species of wild annual grasses (Feldman, 1976). These evolutionary pathways were

identified by studying genomic affinities of large numbers of species, all members of the tribe *Triticea*, which consists of a polyploid series where x = 7. The evolutionary steps were found to be natural hybridization in the wild, chromosome doubling events, from diploid (2n = 2x = 14) to tetraploid (2n = 4x = 28) to hexaploid (2n = 6x = 42). Domestication occurred at each chromosome level that led to cultivated diploids, tetraploids and hexaploid bread wheat *Triticum aestivum* (Feldman, 1976).

The whole story began with *T. urartu* (AA). This is a large-grained wild grass with ears that spontaneously break up when mature, commonly known as brittle ears (Feldman, 1976). The next millennium showed the appearance of non-brittle eared wheat, which indicated a cultivation of wheat, since the seeds would not be able to be dispersed in the wild, whereas an intact ear on the stalk is an obvious convenience for the harvester. This selection pressure led to non-brittle *T. monococcum* var. *monococcum* (Einkorn) spreading widely through the Balkans and Europe (Feldman, 1976). *Triticum urartu* hybridized with *Aegilops speltiodes* (BB) to give rise to *T. turgidum* var. *dicoccum* (Emmer), a non-brittle type wheat (Feldman, 1976). Emmer wheat spread widely in Asia Minor and later Europe, the Mediterranean, India and Central Asia (Feldman, 1976). Bread wheat, *T. aestivum* (AABBDD) appeared when *T. turgidum* (AABB) crossed with *T. tauschii (Aegilops tauschii)* (DD) (Feldman, 1976). *T. tauschii (Aegilops tauschii)* appeared as a weed within the crops or around the margins of cultivation. The final major step in the evolution of bread wheat was the selection of "free-threshing" mutants. In these plants, the grain could be easily separated from its enveloping chaff and this improved the texture of the flour. This whole process of evolution has led to the development of our present-day tetraploid durum or macaroni wheat and the hexaploid bread wheat (Feldman, 1976).

It can be seen from these hybridizations that wheat has a very complex genome. The size of the wheat genome is 17 x 10^9 base pairs per chromosome, which makes it about five times as large as that of human beings (Devos & Gale, 1993). The big and complex genome makes it very difficult to isolate low abundance or low copy-number genes, and this makes it difficult to study.

3. The Russian Wheat Aphid

3.1 Morphology

The first reported occurrence of the Russian wheat aphid (*Diuraphis noxia* Mordvilko; RWA) was done by Grossheim in 1914. He reported the aphid in the Mediterranean region and southern Russia. It only became a recognised serious pest in South Africa in 1978 (Walters *et al.*, 1980). The RWA was discovered in Mexico in 1980 (Gilchrist *et al.*, 1984) and from there found in Texas, USA (Webster *et al.*, 1987) and in Canada in 1988 (Morrison, 1988).

Aphids are relatively small, the wingless form is about 1.4 to 2.3 mm long, spindle shaped, pale yellow-green or grey-green dusted with white wax powder (Fig 1.1). The winged forms (1.5 - 2.0 mm) also have a pale-green abdomen. The RWA differs from other aphids by having shorter antennae, rudimentary cornicles and a secondary tail above the regular cauda (tail) (Bryce, 1994).

Figure 1.1. The Russian wheat aphid (*Diuraphis noxia* (Mordvilko)). Both winged and wingless forms can be seen. *(Photo courtesy of Lieschen Bahlman, University of Pretoria).*

3.2 Life cycle of the Russian wheat aphid

The RWA has two different forms during its lifetime, namely a wingless form (apterous) and a winged form (alate). No male aphids have been identified in South Africa, so reproduction occurs via parthenogenesis (reproduction without fertilisation) (Walters *et al.*, 1980).

Once a plant has grown so much as to be unable to provide a suitable habitat for the aphids, or if a host plant is under stress, the female aphids start developing wings to enable them to move to a new habitat. Once a new suitable host has been found, the aphids begin feeding and reproducing. A female can produce up to four nymphs a day, and these nymphs can start breeding in about 2 weeks. Any new parts of the wheat that appear are immediately infested, until the ears appear, whereupon the females take flight in their search for a new host (Walters *et al.*, 1980).

3.3 Feeding habits

The RWA feeds on the leaf phloem (Ni & Quisenberry, 1997). The stylet penetrates between the cells (intercellular), but as soon as it reaches the phloem, it probes into the cells (Fouche *et al.*, 1984). It seems as if trichome density has an influence on RWA feeding. The resistant wheat line 'Tugela DN' has a higher trichome density compared to susceptible wheat lines. This may act as an obstacle for RWA and so make it more difficult for them to find a good feeding site on the resistant line (Bahlman *et al.*, 2003).

Once infected, a field will show a patchy distribution pattern rather than a uniform infestation. Groups of 5 to 10 plants carrying aphid will occur in groups throughout the field (Walters *et al.*, 1980). A number of field studies proved that if the aphids were removed before the booting and inflorescence stages of the wheat plant, the plants recovered almost completely (Fouche *et al.*, 1984).

3.4 Effects on plants / Symptoms

The symptoms accompanying RWA infestation are white, yellow and purple to reddish-purple longitudinal streaks on the wheat leaves. The leaves also tend to curl inwards at the edges (Walters *et al.*, 1980). Usually, the newest plant growth is affected and the aphids occur there, in the axils of the leaves or inside the curled-up leaves, where they are partially protected from aphicides and predators. If the plants are heavily infested, they sometimes have a flattened appearance (Walters *et al.*, 1980).

The aphid feeding has an enormous effect on the membranes and chloroplasts, causing complete degeneration of both (Van der Westhuizen & Botha, 1993). These changes are most likely due to a phytotoxin injected by the aphid during feeding and not a pathogen as was first thought (Fouche *et al.*, 1984). During infestation, altered protein expression patterns occur in the plants when total proteins and specific pathogenesis-related proteins are differentially expressed, for example chitinases, beta-1-3-glucanases and peroxidases (Van der Westhuizen & Botha, 1993; Botha *et al.*, 1998; Van der Westhuizen *et al* 1998a,b; Bahlman, 2002).

3.5 Pest management

The most common management strategies used against the RWA are chemical control and cultural practices. With cultural practices the main choices are delaying the crop plantings, the control of volunteer plants or even using non-host crops. Systemic aphicides are used during chemical control (Elsidaig & Zwer, 1993). These are more expensive than contact aphicides, but work much better, since the aphids hide in the curled-up leaves and are then partially protected from the aphicide. It is sometimes better to use a combination of aphicides than using a single type (Walters *et al.*, 1980).

In South Africa there are about seven species of ladybirds, seven species of wasps and two species of flies that are natural enemies of the RWA (Hayes, 1998). However, these predators are not very effective as a biological control agent, since the RWA is not their sole source of food. The RWAs are, more likely than not, harder to obtain than other aphids, since they hide in rolled-up leaves (Walters *et al.*, 1980). Prinsloo (1998) reported that a parasitoid, *Aphelinus hordei* (Kurdjumov) introduced into South Africa via Australia from the Ukraine, parasitizes *D. noxia*, without apparent influence on other organisms. It seems as if this may prove to be a useful tool in the control of the RWA problem, but a lot more research is needed before this technique can be implemented in practice.

3.6 Inheritance of wheat resistance to the Russian wheat aphid

To date, ten different resistance genes have been discovered that each confer resistance to the RWA. The different wheat lines containing the different resistance genes have been outlined, as well as the gene response (Table 1.1). The first resistance gene to be identified was the *Dn1* gene (Du Toit, 1989b). This gene occurs in the resistant wheat line 'Tugela DN' and seems to be a single dominant gene (Du Toit, 1989b; Nkongolo *et al.*, 1991a). The *Dn5* gene might be a single dominant gene, but it is commonly believed to be the *Dn1* and *Dn2* genes that occur together (Elsidaig & Zwer, 1993; Marais & Du Toit, 1993; Saidi & Quick, 1996). Not much is known yet about *Dn7 – Dn9* and the latest resistance gene to be identified has been designated *Dnx* (Liu *et al.*, 2001).

Table 1.1. Russian wheat aphid resistance genes identified to date in wheat.

Resistance gene	Line / number	Gene response	Reference
Dn1	PI 137739	Single dominant gene	Du Toit, 1989b; Nkongolo *et al.*, 1991a
Dn2	PI 262660	Single dominant gene	Du Toit, 1989a; Dong & Quick, 1995
Dn3	*Triticum tauschii* (*Aegilops tauschii*)	Single recessive gene	Nkongolo *et al.*, 1991a
Dn4	PI 372129	Single dominant gene	Du Toit, 1989a; Nkongolo *et al.*, 1991b
Dn5	PI 294994	Single dominant gene/ One dominant & one recessive gene/ Two dominant genes	Marais & Du Toit, 1993 Elsidaig & Zwer, 1993 Saidi & Quick, 1996
Dn6	PI 243781	Single dominant gene	Saidi & Quick, 1996
Dn7	Rye	Unknown	Marais *et al.*, 1994
Dn8	PI 294994	Unknown	Liu *et al.*, 2001
Dn9	PI 294994	Unknown	Liu *et al.*, 2001
Dnx	PI 220127	Single dominant gene	Liu *et al.*, 2001

The modes of resistance occurring in these host plants are antixenosis, antibiosis, tolerance or a combination of these factors. Antixenosis can be defined as non-preference, as the aphid will leave a resistant plant in preference to another, less resistant plant (Rafi *et al.*, 1996). Antibiosis leads to a decrease in aphid bodysize, longevity and reproduction. In other words, it has an effect on the aphids' biological

functions (Ungerer & Quisenbury, 1997). Tolerance refers to the amount of damage that occurred in resistant versus susceptible plants after infestation. Resistant plants will show less damage, in other words, be more tolerant of aphid infestation (Kindler *et al.*, 1995).

4. Plant resistance: how, when and where

The sudden occurrence of a new disease or pest may be due to a number of factors. These may include rapid changes in agronomic practice, the release of a widely used, but very susceptible cultivar, the introduction of new pathogen or pest species from outside the zone of interest or the introduction or evolution and increase of new pathotypes/biotypes. Another factor that may occur is a chronic disease or pest infestation that goes unnoticed because it is thought to be part of the environment. Low crop yields are accepted as normal or a different crop is planted. Too high levels of resistance conferred to the plants are also not desirable. This may lead to accelerated evolution in biotypes/pathotypes and so the whole process would have to be started over again. It seems to work better to have less effective but more durable resistance, but by far the best would be to have resistance based on additive gene effects. In the latter case there would probably be no limit to the actual level of resistance that may be obtained (McIntosh, 1998).

The best way to produce crop plants with increased disease resistance is by understanding and utilising their resistance (R) genes. This can be done by classical breeding techniques or by directly engineering the crops (Staskawicz *et al.*, 1995).

The problem with classical breeding techniques is that it takes very long for positive results. Then, after many years of breeding, a resistant plant variety is produced and a few years later the pathogen has evolved in such a way, that the host plant isn't resistant anymore (Staskawicz *et al.*, 1995), due to breakdown of resistance. Population genetic theory predicts that the breakdown of resistance will happen more slowly in mixtures carrying an array of different R genes (Wolfe, 1985). The same pathogen can be overcome by several different R genes, so plant varieties that are made up of a mixture of lines that differ only in the R gene allele they carry, would be more likely to survive (Staskawicz *et al.*, 1995). This is known as gene pyramiding.

4.1 Modes of resistance

Most plants have natural resistance to various pathogens, toxins and other harmful substances. Plant resistance can be divided into the following categories: disease escape, tolerance, resistance mechanisms and genetic resistance (Jones & Clifford, 1983).

Disease escape constitutes plants that literally escape disease by not being a suitable host. This can be due to physical factors such as the type of flowering, for example a closed flower habit prevents a pathogen from entering through the stigma. Disease escape does not figure prominently in the strategy of modern plant-breeding programs (Jones & Clifford, 1983).

Tolerance refers to a plants' ability to recover after becoming diseased. The higher the tolerance, the better the recovery of the plant. Recovery is measured against the crop yield. Tolerance can only be present if the disease and crop loss is not the same concept, as in *Ustilaga nuda* on barley, where the fungus replaces the grain of the infected plant. The phenomenon of compensation can also occur where, for example, an increased number of wheat grains per ear may compensate for damage due to the pathogen (Jones & Clifford, 1983). Tolerance is a very useful mechanism, since it does not place selective pressure on the pathogen. Tolerance is one of the mechanisms by which some wheat lines provide resistance to RWA infestation (Liu *et al.*, 2001).

Resistance mechanisms in plants can further be subdivided into active mechanisms, passive mechanisms and physiological resistance (Jones & Clifford, 1983). Active resistance occurs when a plant actively forms proteins to protect itself against attack. Substances are produced to seal off the infected area via cell division or by repairing damaged tissue. Some plants even produce their own "antibiotics" against invading pathogens. The most frequently described active resistance mechanism is the hypersensitive response (Jones & Clifford, 1983).

Hypersensitive response (HR) occurs once a pathogen has entered the cell wall of a plant. It occurs as a localised programmed cell death (PCD) in the spot where the

pathogen has breached the cell wall. PCD is a number of events (a cascade) that is initiated by among others, pathogen recognition (Jabs, 1999) and is in actual fact, the cell's active participation in its own demise, also referred to as cellular suicide (Jones, 2001). PCD leads to an arrest in the pathogens' progress, either by killing it or by interfering with its nutrient acquisition (Jones & Clifford, 1983). This enables the plant to protect itself by sacrificing a few cells and so stop an invasion that might have led to the destruction of the whole plant.

Passive resistance is a mechanism that is present in the normal plant at all times and not induced by pathogens or other pests. These mechanisms prevent the initial infection and spread of a pathogen, for example a plant may have a very thick cuticle as a structural barrier. Some plants have unique stomatal structures that prevent pathogens entering (Jones & Clifford, 1983). Ni and Quisenberry (1997b) reported that wheat lines with longer trichomes proved to be least preferred by RWAs. These lines had a trichome density that was less than other wheat lines, but the trichomes were positioned along the leaf veins. Bahlman *et al.* (2003) found that the resistant wheat line 'Tugela DN' has more trichomes on the leaf veins, which is the preferred feeding site of RWA, than other non-resistant lines.

Systemic acquired resistance has been observed more than a century ago. When a plant is infected, be it with a virus, pathogen or bacteria, the plant develops a "memory" and upon a second infection with the same, or closely related pathogen, the plant shows an increased resistance (Ryals *et al.*, 1994).

4.2 Genetic resistance

Plant resistance can be categorised as major gene resistance, gene-for-gene resistance, polygenic resistance or general resistance. In plants with major gene resistance, a single gene usually controls resistance. This type of resistance, also known as monogenic resistance, can usually be easily identified, even in seedlings, since it is usually very specific towards a certain pest or pathogen (Jones & Clifford, 1983).

The gene-for-gene concept states that in certain cases where interaction between a plant and a pathogen occurs, the resistance gene (R gene) in the host corresponds to

and is directed against an avirulence gene (*avr* gene) in the pathogen (Flor, 1971). The first time a clarified version of HR-mediated disease resistance was done, was by Flor on flax (1947) that showed that the resistance of flax to the fungal pathogen *Melampsora lini* was due to the interaction of paired cognate genes in the host and the pathogen. This research laid the groundwork for the gene-for-gene hypothesis of plant-pathogen interaction, as well as the basis for cloning of pathogenic *avr*-genes and their corresponding plant R genes.

Gene-for-gene interaction occurs when a plant recognises specific signal molecules (elicitors) produced by an invading pathogen which leads to induction of the plants' defence response and thus to a hypersensitive response (HR) (Fig. 1.2). The *avr*-genes are responsible for directly or indirectly encoding the elicitors and R genes are thought to encode receptors for these elicitors. As soon as the elicitors are recognised, a cascade of host genes are activated and this leads to HR and inhibition of pathogen growth (Keen, 1990; Lamb, 1994; Dangl, 1995; Heath, 1998).

The HR in plants include a rapid oxidative burst, ion fluxes characterised by K^+-H^+ exchange, cellular decompartmentalisation, cross-linking and strengthening of plant cell walls, production of antimicrobial compounds (phytoalexins), and introduction of pathenogenesis-related (PR) proteins such as chitinases and glucanases (Keen, 1990; Dangl, 1995; Lamb, 1994; Heath, 1998). The RWA induces altered protein expression patterns in the infested wheat plants. Total proteins such as chitinases, glucanases and peroxidases are differentially expressed (Botha *et al.*, 1998; Van der Westhuizen 1998a,b). These events are characteristic of a defence response in a plant, irrespective of the pathogen, although variations may occur in the timing, cell autonomy or intensity of the response (Staskawicz *et al.*, 1995).

Polygenic resistance seems to involve a number of genes at different loci, where each has a small individual, but a combined, additive effect (Jones & Clifford, 1983). This type of resistance is usually conferred against all races of a given pathogen. General resistance implies a non-specific resistance against not one, but many pathogens. This can be monogenic or polygenic (Jones & Clifford, 1983).

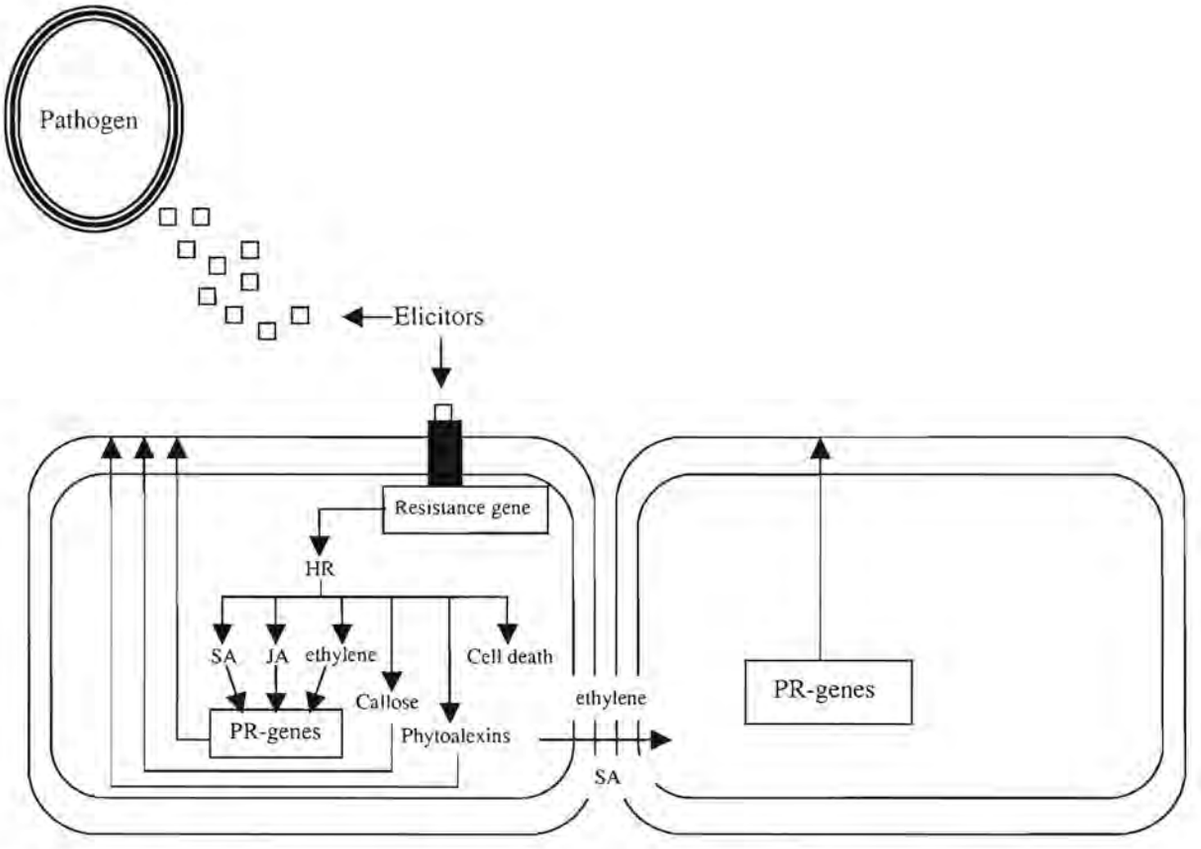


Figure 1.2 Schematic representation of defense responses activated in a plant-pathogen interaction (adapted from Melchers & Stuiver, 2000). On the plants' recognition of pathogen elicitor molecules, defense responses include cell death, callose deposition at the entry site & production of hormones e.g. salicyclic acid (SA), jasmonic acid (JA) and ethylene, which in turn trigger PR (pathogenesis related)-gene product synthesis. This may include production in neighbouring cells.

4.3 Resistance in aphid-plant interaction

Aphids secrete two types of saliva along the stylet path and at the feeding site. The first is a rapidly gelling, sheath saliva and the second, a watery, digestive saliva. The sheath saliva consists of protein, phospholipids and conjugated carbohydrates. This is used to form a protective barrier along the stylet path so that the stylet does not come in direct contact with the plant apoplast. The watery, digestive saliva contains a huge number of enzymes, including pectinases, cellulases, amylases, proteases, lipases, alkaline and acidic phosphatases and peroxidases (Miles, 1999). The elicitors may correspond to one of the salivary components and so induce resistance. A study done on caterpillars showed that their secreted saliva counteracts the amount of toxic nicotine released by *Nicotiana tabacum* due to their feeding (Musser *et al.*, 2002).

Piercing or sucking herbivores, like aphids, lead to very complex signals in plants. Some signals, for example activation of pathenogenesis-related (PR) gene expression, are common between different types of herbivores, whereas other elicitors tend to cause species specific responses (Van de Ven *et al.*, 2000). These signals may be caused by physical damage to the plant or by mechanical stress. This, however, is not always the case. Some plants respond to the herbivore salivary excretions (Musser *et al.*, 2002). The RWA injects a phytotoxin that leads to the resistance response in wheat (Fouche *et al.*, 1984). When an aphid is probing a plant with its stylet, cells may be damaged along the feeding path, which may act as a signal to express certain genes. Even the movement of the stylet between cells may disrupt cell-to-cell contact, which may be seen as physical stress and lead to resistance gene expression (Walling, 2000). However, Botha *et al.* (1998) showed there is a distinct difference between the responses observed after feeding of the RWA, exogenously applied ethylene and mechanical wounding of the plant.

4.4 Resistance genes: their cloning and characterisation

Since the start of molecular cloning of R genes, researchers found that even in genes that confer resistance to different pathogens, the proteins encoded by these genes have several features in common (Staskawicz *et al.*, 1995). This discovery led to the hypothesis that different plants, resistant to different pathogens, may have evolutionary similar signal transduction mechanisms (Staskawicz *et al.*, 1995). The first R gene was only cloned and characterised at molecular level in 1992 (Lamb, 1994). Since then, sequence analysis of cloned R genes showed that most of the encoded proteins contain leucine-rich repeats (LRRs). These motifs are found in many plant and animal proteins and are usually involved in protein-protein interaction (Bent *et al.*, 1994).

The first successful cloning of a plant R gene was the maize *Hml* gene (Johal & Briggs, 1992). This gene is responsible for the resistance to race 1 isolates of *Cochliobolus carbonum*. The gene was identified by transposon tagging, done with the maize (*mu*) transposon. However, the first R gene to be cloned that corresponds to a classic gene-for-gene strategy was the tomato *PTO* gene (Martin *et al.*, 1993).

The *PTO* gene is responsible for resistance to strains of *P. syringae* pv *tomato* (*Pst*) carrying the avirulence gene *avrPto* (Ronald *et al.*, 1992; Fig. 1.3). A map-based cloning strategy was used together with a RFLP marker linked to *PTO* and a yeast artificial chromosome (YAC) was identified that spanned the *PTO* region. This YAC clone was used as a probe to identify cDNAs corresponding to the *PTO* region and the specific clone was identified. The translation product of *PTO* predicts that it encodes a serine-threonine protein kinase that may play a role in signal transduction (Ronald *et al.*, 1992).

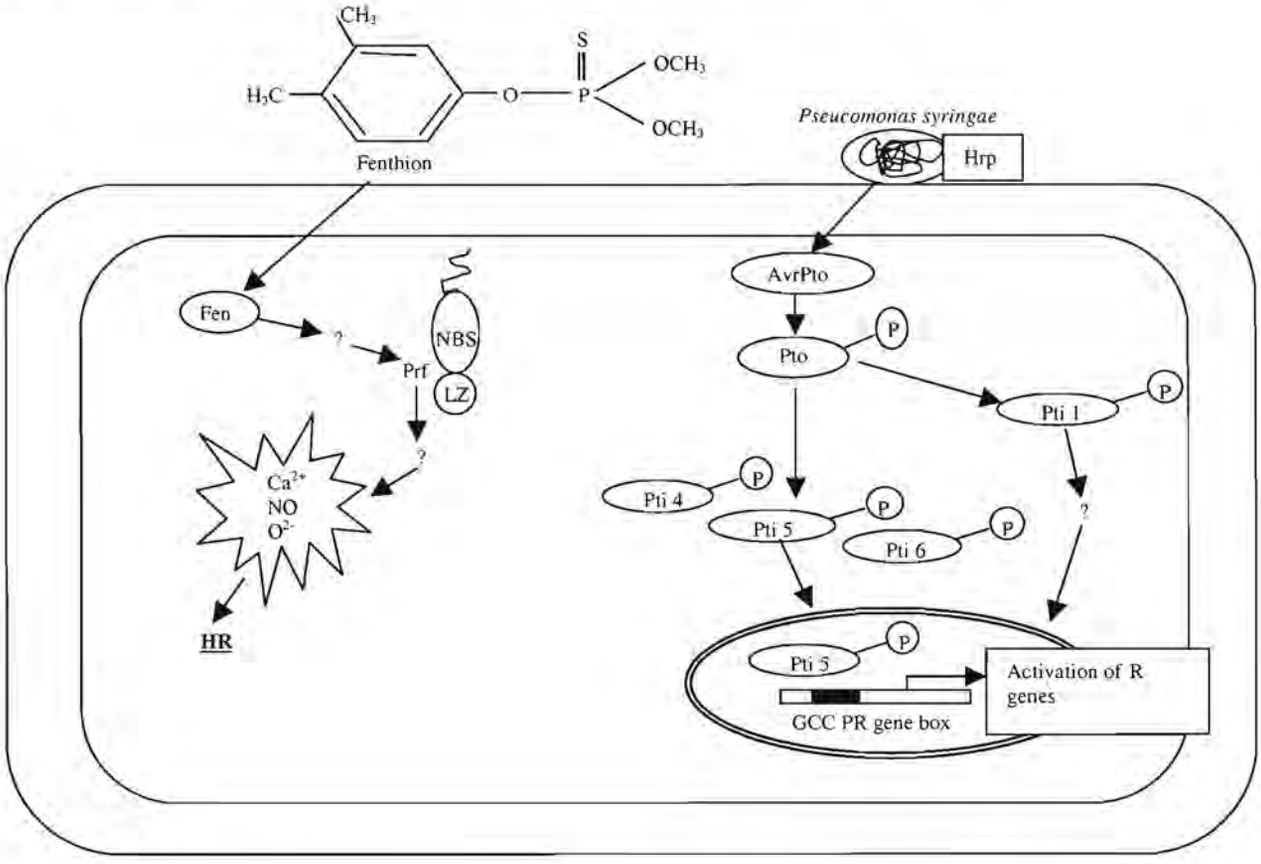


Figure 1.3 Schematic representation of the putative Pto-mediated resistance signaling pathway in tomato (adapted from Hammond-Kosack & Jones, 2000). Research has shown that the direct interaction between Pto kinase and the bacterial AvrPto gene is not necessary for interaction with the Pti proteins or for *in vitro* phosphorylation of Pti 1. The NBS-LRR protein Prf perhaps evolved to "guard" Pto and recognize the AvrPto-Pto complex, so that the defense response could be initiated in addition to those triggered by transcriptional factors Pti1, 4, 5 & 6. Also, the insecticide Fenthion may cause sensitivity, leading to induced cell death, that is caused by Prf recognizing the Fen-complex.

Plant resistance genes can be divided into five defined classes (Table 1.2). Three of the classes of R genes have leucine-rich repeats in common. This leads to the assumption that a major part of the R genes in wheat and other plant species will show homology to leucine-rich repeats. It seems as if disease resistance genes (R genes) are a large group of related sequences in plant genomes and most belong to a gene family that encodes nucleotide-binding proteins (Meyers *et al.*, 1999; Pan *et al.*, 2000).

Table 1.2 The different classes of R genes.

Class	Gene	Plant	Working of R protein	First Identified by
1	*Hm1*	Maize	Detoxifying enzyme	Johal & Briggs, 1992
2	*Pto*	Tomato	Intracellular serine/ threonine protein kinase	Martin *et al.*, 1993
3a	*RPS2*	*Arabidopsis*	NBS-LRR	Bent *et al.*,1994
3b	*N*	Tobacco	Toll-NBS-LRR	Whitham *et al.*,1994
4	*Cf-9*	Tomato	Extracellular LRR protein	Jones *et al.*,1994
5	*Xa-21*	*X. oryzae*	Extracellular LRR protein	Song *et al.*, 1995

Serine- Threonine kinases play an important role in signal transduction during gene-for-gene plant disease resistance (Bent, 1996). A large number of protein kinases have been biochemically confirmed and these sequences are available. The kinase structures identified to date are 11 subdomains, 15 invariant amino acid residues and conserved regions responsible for phosporylation of serine-threonine residues (Bent, 1996). A good example of the above is the tomato *Pto* gene (Martin *et al.*, 1993).

4.5 Nucleotide binding sites- Leucine-Rich Repeats (NBS-LRRs)

The largest group of R genes have been identified as the nucleotide binding site-leucine-rich repeats (NBS-LRRs) (Meyers *et al.*, 2003). There is a hypothesis called the "guard hypothesis" that states that NBS-LRR proteins act as guards of the plant against pathogen effector proteins, where the pathogen products fulfil the role of virulence factors, thus augmenting the plants' vulnerability when no recognition occurs (Dangl & Jones, 2001).

The nucleotide-binding domain (NBS) consists of a number of short amino-acid motifs that appear to be highly conserved regions (Meyers *et al.*, 1999; Pan *et al.*, 2000). These domains occur in diverse proteins with ATP or GTP binding activity. This nucleotide binding ability has been demonstrated in studies on the tomato I2 and Mi R proteins (Tameling *et al.*, 2002). Since this highly conserved region also occurs in some R gene products, it may indicate that their activity is dependent upon the binding of a nucleotide triphosphate (Bent, 1996). The specific mechanism by which the NBS functions in plant defence is still unknown.

The NBS proteins can be divided into sub-classes. The first class has a toll/intereukin-1 receptor homology N-terminal to the NBS that gives it the name of toll-interleukin receptor (TIR) class (Hoffman *et al.*, 1999; Pan *et al.*, 2000; Richly *et al.*, 2002). The second class does not have the TIR region, but rather a leucine zipper sequence (LZ) between the NBS and LRR domain. These leucine zippers function to form the coiled-coiled (CC) structure (Bent, 1996; Meyers *et al.*, 1999; Cannon *et al.*, 2002). These CC's comprise of two to five helixes with distinctive amino-acid side chains at the helix-helix interface (Lupas, 1996). No TIR-type genes have been detected in any grass species (Meyers *et al.*, 1999). The NBS domain contains amino-acid motifs that can be used to distinguish between TIR and non-TIR NBS-LRRs (Young, 2000). Plant NBS domains are almost always similar in several amino-acid motifs that appear to be highly conserved. These domains are the P-loop, Kin-1a and 'GLPL' sites (Meyers *et al.*, 1999; Pan *et al.*, 2000). In TIR-NBS-LRRs there is a conserved region situated directly after the P-loop, containing the amino acids LOKKLLSKLL, with a further motif just before the LRR domain with an amino-acid sequence of FLHIACFF (Young, 2000). Neither of these motifs appears in non-TIR NBS-LRRs. However, the latter contain a conserved motif near the P-loop with the amino-acids FDLxAWVCVSQxF and another near the carboxy-terminus of the NBS domain with the amino-acids CFLYCALFP. These amino-acid motifs are absent from TIR-NBS-LRRs.

Due to the conservancy of these domains between species, the use of a PCR-based technique using degenerate primers, allows the amplification and identification of possible resistance genes in other plant species (Meyers *et al.*, 1999). This will lead

to faster identification and mapping of various resistance genes. However, the function of the identified gene will still have to be confirmed.

Leucine-rich repeats (LRR's) are made up of numerous repeats of a motif that is about 24 amino acids in length and that contains hydrophobic leucines at regular intervals (Bent *et al.*, 1994). The function of different LRR's seems to be determined by the other amino acids that occur between the conserved LRR's, while the LRR's only lend the characteristic structure (Bent, 1996). Many available LRR sequences show deviations from the regulated structure expected and some R gene products are modified in such a way that they seem unlikely to form an LRR structure. The functions of LRR's in numerous organisms, ranging from humans to yeast, show that they are involved in protein-protein interaction (Kobe & Deisenhofer, 1994), ligand binding and pathogen recognition (Young, 2000).

Numerous genetic analysis done throughout the years show that most R genes occur as clusters in plant genomes (Hulbert *et al.*, 2001). Studies done on the genomes of *Arabidopsis* and rice indicate this clustering phenomenon also occurring in NBS-LRR genes (Bai *et al.*, 2002; Richly *et al.*, 2002). This clustering may enable the plant to engender novel resistance responses by recombination of these clustered genes (Hulbert *et al.*, 2001). A study done on the *Arabidopsis* genome characterised all the NBS encoding genes with relation to plant resistance (Meyers *et al.*, 2003). This study found 149 NBS-LRR-encoding genes as well as a further 58 NBS genes without LRRs. The TIR and non-TIR classes were then further divided into subgroups according to their intron numbers, encoded protein motifs and phylogenetic analysis.

Non-NBS predicted extracellular LRR proteins do not contain a NBS domain, but only a LRR domain, for example the *Cf-9* gene in tomato (Jones *et al.*, 1994). Analysis of this gene showed that it encodes 28 LRR's and most of them are extracellular. These LRR's seem to be more conserved than those found in the NBS-LRR (Bent, 1996). The N-terminus of these genes seems to have a peptide sequence that facilitates trans-membrane signalling and transport (Bent, 1996).

5. Molecular tools employed to study R genes

Initially, genomic/molecular mapping was the most popular way of studying R genes in plant genomes (Lagundah, 1997; Faris *et al.*, 1999; Li *et al.*, 1999). This involved the mapping of the physical position of a specific R gene onto a genome map. However, no sequence data and thus, functional analysis of these genes could be obtained through this technique.

A common approach used to obtain more functional data on R genes was by using techniques such as AFLPs, RAPDs and RFLPs. Markers were isolated via these techniques and used to screen populations for the complete sequences of specific genes. A study done on barley resistance to powdery mildew, using these techniques (Wei *et al.*, 1999) revealed three NBS-LRR gene families involved in the plant resistance response. Markers derived from AFLPs, RAPDs and RFLPs were used to construct a physical contig of YAC and BAC clones spanning the resistance gene cluster. From this data, 3 NBS-LRR families appeared to be the key elements in the barley plant resistance.

However, the most recent approaches to studying R genes are large-scale genome sequencing, re-sequencing of R-gene clusters, use of degenerate PCR primers to harvest R-gene candidates (as done in this study), comparative genomics and phylogenetic analysis (Young, 2000). *Arabidopsis* has been one of the plants most frequently used for genome studies (Young, 2000; Meyers *et al.*, 2003). Large-scale genome sequencing has been used for the analysis of NBS-LRRs in *Arabidopsis* (Meyers *et al.*, 1999). Based on the results from this study, it was found that TIR-containing sequences outnumbered non-TIR sequences by about three to one. The re-sequencing of gene clusters, also in *Arabidopsis*, led to the discovery that the NBS-LRR sequences on chromosome IV are co-localised with the previously defined R-gene cluster MRC-H (Holub, 1997, Meyers *et al.*, 1999). The use of degenerate primers for the amplification of R-gene candidates has also been employed extensively. A conclusive study done was on *Arabidopsis RPS2* gene and the tobacco *N* gene using NBS degenerate primers from these two plants to amplify the NBS domain in rice, potato, soybean, barley and tobacco (Yu *et al.*, 1996). Comparative genomics and phylogenetic analysis supports the finding of that two distinct NBS-LRR groups exist, namely TIR and non-TIRs (Meyers *et al.*, 1999; Pan *et al.*, 2000).

Another interesting finding resulting from these studies is the fact that no TIR-containing sequences have ever been found in Poaceae.

Much has been achieved in this field of study but since the RWA resistance genes in wheat have not yet been cloned or characterised, to our knowledge, this thesis is aimed at the gathering of information about the genes involved in response to RWA infestation, as well as the involvement of NBS-LRRs in this response.

6. References

Bahlman L. 2002. *Factors affecting the resistance mechanism of the Russian wheat aphid (Diuraphis noxia) on wheat.* M.Sc. Thesis, University of Pretoria, South Africa, pp. 117.

Bahlman L., Govender P. & Botha A-M. 2003. *Leaf epicuticular wax ultrastructure and trichome presence on Russain wheat aphid (Diuraphis noxia) resistant and susceptible leaves.* African Entomology 11 (1): 59-64.

Bai J., Pennill L.A., Ning J., Lee S.W., Ramalingam J., Webb C.R., Zhao B., Sun Q. Nelson J.C., Leach J.E. & Hulbers S. H. 2002. *Diversity in nucleotide binding site-leucine-rich repeat genes in cereals.* Genome Research 12: 1871-1884.

Bent A.F. 1996. *Plant disease resistance genes: Function meets structure.* The Plant Cell 8: 1757-1771.

Bent A.F., Kunkel B.N., Dahlbeck D., Brown K.L., Schmidt R., Giraudat J., Leung J. & Staskawicz B.J. 1994. *RPS2 of Arabidopsis thaliana: A leucine-rich repeat class of plant disease resistance genes.* Science 265: 1856-1860.

Botha A-M., Nagel M.A.C., Van der Westhuizen A.J. & Botha F.C. 1998. *Chitinase isoenzymes in near-isogenic wheat lines challenged with Russian wheat aphids, exogenous ethylene and mechanical wounding.* Botanicla Bulletin of Academia Sinica 39: 99-106.

Bryce B. 1994. *The Russian wheat aphid Diuraphis noxia (Mordvilko).* Factsheet: http://www.ceris.purdue.edu/napis/pests/rwa/facts.txt.

Cannon S.B., Zhu H., Baumgarten A.M., Spangler R., May G., Cook D.R. & Young N.D. 2002. *Diversity, distribution, and ancient taxonomic relationships within the TIR and non-TIR NBS-LRR resistance gene subfamilies.* Journal of Molecular Evolution 54: 548-562.

Dangl J.L. 1995. *Piece de resistance: Novel classes of plant disease resistance genes.* Cell 80: 363-366.

Dangl J.L. & Jones J.D. 2001. *Plant pathogens and integrated defence responses to infection.* Nature 411: 826-833.

Devos K. & Galle M. 1993. *The genetic maps in wheat and their potential in plant breeding.* Outlook in Agriculture 22: 93-99.

Dong H. & Quick J.S. 1995. *Inheritance and allelism to the Russian wheat aphid in seven wheat lines.* Euphytica 81: 299-303.

Du Toit F. 1989a. *Inheritance of resistance in two Triticum aestivum lines to Russian wheat aphid (Homoptera: Aphididae).* Journal of Economic Entomology 82 (4): 1251-1253.

Du Toit F. 1989b. *Components of resistance in 3 bread lines to Russian wheat aphid (Homoptera: Aphididae) in South Africa.* Journal of Economic Entomology 82 (6): 1779–1781.

Elsidaig A.A. & Zwer P.K. 1993. *Genes for resistance to Russian wheat aphid in PI 294994.* Wheat Crop Science 33: 998-1001.

Faris D., Li W.L., Liu D.J., Chen D. & Gill B.S. 1999. *Candidate gene analysis of quantitative resistance in wheat* . Theoretically Applied Genetics 98: 219-225.

Feldman M. 1976. *Wheats, Triticum spp (Gramineae - Triticinae).* In Evolution of Crop Plants (Ed. N.W. Simmonds). Longman. London.

Flor H.H. 1947. *Inheritance of rust reaction in flax.* Journal of Agricultural Research 74: 241-262.

Flor H.H. 1971. *Current Status of gene-for-gene concept.* Annual Review of Phytopathology 9: 275-296.

Fouche A., Verhoefen R.L., Hewitt P.H., Walters M.C., Kriel C.F. & De Jager J. 1984. *Russian aphid (Diuraphis noxia) feeding damage on wheat, related cereals and a Bromus grass species.* Technical Communications. Department of Agriculture, Republic of South Africa 191: 22-33.

Gilchrist L.I., Rodriguez R & Burnett P.A. 1984. *The extent of Freestate streak and Diuraphis noxia in Mexico.* Proceedings of the workshop Barley Yellow Dwarf. Mexico.

Grossheim N.A. 1914. *The aphid Brachycolus noxius, Mordvilko.* Memoirs of the Natural History Museum of the Zemstvo of the Government of Tavrida, Simperopol 3: 35-78 (In Russian).

Hammond-Kosack K. & Jones J.D.G. 2000. *Responses to Plant Pathogens.* In Biochemistry and Molecular Biology of Plants. (Editors Buchanan B., Gruissen W., Jones R; USA) pp 1102-1156.

Hayes P. 1998. *Beating Russian wheat aphid: the farmers' allies.* Farmers' Weekly January 23: 22-25.

Heath M. 1998. *Apoptosis, programmed cell death and the hypersensitive response.* European Journal of Plant Pathology 104: 117-124.

Hoffman J.A., Kafatos F.C., Janeway CA & Ezkowitz RAB. 1999. *Phylogenetic perspectives in innate immunity.* Science 284: 1313-1317.

Holub E. 1997. *Organization of resistance genes in Arabidopsis.* In The Gene-for Gene Relationship in Plant-Parasite Interactions. (Edited by Crute I.R., Holub E.B., Burdon J.J., Farnham, U.K; CAB International pp. 5-26.)

Hulbert S.H., Webb C.A., Smith S.M & Sun Q. 2001. *Resistance gene complexes: Evolution and utilization.* Annual Review of Phytopathology 39: 285-312.

Jabs T. 1999. *Reactive oxygen intermediates as mediations of programmed cell death in plants and animals.* Biochemical Pharmacology 57: 231-245.

Johal G.S. & Briggs S.P. 1992. *Reductase activity encoded by the HM1 disease resistance gene in maize.* Science 258: 985-987.

Jones A.M. 2001. *Programmed cell death in development and defense.* Plant Physiology 125: 94-97.

Jones D.G. & Clifford B.C. 1983. *Cereal Diseases: their pathology and control.* Wiley-Interscience Publications. UK.

Jones D.A., Thomas C.M., Hammond-Kosack K.E., Balint-Kurti P.J. & Jones J.D.G. 1994. *Isolation of the tomato Cf-9 gene for resistance to Cladosporium fulvum by transposon tagging.* Science 266: 789-793.

Keen N.T. 1990. *Gene-for-gene complementarity in plant-pathogen interactions.* Annual Review of Genetics 24: 447-463.

Kindler S.D., Springer T.L. & Jensen K.B. 1995. *Detection and characterization of the mechanisms of resistance to Russian wheat aphids (Homoptera: Aphididae) in tall wheatgrass.* Journal of Economic Entomology 88: 1503-1509

Kobe B. & Deisenhofer J. 1994. *The leucine-rich repeat: A versatile binding motif.* Trends in Biochemical Science 19: 415-421.

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

Lamb C.J. 1994. *Plant disease resistance genes in signal perception and transduction.* Cell 76: 419-422.

Li W.L., Faris J.D., Chittoor J.M., Leach J.E., Hulbert S.H., Liu D.J., Chen P.D. & Gill B.S. 1999. *Genomic mapping of defense response genes in wheat.* Theoretically Applied Genetics 98: 226-233.

Liu X.M., Smith C.M., Gill B.S. & Tolmay V. 2001. *Microsatellite markers linked to six Russian wheat aphid resistance genes in wheat.* Theoretically Applied Genetics 102: 504-510.

Lupas A. 1996. *Coiled coils: new structures and new functions.* Trends in Biochemical Science 21: 375-382.

Marias G.F. & Du Toit F. 1993. *A monosomic analysis of Russian wheat aphid resistance in the common wheat PI294994.* Plant Breeding 111: 246-248.

Marias G.F., Horn M. & Du Toit F. 1994. *Intergenic transfer (rye to wheat) of a gene(s) for Russian wheat aphid resistance.* Plant Breeding 113: 265-271.

Martin G.B. , Brommonschenkel S.H., Chunwongse J., Frary A., Ganal M.W., Spivey R., Wu T., Earle E.D. & Tanksley S.D. 1993. *Map-based cloning of a protein kinase gene conferring disease resistance in tomato.* Science 262: 1432-1436.

McIntosh R.A. 1998. *Breeding wheat for resistance to biotic stresses.* Euphytica 100: 19-34.

Melchers L.S. & Stuiver M.H. 2000. *Novel genes for disease-resistance breeding.* Current Opinion in Plant Biology 3: 147-152.

Meyers B.C., Dickermann A.W., Michelmore R.W., Sivaramakrishnan S., Sobral B.W.& Young N.D. 1999. *Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide binding superfamily.* Plant Journal 20: 317-332.

Meyers B.C., Kozik A., Griego A., Kuang H. & Michelmore R.W. 2003. *Genome-wide analysis of NBS-LRR-encoding genes in Arabidopsis.* Plant Cell 15: 809-834.

Miles P.W. 1999. *Aphid saliva.* Biological Review 74: 41-85.

Morrison W.P. 1988. *Current distribution and economic impact.* Proceedings of the 2nd Russian wheat aphid workshop. USA.

Musser R.O., Hum-Musser S.M., Eichenseer H., Peiffer M., Ervin G., Murphy J.B. & Felton G.W. 2002. *Caterpillar saliva beats plant defences.* Nature 416: 599-600.

Ni X & Quisenberry S.S. 1997. *Distribution of Russian Wheat Aphid (Homoptera: Aphididae) salivary sheaths in resistant and susceptible wheat leaves.* Journal of Economic Entomology 90: 848-853.

Nkongolo K.K., Quick J.S., Limin A.E. & Fowler D.B. 1991a. *Sources and inheritance of resistance to Russian wheat aphid in Triticum species amphiploids and Triticum tauschii.* Canadian Journal of Plant Science 71: 703-708.

Nkongolo K.K., Quick J.S., Peairs F.B. & Meyer W.L. 1991b. *Inheritance of resistance of PI 372129 wheat to the Russian wheat aphid.* Crop Science 31: 905-907.

Pan Q., Wendel J. & Fluhr R. 2000. *Divergent evolution of plant NBS-LRR resistance gene homologues in dicot and cereal genomes.* Journal of Molecular Evolution 50: 203-213.

Perrino P., Laghetti G., D'Antuono L.F., Al-Ajlouni M., Kanbertay M., Szabo A.T. & Hammer K. 1995. *Ecogeographical distribution of hulled wheat species.* Proceedings of the first international workshop on hulled wheats. Italy.

Prinsloo G.J. 1998. *Aphelinus hordei (Kurdjumov) (Hymenoptera: Aphelinidae), a parasitoid released for the control of Russian wheat aphid, Diuraphis noxia (Kurdjumov) (Homoptera: Aphididae), in South Africa.* African Entomology 6(1): 147-149.

Rafi M.M., Zemetra R.S. & Quisenberry S.S. 1996. *Interaction between Russian wheat aphids (Homoptera: Aphididae) and resistant and susceptible genotypes of wheat.* Journal of Economic Entomology 89: 239-246.

Richly E., Kurth J. & Leister D. 2002. *Mode of amplification and reorginazation of resistance genes during recent Arabidopsis thaliana evolution.* Molecular and Biological Evolution 19: 76-84.

Ronald P.C., Salmeron J.M., Carland F.M. & Staskawicz B.J. 1992. *The cloned avirulence gene avrPto induces disease resistance in tomato cultivars containing the Pto resistance gene.* Journal of Bacteriology 174: 1604-1611.

Ryals J., Uknes S. & Ward E. 1994. *Systemic acquired resistance.* Plant Physiology 104: 1109-1112.

Saidi A. & Quick J.S. 1996. *Inheritance and allelic relationships among Russian wheat aphid resistance genes in winter wheat.* Crop Science 36: 256-258.

Song W-Y, Wang G-L, Chen L-L, Kim H-S, Pi L-Y, Holsten T., Gardner J., Wang B., Zhai W-X, Zhu L-H, Fauquet C & Ronald P. 1995. *A receptor kinase-like protein encoded by the rice disease resistance gene, Xa21.* Science 270: 1804-1806.

Staskawicz B.J., Ausubel F.M., Baker B.J., Elllis J.G. & Jones J.D.G. 1995. *Molecular genetics of plant disease resistance.* Science 268: 661 - 667.

Tameling W.I., Elzinga S.D., Darmin P.S., Vossen J.H., Takken F.L., Haring M.A. & Cornelissen B.J. 2002. *The tomato R gene products I-2 and MI-2 are functional ATP binding proteins with ATPase activity.* Plant Cell 14: 2929-2939.

Ungerer L.M & Quisenbury S.S. 1997. *Effect of antibiotic plant resistance on the reproductive fitness of the Russian wheat aphid (Homoptera: Aphididae).* Journal of Economic Entomology 90: 1697-1701.

Van de Ven W.T.G., LeVesque C.S., Perring T.M. & Walling L.L. 2000. *Local and systemic changes in squash gene expression in response to Silverleaf whitefly feeding.* Plant Cell 12: 1409-1423.

Van der Westhuizen A.J. & Botha F.C. 1993. *Effect of the Russian wheat aphid on the composition and synthesis of water soluble proteins in resistant and susceptible wheat.* Journal of Agronomy and Crop Science 170: 322-326.

Van der Westhuizen A.J., Qian X-M & Botha A-M. 1998a. *β-1,3-Glucanases in wheat and resistance to the Russian wheat aphid.* Physiologia Plantarum 103: 125-131.

Van der Westhuizen A.J., Qian X-M & Botha A-M. 1998b. *Differentail induction of apoplastic peroxidase and chitinase activities in susceptible and resistant wheat cultivars by Russian wheat aphid infestation.* Plant Cell Reports 18: 132-137.

Walling L.L. 2000. *The myriad plant responses to herbivores.* Journal of Plant Growth Regulation 19: 195-216.

Walters M.C., Penn F., Du Toit F., Botha T.C., Aalsberg K., Hewit P.H. & Broodryk S.W. 1980. *The Russian Wheat Aphid.* Farming in South Africa leaflet series : Wheat C3.

Webster J.A., Starks K.J. & Burton R.L. 1987. *Plant resistance studies with Diuraphis noxia (Mordvilko) (Homoptera: Aphididae), an new United States wheat pest.* Journal of Economical Entomology 80: 944-949.

Wei F., Gobelman-Werner K., Morrol S.M., Kurth J., Mao L., Wing R., Leister D., Schultze-Lefert P. & Wise R.P. 1999. *The Mla (powdery mildew) resistance cluster is associated with three NBS-LRR gene families and suppressed recombination within a 240-kb DNA interval on chromosome 5S (1HS) of Barley.* Genetics 153: 1929-1948.

Whitham S., Dinish-Kumar S.P., Choi D., Hehl R., Corr C. & Baker B. 1994. *The product of the tobacco mosaic virus resistance gene N: similarity to Toll and the interleukin-1 receptor.* Cell 78: 1101-1115.

Wiese M.V. 1977. *Compendium of Wheat diseases.* American Phytopathological Society. USA.

Wolfe M.S. 1985. *The current status and prospects of multilane cultivars and variety mixtures for disease resistance.* Annual Review of Phytopathology 23: 251-273.

Young N.D. 2000. *The genetic architecture of resistance.* Current Opinion in Plant Biology 3: 285-290.

Yu Y.G., Buss G.R. & Maroof M.A.S. 1996. *Isolation of a superfamily of candidate disease-resistance genes in soybean based on a conserved nucleotide-binding site.* Proclamation of the National Academy of Science, USA 93: 11751-11756.

CHAPTER 2

ANALYSIS OF EXPRESSED SEQUENCE TAGS OBTAINED FROM THE WHEAT LINE 'TUGELA DN'

This data has been published as *Functional and comparative analysis of expressed sequences from Diuraphis noxia infested wheat obtained utilizing the conserved Nucleotide Binding Site* in African Journal of Biotechnology 2(4): 75-81.

Chapter 2:

Analysis of Expressed sequence tags obtained from the wheat line 'Tugela DN'.

1. Abstract

When the resistant wheat line, 'Tugela DN' (*Dn1*, SA1684 / Tugela*5) is infested with the Russian wheat aphid (*Diuraphis noxia* Mordvilko; RWA), differential expression of gene sequences occurs due to the feeding of the aphid. Nucleotide binding sites (NBS) have been found to be conserved regions in many resistance genes from different plant species. In this study, a PCR-based approach was followed using degenerate primers to target and amplify NBS sites from cDNA synthesized from RNA isolated from leaf tissue after infestation with RWA. All amplified fragments were isolated, cloned into the pGem-T Easy Vector system and 80 selected clones were sequenced. All sequences obtained were submitted to GenBank for identification. The amplified sequences grouped into six categories. After analysis of the sequences, it was found that the metabolism category consisted of 38%, resistance comprised 19%, miscellaneous had 16%, structural comprised 17%, regulatory consisted of 9% and protein synthesis had only 1% of the total number of sequences.

2. Introduction

Expressed sequence tags (ESTs) are partial cDNA sequences and are an economical way to gain information about expressed genes in a variety of organisms. ESTs have proven to be a rapid and efficient method of characterising the subset of genes that are expressed in tissue or life-stage specific manner in a wide variety of organisms (Ajioka, 1998). The worldwide effort to sequence expressed genes and chromosomal DNA in rice and *Arabidopsis*, that has been completed, paved the way for using this valuable and informative technique in more complex plants such as wheat and barley (Clarke *et al.*, 1998).

Since it is very important to identify certain desirable phenotypes in wheat, ESTs are a powerful tool. Identification of gene structure and possible function lays the groundwork for identifying the roles of the respective genes. Any sequence discovered can be compared to existing databases. These databases catalogue the structure of genes expressed in certain tissue. International projects have all contributed to these databases (Clarke *et al.*, 1998). One of the most popular search tools is called BLAST (Basic Local Alignment Search Tool) (Altschul *et al.*, 1997).

A problem that occurs is the number of redundant sequences that appear in the existing databases. This number may even be larger than first expected, since partial sequencing sometimes occurs. This involves the sequencing of incomplete cDNA clones which leads to segments of the same gene being classified into independent groups, based on sequence similarity (Yamamoto & Sasaki, 2000). A solution for this problem could be the use of normalised libraries, which brings the representation of the most abundant and least abundant clones to within one order of magnitude of each other within the library (Bonaldo *et al.*, 1996).

A study done on wheat that has been infected with *Fusarium graminearum* (Fusarium head blight) utilised ESTs obtained from a cDNA library to identify genes expressed during *F. graminearum* infection (Kruger *et al.*, 2002). From over 4000 EST's generated, four sets of non-redundant sequences were identified. The first set consisted of biotic and abiotic stress-related genes, the second set contained sequences originating from *F. graminearum*, the third set was sequences with unknown function

and the fourth set had sequences that might have something to do with plant-fungal interaction establishment (Kruger *et al.*, 2002).

Recent studies show that at least 1% of the *Arabidopsis* genome consists of NBS-LRR genes (Sandhu & Gill, 2002). To date, only non-TIR (Toll-Interleukin-I repeats) and no TIR-type genes have been detected in any grass species (Meyers *et al.*, 1999). Due to the similarity between these resistance genes and their encoded proteins, the use of a PCR-based technique using degenerate primers designed to target the conserved NBS regions, allows the amplification and identification of possible resistance gene family members or homologs in other plant species (Meyers *et al.*, 1999). This strategy has been employed in other plants such as potato (Leister *et al.*, 1996), soybean (Yu *et al.*, 1996) and citrus (Deng *et al.*, 2000). A number of studies have been done on the identification and analysis of expressed sequence tags (ESTs) and through this approach a great number of genes expressed in plants have been identified (White *et al.*, 2000; Yamamoto & Sasaki, 2000). This is a very suitable approach for an organism with such a complex genome as wheat. All ESTs generated can also be deposited into the EST databank and so provide a resource for other studies.

The objective of this chapter was to isolate ESTs from the wheat genome using a PCR-based strategy with degenerate NBS primers to target resistance gene candidates. All ESTs generated were deposited into GenBank and a comparative study was done to determine the amount of resistance gene homologs to other housekeeping genes isolated.

3. Materials and Methods

3.1 Plant material

The 'Tugela DN' (SA1684 / Tugela*5) (Du Toit, 1989) wheat line, which is a line resistant to RWA, was used for this study. Wheat was planted in a peat-sand (1:1) mix in a greenhouse, with the temperature kept at an average of 24°C. The plants were watered daily and allowed to grow until the second leaf stage. The wheat plants were infested with Russian wheat aphids (*Diuraphis noxia,* Mordvilko; RWA). Approximately five aphids were applied to each plant with a fine brush and were allowed to feed for three to five days, where after second and third leaves from the infested plants were harvested. All the aphids were removed by rinsing the leaves with water and then wiping the leaves to ensure no aphids were left. This prevented contamination of the sample with aphid nucleic acids. Excess water was removed and the leaves were used immediately for RNA isolation.

3.2 Treatment of equipment and solutions

Care was taken that all glassware, plastic ware and solutions used during the RNA and mRNA procedures were as RNase free as possible. All containers, mortars and pestles were treated overnight in a 0.1% (v/v) diethyl pyrocarbonate (DEPC) solution, covered in foil and outoclaved for 20 minutes at 121°C and finally baked for at least 4 hours at 200°C (Sambrook *et al.*, 1989). All solutions were DEPC-treated, except those that contained Tris (2-Amino-2-(hydroxymethyl)-1,3-propandiol), which were only outoclaved.

3.3 Isolation of total RNA

A modified version of the Chomczynski and Sacchi (1987) RNA extraction method was used as described by Gehrig *et al.* (2000). Collected wheat leaves were frozen in liquid nitrogen and ground to a fine powder. Extraction buffer (4M guanidinium isothiocyanate, 25 mM sodium citrate, 0.5% (w/v) N-lauriel-sarcosine, 2% (w/v) PEG, 0.1M β-mercaptoethanol) was added to the ground leaves. About 1 ml buffer was added to 100 mg sample. The sample was left at room temperature for 10 minutes and then centrifuged for 20 minutes at 10 000 rpm. The supernatant was

transferred to a new tube and 50 µl of 2M NaOAc (pH 4) and 500 µl of phenol:chloroform was added. The sample was left at room temperature for 10 minutes and then centrifuged for 10 minutes at 10 000 rpm, where after the supernatant was transferred to a new tube. An equal volume of isopropanol was added and the RNA was precipitated at –20°C for 1 hour. The sample was centrifuged for 30 minutes at 13 000 rpm and the resulting pellet was washed with 500 µl of 75% EtOH and then centrifuged again for 10 minutes at 10 000 rpm. The pellet was air-dried for 10 minutes and the dissolved in 20 µl of DEPC water. All centrifugation steps were done at 4°C.

The concentration of the RNA was determined spectrophotometrically at an absorbance of 260 nm. The level of protein contamination was determined by the 260/280 ratio of the sample (Sambrook *et al.*, 1989). All samples were analyzed on a 1% (w/v) agarose gel containing ethidium bromide. The samples were developed at 100V for 15 minutes and visualised under UV light. All samples were stored in a freezer at –70°C.

3.4 mRNA purification and cDNA synthesis

mRNA was purified using an Oligotex mRNA midi kit from Qiagen (USA), following the manufacturers instructions. First and second strand cDNA was synthesized by using the Roche Molecular Biochemicals (Germany) cDNA Synthesis System according to the protocol supplied with the kit. An initial concentration of 2µg total RNA was used, since about 10% of this consists of mRNA (Sambrook *et al.*, 1989); mRNA from RWA infested 'Tugela DN' was used for the cDNA synthesis. The cDNA concentration was determined spectrophotometrically at an absorbance of 260 nm and stored at –20°C.

3.5 PCR amplification

Nucleotide binding-site (NBS) degenerate primers were used for the amplification of fragments from the cDNA (Yu *et al.*, 1996). The primers sequences were NB1-5'-GGA-ATG-GG(AGCT)-GG(AGCT)-GT(AGCT)-GG(AGCT)-AA(AG)-AC-3' and NB2-5'-(CT)CT-AGT-TGT-(AG)A(CT)-(AGT)AT-(AGT)A(CT)-(CT)(CT)T-AG)

C-3' (Yu *et al.*, 1996). All PCR reactions were done on a Perkin-Elmer GeneAmp PCR System 9700 DNA thermal cycler (Applied Biosystems). The PCR reactions (20 µl total volume) were each composed of 10 ng DNA, 1x PCR buffer (10 mM Tris-HCl pH 9.0, 50 mM KCl, 0.1% (v/v) Triton X-100), 10 µM dNTPs, 1 mM $MgCl_2$, 10 pmol of each primer, 0.5 U Taq polymerase and Sabax water. The PCR program consisted of (94°C for 5 minutes) x 1 cycle; (94°C for 1 minute, 55°C for 1.5 minutes, 72°C for 1 minute) x 30 cycles and (72°C for 7 minutes, 4°C hold). After amplification, the PCR reaction results were seperated on a 1% (w/v) agarose gel containing ethidium bromide at 100 V for 25 minutes. The resulting bands were visualised under UV light, carefully cut from the gel with a clean, sharp blade and purified with the Geneclean III kit from Bio101 (USA), following the manufacturer's instructions.

3.6 Construction of library

The amplified fragments from the NBS-PCRs were cloned into the pGEM-T Easy Vector system, following the manufacturers' instructions, transformed into commercial competent *E. coli* cells (JM109; $>1x10^8$ cfu/µg DNA transformed) supplied by Promega (USA) and plated onto LB agar plates containing ampicillin (0.05 mg/ml) / IPTG (0.25 g/ml) / X-Gal (20 mg/ml). This system is based on blue/white colony screening, where white colonies indicate a plasmid with an insert and blue colonies indicate a plasmid without an insert. The transformation efficiency was determined (x cfu/transformation dilution plated = x cfu/µg DNA).

3.7 Colony PCR

The inserts in the white colonies were amplified via colony PCR (Gussow and Clackson, 1989) using T7 (5'-TAA-TAC-GAC-TCA-CTA-TAG-GG-3') and Sp6 (5'-ATT-TAG-GTG-ACA-CTA-TAG-AA-3') as primers. Each PCR reaction (10 µl total volume) contained 1x PCR buffer (10 mM Tris-HCl pH 9.0, 50 mM KCl, 0.1% (v/v) Triton X-100), 10 µM dNTPs, 1 mM $MgCl_2$, 10 pmol of each primer, 0.5 U Taq polymerase and Sabax water. A white colony was picked off of the LB agar plate with a toothpick and the toothpick was then swirled in the PCR mix. This method provided sufficient DNA for amplification of the insert. The PCR program consisted of (94°C for 3 minutes) x 1 cycles; (94°C for 30 seconds, 50°C for 30 seconds, 72°C

for 1 minute) x 30 cycles and (72°C for 1.5 minutes, 25°C for 30 seconds) x 1 cycle. After amplification, the PCR reactions were separated on a 1% (w/v) agarose gel containing ethidium bromide at 100 V for 25 minutes. The resulting bands were visualised under UV light, carefully cut from the gel with a clean, sharp blade and purified with the Geneclean III kit, supplied by Bio101 (USA).

3.8 Sequencing

The amplified PCR bands were sequenced using Sp6 and T7 primers for both forward and reverse sequencing. The sequencing reaction consisted of 25 ng DNA, 10 pmol Sp6 or T7 primer and 2 µl Big Dye Terminator Sequencing Reaction mix supplied by Perkin-Elmer. The PCR program consisted of (96°C for 10 seconds, 50°C for 5 seconds, 60°C for 4 minute) x 25 cycles. The sequencing reaction was done on an ABI-3100 Prism Automated sequencer.

3.9 Sequence data analysis

All results were analyzed using the Sequence Navigator programme version 1.0.1 (Applied Biosystems) on the Apple MacIntosh computer. The sequence identities were obtained after subjecting the sequences to BLASTX (www.ncbi.nlm.nih.gov/BLAST) for alignment to other published sequences in GenBank (Altschul *et al.*, 1997). Functions were assigned to the sequences based on the results returned from searches using the BLASTX algorithm. Any ESTs that did not produce a BLASTX hit were considered to have an unknown function. The e-value indicated the likelihood of the query sequence being the same as the sequence it showed homology to. The lower the e-value the more likely it is that the sequences are the same. The lengths of the homologous sequences were also taken into account. If less than 20% of the submitted sequence showed homology to any gene, it was regarded as an unknown sequence (Kruger *et al.*, 2002). Sequences that produced hits to proteins with e-values greater than 10^{-5} were also considered to have unknown functions (Kruger *et al.*, 2002).

All sequences were divided into categories. The different categories were quite broad and were based on the homology of the sequences. If, for example, the sequence were

homologous to chloroplast, it would be grouped in the metabolism category. The different categories were resistance, metabolism, protein synthesis, structural and miscellaneous. Sequences with hits to proteins with no discernable function were placed into the miscellaneous category and sequences with hits to any plant defence were placed into the resistance group.

4. Results

All the RNA samples were visualised on a 1% (w/v) agarose gel to confirm their integrity. Visible 25S and 18S rRNA bands indicate an intact RNA sample (Fig.2.1). If DNA contamination was observed, the sample was discarded. The RNA was subjected to mRNA purification (as described above) and cDNA was synthesized. After cDNA synthesis, the cDNA is only visible as a smear on a gel (not shown).

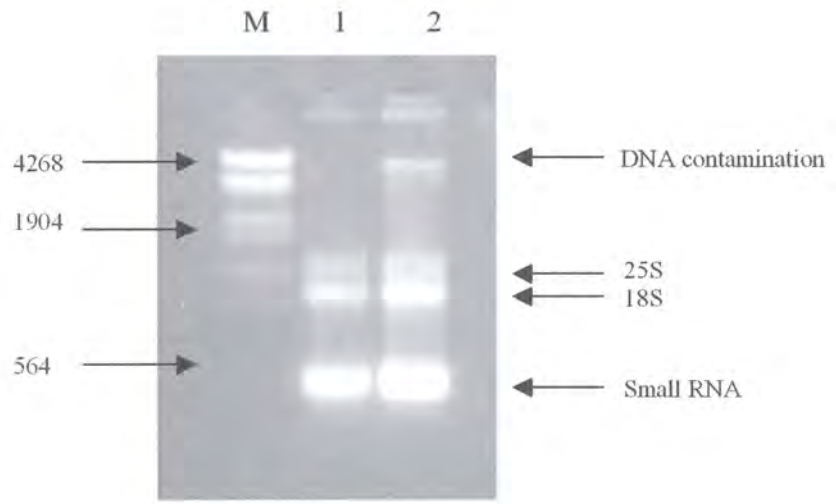


Figure 2.1. Total RNA extracted from infested (1) and uninfested (2) 'Tugela DN'. Lane M = marker III (λ DNA digested with *Eco*R1 and *Hind*III).

The concentrations of the samples were determined. A working concentration of 10 ng cDNA was used in the PCR reactions. The NB1 and NB2 primers were used to amplify possible NBS-LRR sites from the constructed cDNA (Fig. 2.2). All

amplified products between 500 and 600 bp in size were isolated from the gel, cloned into the pGEM-T Easy Vector system and *E. coli* was transformed.

Figure 2.2. PCR amplification products obtained with NBS primers on cDNA obtained from infested 'Tugela DN'. Lane M = marker III (λ DNA digested with *Eco*R1 and *Hind*III), lane 2 & 3 = PCR product obtained with NBS primers. The background smear is due to overloading of the samples.

A library was constructed by using the pGEM-T Easy Vector cloning kit and by using cDNA from the 'Tugela DN' wheat line (resistant) for PCR amplification with NBS primers. A transformation efficiency of 5 x 10^7 cfu/µg was obtained. Colony PCR was done on all white colonies and the resulting fragments were approximately 564 bp in size (Fig. 2.3). These PCR products were isolated from the gel and sequenced.

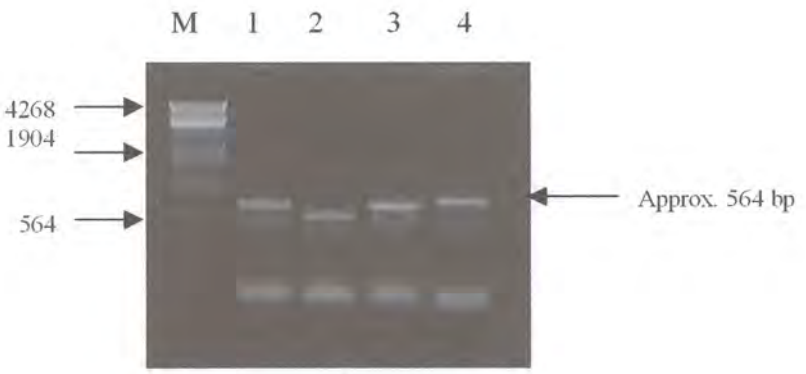


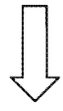
Figure 2.3. Colony PCR done with Sp6 and T7 primers on white colonies obtained after cloning NBS –PCR products. Lane M = marker III (λ DNA digested with *Eco*R1 and *Hind*III), lane 1-4 = PCR products with approximate sizes of 564 bp amplified by Sp6 and T7 primers during colony PCR.

a. gi|343676|gb|M16843.1|WHTCPATPB Wheat chloroplast ATP synthase gene

Expect = 0.0 (e-value)

Nucleotide-nucleotide search:

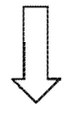
Query: 140 ttctttcttgcaaagaacccatttctgtactaagagtaggttgataacccactgcggaag **199**
Sbjct: 1667 ttctttcttgcaaagaacccatttctgtactaagagtaggttgataacccactgcgga-g **1609**



Query: 500 tgagctttagcaatgttattgattaattccatgatcagtactgtttt **546**
Sbjct: 1308 tgagctttagcaatgttattgattaattccatgatcagtactgtttt **1262**

Nucleotide translated to protein search:

Query:
351GVGKTVLIMELINNIAKAHGGVSVFGGVGERTREGNDLYMEMKESGVINEKNIEESK**372**
GVGKTVLIMELINNIAKAHGGVSVFGGVGERTREGNDLYMEMKESGVINEKNIEESK
175 GVGKTVLIMELINNIAKAHGGVSVFGGVGERTREGNDLYMEMKESGVINEKNIEESK**234**
Sbjct



Query:
371 VYGQMNEPPGARMRVGLTALTMAEYFRDVNKQDVLLFIDNIFRFVQAGSEVSAL **192**
VYGQMNEPPGARMRVGLTALTMAEYFRDVNKQDVLLFIDNIFRFVQAGSEVSAL
235 VYGQMNEPPGARMRVGLTALTMAEYFRDVNKQDVLLFIDNIFRFVQAGSEVSAL **294**
Sbjct

Figure 2.4. An example of part of a BLASTX result obtained after a sequence was submitted. The query line represents the submitted sequence obtained from the experiment and the subject line shows the sequence to which homology was found in GenBank. Only the first and last parts of the sequences are shown. The e-value = 0 that indicates a 100% probability that the submitted sequence is part of the GenBank gene, in this case a wheat chloroplast ATP syntase gene. a= GenBank accession number is given in the first line. (-- indicates gaps in the sequence homology.)

The sequence data was submitted to BLASTX for further analysis. The obtained results indicated homology to a number of other sequences in GenBank (Table 2.1). Some submitted sequences showed an e-value = 0. This indicated complete homology of the submitted sequence to an existing sequence in GenBank (Fig. 2.4). Many of these sequences proved to be housekeeping genes, for example chloroplast sequences. All results obtained had an accession number to identify the sequence for further analysis. All of the sequences showing e-values higher than 10^{-5} were considered having an unknown function (Table 2.1; Kruger *et al.*, 2002).

Fifteen of the sequences in this study showed homology with resistance genes, for example leucine-rich-like proteins (2e-57) (Anderson *et al.*, 2002). Seven putative resistance genes (RGA-2) were identified from wheat (Wicker *et al.*, 2001) with significant e-values (e.g. 4e-16), as well as a resistance gene analogue in rice (3e-11) (Tada, 1999). All sequences that showed homology to resistance genes were considered important.

Table 2.1. 'Tugela DN' sequences derived from cDNA amplified with NBS primers. cDNA was synthesized from purified mRNA obtained from wheat infested with RWA. The origin of the homologous sequence is given in italics after the sequence identity (BLASTX annotation).

BLASTX annotation	EST accession no.[a]	e-value	NCBI accession no.[b]	No. of hits[c]
Wheat chloroplast ATP syntase gene *(Triticum aestivum)*	B412198, CB412200, CB412201	0	AB042240	16
Resistance gene analogue 2 (RGA2) *(Triticum monococcum)*	CB412202, CB412203 CB412204,CB412205 CB412206, CB412207	7e-13	AF326781	6
Leucine-rich-like protein *(Aegilops tauschii)*	CB412208, CB412209 CB412210, CB412211,CB412212 CB412213, CB412214	2e-57	AF497474.1	7
Retrotransposon MITE *(Hordeum vulgare)*	CB412215, CB412216 CB412217	3e-76	AB022688	3
Noduline-like protein *(Triticum monococcum)*	CB412218, CB412219	2e-43	AF326781	2
ATP synthase *(Aegilops crassa)*	CB412220, CB412221 CB412222, CB412223 CB412224,	0	D11099	6
Actin (ACT-1) gene, partial cds *(Triticum monococcum)*	CB412230, CB412231 CB412232,CB412233 CB412234	4e-16	AF326781	5

BLASTX annotation cont.	**EST accession no.**[a]	**e-value**	**NCBI accession no.**[b]	**No. of hits**[c]
Putative chromosome condensation factor (CCF) *(Triticum monococcum)*	CB412235, CB412236 CB412237, CB412238 CB412239	4e-16	AF326781	5
Chloroplast matK gene for maturase *(Cycas pectinata)*	CB412240,CB412241	6e-26	AB076238	6
Ty1-copia-like retrotransposon partial pol pseudogene, clone Tbn-1 *(Beta nana)*	B412242,CB412243 CB412247	7e-29	AJ489202	1
Predicted membrane protein *(Clostridium acetobutylicum)*	CB412248	1e-03	AE007615-4	1
Conserved hypothetical protein *(Bacillus subtilis))*	CB412249, CB412250	2e-05	E69978	2
Integrase/recombinase *(Brucella melitensis)*	CB412251	1e-03	AE009541	1
Genomic sequence BAC F27F5 from chromosome I *(Arabidopsis thaliana)*	CB412252	1e-08	AC007915.3	1
Genomic DNA, chromosome I *(Oryza sativa)*	CB412254, CB412255 CB412256	e-43	AP004326.3	3
Microsatelite DNA, CA-repeat *(S. salar)*	CB412257, CB412258	5e-30	Y11455	2
Clone tac 923.8 3' Ac insertion site sequence *(Zea mays)*	CB412253	1e-29	AY065582.1	1
Chloroplast matK gene for maturase *(Amia angustifolia)*	CB412199	6e-26	AB076567.1	3
putative resistance protein (RGA-2) *(Triticum monococum)*	CB412225	4e-16	AF326781	1
putative nodulin-like protein (Nll) gene *(Triticum monococum)*	CB412245	4e-16	AF326781	1
PCR-amplified resistance gene analogs linked to resistance loci in rice *(Oryza sativa)*	CB412244	3e-11	AB022168	1

a. Accession numbers allocated to EST sequences sumitted to dbEST/GenBank. (Van Niekerk C. & Botha A.M. *Isolation and Characterisation of cDNA Sequences from Russian wheat aphid induced 'Tugela DN' (Dn1) libraries;* www.ncbi.nlm.nih.gov/BLAST)

b. NCBI accession number of homologous sequences.

c. Number of ESTs that gave BLAST annotations with identical e-values.

The results were combined and sorted into six functional categories, namely resistance, metabolism, protein synthesis, structural, regulatory and miscellaneous (Tabel 2.2).

Table 2.2. Functional categories that all the obtained sequences were sorted into.

BLASTX annotation	No. of sequences	Category [a]
Wheat ATP synthase gene *(Triticum aestivum)*	16	Metabolism
Resistance gene analogue 2 (RGA2) *(Triticum monococcum)*	7	Resistance
Leucine-rich-like protein *(Aegilops tauschii)*	7	Resistance
Retrotransposon MITE *(Hordeum vulgare)*	3	Miscellaneous
Noduline-like protein *(Triticum monococcum)*	2	Structural
ATP synthase subunit *(Aegilops crassa)*	6	Metabolism
Actin (ACT-1) gene, partial cds *(Triticum monococcum)*	5	Structural
Putative chromosome condensation factor (CCF) *(Triticum monococcum)*	5	Structural
Chloroplast matK gene for maturase *(Cycas pectinata, Amia angustifolia))*	9	Metabolism
Met-tRNA gene *(Triticum aestivum)*	1	Protein synthesis
Ty1-copia-like retrotransposon partial pol pseudogene, clone Tbn-1 *(Beta nana)*	1	Miscellaneous
Predicted membrane protein *(Clostridium acetobutylicum)*	1	Structural
Conserved hypothetical protein *(Bacillus subtilis)*	2	Miscellaneous
Integrase/recombinase *(Brucella melitensis)*	1	Regulatory
Genomic sequence BAC F27F5 from chromosome I *(Arabidopsis thaliana)*	1	Miscellaneous
Genomic DNA, chromosome I *(Oryza sativa)*	3	Miscellaneous
Microsatelite DNA, CA-repeat *(S. salar)*	2	Miscellaneous
Clone tac 923.8 3' Ac insertion site sequence *(Zea mays)*	1	Miscellaneous
putative nodulin-like protein (Nll) gene *(Triticum monococum)*	1	Structural
PCR-amplified resistance gene analogs linked to resistance loci in rice *(Oryza sativa)*	1	Resistance

a. Sequences were grouped according to their function allocated during the GenBank search.

The sequences were grouped into categories (Table 2.2). Categories were determined by the sequence function obtained during the BLASTX search. The numbers were determined and the percentages of the different categories were calculated (Fig 2.5).

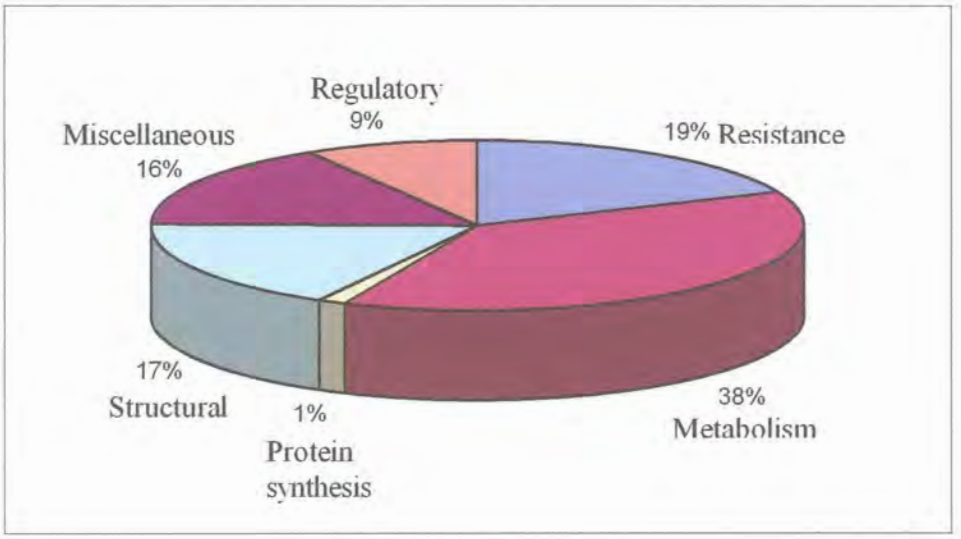


Figure 2.5. The functional categories of the obtained cDNA sequence data after doing a BLASTX search with the sequences. Sequence homology indicated to which group the sequence would be allocated. Metabolism = proteins with a defined metabolic function e.g. energy, lipid or carbohydrate metabolism; Protein synthesis = tRNA ligases.; Structural = cytoskeleton, ribosomal and membrane-bound proteins; Regulatory = integrase & maturase; Miscellaneous = proteins with unknown function or that showed low homology (e-value > 10^{-5}) to sequences in GenBank.

The sequences were grouped into six broad categories. All sequences that showed homology to any resistance genes were grouped to the resistance category, while sequences with unknown functions, as well as those with e-values higher than 10^{-5} fell into the miscellaneous category. Any sequences with a metabolic function, such as energy (ATP) synthesis grouped to the metabolic category. The protein synthesis category consisted of tRNA, the regulatory category contained maturase and integrase, and the structural category contained all sequences involved in the maintenance and building of cells. Obtained results indicated that the most sequences grouped to the metabolism category (38%) (Fig. 2.5). The resistance category was the second largest group with 19% of the sequences, while the miscellaneous group had 16% of the sequences. The structural category contained 17%, regulatory had 9% and protein synthesis comprised only 1% of the sequences.

5. Discussion

A large number of plant disease resistance genes (R genes) that have been identified and sequenced belong to the NBS-LRR family (Young, 2000). Numerous studies have shown that these genes are found in monocotyledons as well as in dicotyledeous plants (Bent *et al.*, 1994; Lagundah *et al.*, 1997). PCR-based methods have been used in the past to isolate NBS genes from plant genomes. This was done by using degenerate primers designed from the NBS regions of known disease resistance genes from different plant species (Leister *et al.*, 1996; Garcia-Mas *et al.*, 2001). By using these primers possible resistance genes could be amplified from various plants.

A study was done on the complete *Arabidopsis* genome to isolate NBS and other resistance genes (Meyers *et al.*, 2003). From this study 149 NBS-LRR genes were identified and 58 NBS genes without LRRs. These genes were divided into two main groups, Toll-Interleukin-1 Receptor (TIR)-NBS-LRR (TNL) and N-terminal coiled-coil (CC)-NBS-LRR (CNL). Four subgroups of CNL and eight subgroups of TNL were defined. This study showed that an individual genotype contains a large number of diverse recombination molecules related to plant resistance.

In our study we also tested the use of such a PCR-based method on a genome as complex as that of wheat. Degenerate oligonucleotide primers designed from conserved motifs in the NBS domain, were used for the amplification of these NBS-LRR regions in the resistant wheat line 'Tugela DN'. Wheat plants were infested with RWAs to create a resistance response in the plants. All amplifications were done on cDNA to ensure expressed genes were amplified. A cDNA library was constructed from the obtained sequences, enabling the screening thereof. The obtained sequences were grouped into categories according to their functions and 19% showed homology to resistance genes. In other more randomised studies done on *Arabidopsis* (White *et al.*, 2000) and rice (Yamamoto & Sasaki, 2000), this was not the case. These studies showed the largest number of ESTs fall into the miscellaneous class as well as the metabolism class. This can be explained by the fact that these studies used random primers, compared to our study that used degenerative specific primers.

However, the large number of ESTs grouped into the resistance class does not necessarily mean they all perform resistance functions, merely that they show homology to published resistance genes. Even though several disease resistance genes have conserved NBS and LRR regions, these regions also occur in other non-resistant genes (Yu *et al.*, 1996). Further studies involving gene expression will have to be done on the identified clones to prove their involvement in resistance response (see Chapter 4). All the ESTs generated were submitted to GenBank and this will prove a valuable source of information for future studies on gene expression or profiling experiments.

Previously, cloned genes were used as probes to identify related genes (Yu *et al.*, 1996). However, utilising PCR for gene isolation is a much more sensitive approach, especially if the target is a conserved region of a gene. This method proved to be an easy-to-use approach, since no difficulties were experienced in the application thereof. The results obtained in this chapter contained no TIR-type genes, confirming previous studies (Meyers *et al.*, 1999). The ESTs in this chapter were used in the microarray study as described in Chapter 4.

Another approach to identify novel sequences found in a specific genome is suppressive subtractive hybridization (SSH). Chapter 3 will discuss the application and results of this technique when applied to wheat.

6. References

Ajioka J.W. 1998. *Toxoplasma gondii: ESTs and gene discovery.* International Journal of Parasitology 28: 1025-1031

Altschul S.F., Madden T.L., Schaffer A.A., Zhang J., Zhang Z., Miller W. & Lipman D.J. 1997. *Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.* Nucleic Acids Research 25: 3389-3402. www.ncbi.nlm.nih.gov/BLAST

Anderson O.D., Rausch C., Moullet O. &Lagudah E.S. 2002. *Characterization of wheat D-genome BAC containing two paralogous HMW-glutenin genes: distribution of genes and retrotransposon clusters.* Submitted to GenBank unpublished.

Bent A.F., Kunkel B.N., Dahlbeck D., Brown K.L., Schmidt R.L., Giraudat J., Leung J.L. & Staskawicz B.J. 1994. *RPS2 of Arabidopsis thaliana: a leucine-rich repeat class of plant disease resistance genes.* Science 265: 1856-1860.

Bonaldo M de F., Lennon G. & Soares M.B. 1996. *Normalisation and subtraction: two approaches to facilitate gene discovery.* Genome Research 6: 791-806.

Chomczynski P. & Sacchi N. 1987. *Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction.* Analytical Biochemistry 162: 156-159.

Clarke B.C., Taylor W., Morell M., Li Z., Rahman S., Ali S. & Appels R. 1998. *Gene transfer and regulation.* Proceedings of the 9th international wheat genetics symposium. Canada.

Deng Z., Huang S., Ling P., Chen C., Yu C., Weber C.A., Moore G.A. & Gmitter F.G. 2000. *Cloning and characterisation of NBS-LRR class resistance-gene candidate sequences in citrus.* Theoretical and Applied Genetics 101: 814-822.

Du Toit F. 1989. *Inheritance of resistance in two Triticum aestivum lines to Russian wheat aphid (Homoptera: Aphididae).* Journal of Economic Entomology 82 (4): 1251-1253.

Garcia-Mas J., Van leeuwen H., Monfort A., De Vincente M.C., Puigdomenech P. & Arus P. 2001. *Cloning and mapping of resistance gene homologues in melon.* Plant Science 161: 165-172.

Gehrig H.H., Winter K., Cushman J., Borland A. & Taybi T. 2000. *An improved RNA isolation method for succulent plant species rich in polyphenols and polysaccharids.* Plant Molecular Biology Reporter 18: 369-376.

Gussow D. & Clackson T. 1989. *Direct clone characterisation from plaques and colonies by the polymerase chain reaction.* Nucleic Acid Research 17: 4000.

Kruger W.M., Pritsch C., Chao S. & Muehlbauer G.J. 2002. *Functional and comparative analysis of expressed genes from wheat spikes infected with Fusarium graminearum.* MPMI 15 (5): 445-455.

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

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

Meyers B.C., Dickermann A.W., Michelmore R.W., Sivaramakrishnan S., Sobral B.W. & Young N.D. 1999. *Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide binding superfamily.* Plant Journal 20: 317-332.

Meyers B.C., Kozik A., Griego A., Kuang H. & Michelmore R.W. 2003. *Genome-wide analysis of NBS-LRR-encoding genes in Arabidopsis.* Plant Cell 15: 809-834.

Sambrook J., Fritsch E.F. & Maniatis T. 1989. *Molecular Cloning: A Laboratory Manual.* Second Edition. Cold Spring Harbour Laboratory, New York.

Sandhu D. & Gill K.S. 2002. *Gene-containing regions of wheat and the other grass genomes.* Plant Physiology 128: 803-811.

Tada Y. 1999. *PCR-amplified resistance gene analogs link to resistance loci in rice.* Breeding Science 49: 267-273.

White J.A., Todd J., Newman T., Focks N., Girke T., De Llarduya O.M., Jaworski J.G., Ohlrogge J.B. & Benning C. 2000. *A new set of Arabidopsis expressed sequence tags from developing seeds. The metabolic pathway from carbohydrates to seed oil.* Plant Physiology 124: 1582-1594.

Wicker T., Stein N., Albar L., Feuillet C., Schlagenhauf E. & Keller B. 2001. *Analysis of a contiguous 211 kb sequence in diploid wheat (Triticum monococcum L.) reveals multiple mechanisms of genome evolution.* Plant Journal 26(3): 307-316.

Yamamoto K. & Sasaki T. 2000. *Large-scale EST sequencing in rice.* Plant Molecular Biology 35: 1350-1364.

Young N.D. 2000. *The genetic architecture of resistance.* Current Opinion in Plant Biology 3: 285-290.

Yu Y.G., Buss G.R. & Maroof M.A.S. 1996. *Isolation of a superfamily of candidate disease-resistance genes in soybean based on a conserved nucleotide-binding site.* Proceedings of the National Academy of Science, USA 93: 11751-11756.

CHAPTER 3

USING SUPPRESSION SUBTRACTIVE HYBRIDIZATION (SSH) TO SCREEN FOR NOVEL SEQUENCES EXPRESSED IN RESPONSE TO RUSSIAN WHEAT APHID INFESTATION

Chapter 3

Using Suppression Subtractive Hybridization (SSH) to Screen for Novel Sequences Expressed in Response to Russian Wheat Aphid Infestation

1. Abstract

Breeding efforts have led to the development of wheat lines that are resistant to RWA infestation, for example 'Tugela DN'(*Dn1*, SA1684 / Tugela*5). By making use of 'Tugela DN' and suppression subtractive hybridization (SSH), fragments were isolated that may be involved in the wheat plant's resistance to RWA infestation. Two subtractions were done using infested 'Tugela DN' as tester with infested 'Tugela' as driver in SSHa and uninfested 'Tugela DN1' as driver in SSHb. The SSH fragments were cloned into the vector pGEM-T Easy and randomly selected clones were sequenced. Obtained sequences were subjected to a GenBank database search using the BLASTX algorithm. All sequences from SSHa and SSHb showed no significant homology (e-value $< 10^{-5}$) with any known proteins. However, Real-Time PCR and Northern blot analysis indicated involvement of three selected sequences in the RWA resistance response through up-regulation from 5-fold to 5.4-fold.

2. Introduction

The isolation of differentially expressed genes can be difficult. Methods used to identify these genes are representational difference analysis (RDA) (Lisitsyn *et al.*, 1993), differential display reverse transcriptase PCR (DDRT-PCR) (Liang & Pardee, 1992) and cDNA-amplified length polymorphisms (cDNA-AFLP) (Money *et al.*, 1996). A big drawback to some of these methods, however, is the difficulty in isolating rare messages. Suppression subtractive hybridization (SSH) is a useful method for the detection of over-expressed or exclusively expressed genes in one cDNA population compared to another (Desai *et al.*, 2000). Two cDNA populations are needed, one in which the specific gene occurs (tester) and another in which the gene is absent or not expressed (driver). Both populations undergo restriction enzyme digestion after which the tester population is divided into two equal parts. Each of these two sub-populations is ligated to a different adaptor, whereupon an excess of driver cDNA is added to each, in this way isolating and removing all common genes between the tester and driver populations. This allows equalization of high and low copy number cDNA's. The two reactions are then combined and allowed to hybridize further. The sticky ends are filled in and two subsequent PCR reactions are done. For the primary PCR, primers that anneal to each of the adaptors are used. This ensures that only hybridized fragments that have both adaptors are amplified. A nested PCR follows to further increase the specificity of the reaction. The products that form can then be cloned and studied (Birch *et al.*, 2000).

SSH has been widely used in the medical field. A study was done on a DNA library using SSH acquired cDNA sequences from testis material as probes. The homologous sequences acquired from the library through this study proved to be unique sequences that are only expressed in the testis (Diatchenko *et al.*, 1996). This study suggests that SSH is very specific and can be applied for various studies including identification of disease or even isolation of differentially expressed genes.

SSH has also been applied in botanical studies. Similar results as above were obtained with the tropical legume *Sesbania rostrata*, where sequences that stimulate root

primordia to cause root outgrowth were isolated. From the 192 SSH clones identified, 26 sequences showed putative up-regulation (Caturla *et al.*, 2002). Another study compared two *E. coli* strains. One strain was uropathogenic (*E. coli* strain 536) and the other was non-pathogenetic (*E. coli* K-12 strain MG 1655). From the SSH 22 fragments were identified, from which 5 showed homology to known virulence determinants and seven were unknown proteins (Janke *et al.*, 2001).

In this chapter SSH was applied in an attempt to isolate novel sequences that were being expressed in the wheat plants after RWA infestation. By using SSH, two cDNA populations could be compared with one another and any novel expressed sequences could be isolated. The first comparison was made between two cDNA populations originating from RWA infested near isogenic wheat lines, 'Tugela DN' (*Dn1*, SA1684 / Tugela*5) (Du Toit, 1989a) that was used as tester cDNA and 'Tugela' that was used as driver cDNA. 'Tugela DN' is a wheat line resistant to RWA infestation and 'Tugela' is susceptible to infestation. The second comparison was made between two cDNA populations from the same resistant wheat line, 'Tugela DN', but where the tester cDNA population was obtained from RWA infested wheat material and the driver population came from uninfested wheat material. The two different subtractions would result in different genes being isolated, since the two driver populations used were different.

3. Materials and Methods

3.1. Plant material

The different wheat lines were planted in a 1:1 peat-soil mixture. The wheat cultivars and lines used in this study were 'Tugela DN' (resistant line), 'Tugela' (susceptible line), *Aegilops speltoides* (BB genome) and *Triticum urartu* (AA genome). The plants were kept in a greenhouse, at a constant temperature of 24°C. The plants were watered daily and allowed to grow until the second leaf stage. The wheat plants were infested with RWA. Approximately five aphids were applied to each plant with a fine brush and were allowed to feed for 3-5 days, where after second and third leaves from the infested plants

were harvested. All the aphids were removed by rinsing the leaves with water and then picking off the remaining aphids by hand. This prevented contamination of the sample with aphid nucleic acids. The excess water was removed from the leaves and nucleic acid isolation was done immediately.

3.2 Treatment of equipment and solutions

Care was taken that all glassware, plastic ware and solutions used during the RNA and mRNA procedures were as RNAse free as possible. All containers, mortars and pestles were treated in a 0.1% (v/v) diethyl pyrocarbonate (DEPC) solutions overnight, covered in foil and outoclaved for 20 minutes at 121°C and finally baked for at least 4 hours at 200°C (Sambrook *et al.*, 1989). All solutions were DEPC-treated, except those that contained Tris (2-amino-2-(hydroxymethyl)-1,3-propandiol), that were only outoclaved.

3.3 Isolation of total RNA

A modified version of the Chomczynski & Sacchi (1987) RNA extraction method was used, as described by Gehrig *et al.* (2000) for total RNA isolation. Different tissue types were sampled from the wheat plants. Leaf tissue was obtained from uninfested 'Tugela', RWA infested and uninfested 'Tugela DN', while stem tissue was only procured from uninfested and RWA infested 'Tugela DN'. The material was frozen in liquid nitrogen and ground to a fine powder. Extraction buffer (4M guanidinium isothiocyanate, 25 mM sodium citrate, 0.5% (w/v) N-lauryl-sarcosine, 2% (w/v) PEG, 0.1M β-mercaptoethanol) was added to the ground leaves. Buffer (1 ml) was added to 100 mg sample. The sample was left at room temperature for 10 minutes and then centrifuged for 20 minutes at 10 000 rpm. The supernatant was transferred to a new tube and 50 µl of 2M NaOAc (pH 4) and 500 µl of phenol:chloroform (1:1) was added. The sample was left at room temperature for 10 minutes and then centrifuged for 10 minutes at 10 000 rpm, where after the supernatant was transferred to a new tube. An equal volume of isopropanol was added and the RNA was precipitated at -20 °C for 1 hour. The sample was centrifuged for 30 minutes at 13 000 rpm and the resulting pellet was washed with 500 µl of 75% EtOH and then centrifuged again for 10 minutes at 10 000 rpm. The pellet was air-dried

for 10 minutes and redissolved in 20 µl of DEPC-treated water. All centrifugation steps were done at 4°C.

The concentration of the RNA was determined spectrophotometrically at an absorbance of 260 nm. The level of protein contamination was determined by the 260/280 ratio of the sample (Sambrook *et al.*, 1989). All samples were analyzed on a 1% (w/v) agarose gel containing ethidium bromide. The samples were run at 100 V for 15 minutes and visualised under UV light. All samples were stored at –70°C.

3.4. mRNA purification and cDNA synthesis

mRNA was isolated using an Oligotex mRNA midi kit obtained from Qiagen (USA) and by following the manufacturers instructions. An initial concentration of 11 µg RNA was used for the mRNA purification to ensure sufficient yield, since only 10% of the total RNA is mRNA (Sambrook *et al.*, 1989). First and second strand cDNA was synthesized by using the Roche Molecular Biochemicals (Germany) cDNA Synthesis System as described in the accompanying protocol. mRNA from all the wheat samples was used for the cDNA synthesis. The cDNA concentration was determined spectrophotometrically at an absorbance of 260 nm and stored at –20°C.

3.5 DNA isolation

Fresh tissue was collected from the wheat plants in the greenhouse. Leaf tissue was collected from 'Tugela DN', *Triticum urartu* and *Aegilops speltoides*, frozen in liquid nitrogen and ground to a fine powder. A DNA isolation method, DEB (DNA Extraction Buffer) was used (Raeder & Broda, 1985). Extraction buffer (200 mM Tris-HCl, 150 mM NaCl, 25 mM EDTA pH 8, 0.5% (w/v) SDS) was added together with an equal volume of chloroform-isoamylalcohol (24:1), mixed and the sample was centrifuged at 10 000 rpm for 30 minutes at 4°C. The supernatant was transferred to a clean tube, chloroform-isoamylalcohol (24:1) was added and the sample was centrifuged at 10 000 rpm for 30 minutes at 4°C. This step was repeated once more. One volume of isopropanol was added, mixed and the DNA was allowed to precipitate at –20°C overnight. The resulting DNA was pelleted at 10 000 rpm for 30 minutes, washed with

70% EtOH and centrifuged at 10 000 rpm for 10 minutes. The pellet was allowed to dry and redissolved in Sabax water. The concentrations of the DNA samples were determined on a spectrophotometer at an absorbency of 260 nm and the protein contamination was calculated with the 260/280 nm ratio (Sambrook *et al.*, 1989). The samples were also run on a 1% (w/v) agarose gel containing ethidium bromide and visualised under UV light. All samples were stored at –70°C.

3.6 Suppression subtractive hybridization

The suppression subtractive hybridization (SSH) was done with the Clontech (USA) PCR-Select cDNA subtraction kit, according to the manufacturers protocol. A schematic representation is given of the procedure (Fig. 3.1). Equal concentrations of the tester and driver cDNA were separately digested to obtain shorter, blunt-ended molecules. The tester was divided into two populations that each received a different adapter, namely adapter 1 (5'-GTA-ATA-CGA-CTC-ACT-ATA-GGG-CTC-GAG-CGG-CCG-CCC-GGG-CAG-GT-3'; 3'- CCC-GTC-CA-5') and adapter 2 (5'- TGT-AGC-GTG-AAG-ACG-ACA-GAA-AGG-GCG-TGG-TGC-GGA-GGG-CGG-T-3'; 3'-GCC-TCC-CGC-CA--5') (Diachenko *et al.*, 1996). The driver cDNA had no adaptors. An adaptor ligation test was done to determine if the adaptor ligation was successful. Two PCR reactions were done, one for each adapter-ligation reaction. These PCR reactions (25 µl) were made up of 1 µl of a 1:200 dilution of tester cDNA, 10 µM of primer PN1 (5'-TCG-AGC-GGC-CGC-CCG-GGC-AGG-T-3'; Diachenko *et al.*, 1996), 10 µM of primer PN2 (5'-AGG-GCG-TGG-TGC-GGA-GGG-CGG-T-3'; Diachenko *et al.*, 1996), 1x PCR buffer (10 mM Tris-HCl pH 9.0, 50 mM KCl, 0.1% (v/v) Triton X-100), 10 mM dNTP mix, 0.5 U Taq polymerase and Sabax water. The PCR program consisted of (5 minutesat 75°C) x1 cycle, (30 secondsat 94°C, 30 secondsat 65°C, 2.30 minutesat 68°C) x 20 cycles and (7 minutesat 68°C, 4°C hold) x1 cycle. The results were separated on a 1% (w/v) agarose gel containing ethidium bromide at 100 V for 20 minutes. The gel was visualized under UV light.

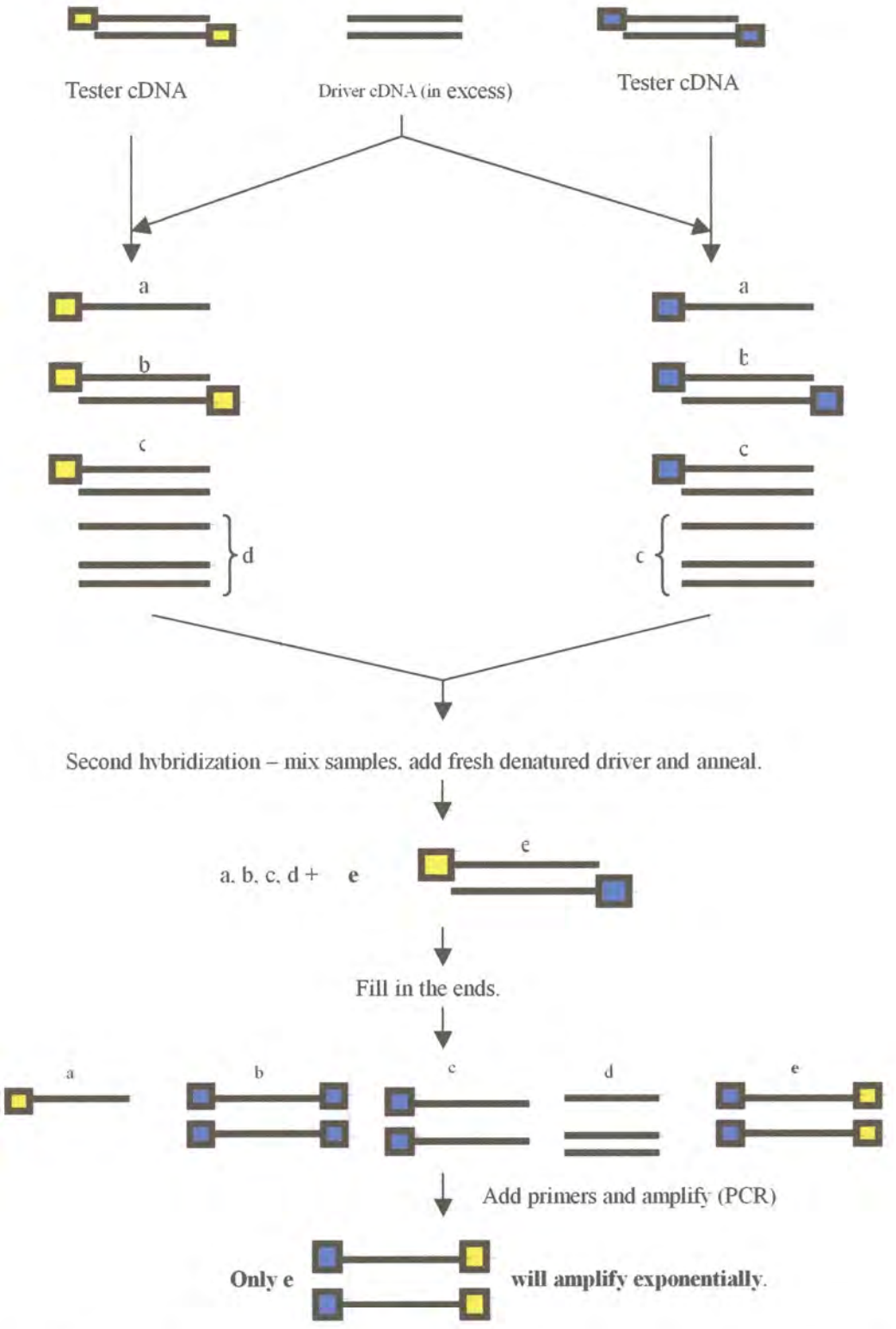


Figure 3.1. Schematic representation of the SSH method (Clontech, 1999). Only sample e will amplify during PCR.

After ligation, 1 µl of each ligation sample was taken, combined and transferred to a new tube. This served as an unsubtracted control reaction during the primary and secondary PCRs. The remaining tester samples were diluted to a concentration 150-fold less than that of the driver. The tester (1.5 µl) and driver (1.5 µl) samples were mixed and denatured at 98°C for 1.5 minutes. One µl hybridization buffer (50mM Hepes pH 8.3, 0.5M NaCl, 0.02 mM EDTA pH 8.0, 10% (w/v) PEG 8000) was added and the samples were allowed to hybridize for 8 hours at 68°C in a PCR machine. After this first hybridization, the two samples were mixed together and 1 µl fresh denatured driver cDNA was added. The reaction was incubated at 68°C overnight. Hybridization kinetics led to equalisation and enrichment of differentially expressed sequences, from which templates for PCR amplification were generated. By using suppression PCR, only differentially expressed sequences were amplified exponentially (Fig 3.1). The PCR reactions (25 µl) contained 1 µl cDNA diluted 1:100 in dilution buffer (20 mM Hepes pH 8.3, 50 mM NaCl, 0.2 mM EDTA), 1x PCR buffer (10 mM Tris-HCl pH 9.0, 50 mM KCl, 0.1% (v/v) Triton X-100), 10 mM dNTP mix, 10 µM Primer 1 (5'-GTA-ATA-CGA-CTC-ACT-ATA-GGG-C-3'; Diachenko *et al.*, 1996), Advantage cDNA Polymerase mix (1x; included in Clontech kit). The PCR cycle consisted of (5 minutes at 72°C) x1, (30 seconds at 91°C, 30 seconds at 54°C, 2.30 minutes at 72°C) x 27 and (7 minutes at 68°C, 4°C hold) x1. Background was further reduced with nested PCR. The PCR reactions (25 µl) contained 1 µl of a 1:10 dilution of primary PCR product, 1x PCR buffer (10 mM Tris-HCl pH 9.0, 50 mM KCl, 0.1% (v/v) Triton X-100), 10 mM dNTP mix, 10 µM primer PN1 (5'-TCG-AGC-GGC-CGC-CCG-GGC-AGG-T-3'; Diachenko *et al.*, 1996), 10 µM primer PN2 (5'-AGG-GCG-TGG-TGC-GGA-GGG-CGG-T-3'; Diachenko *et al.*, 1996), Advantage cDNA Polymerase mix (1x; included in kit) and Sabax water. The PCR conditions were (30 seconds at 94°C, 30 seconds at 64°C, 2.30 minutes at 72°C) x 30 cycles and (7 minutes at 72°C, 4°C hold) x1 cycle. All PCR reactions were done on a Perkin-Elmer GeneAmp PCR System 9700 DNA thermal cycler (Applied Biosystems, USA). After amplification, the PCR reactions were separated on a 1% (w/v) agarose gel containing ethidium bromide at 100 V for 25 minutes. The resulting bands were visualised under UV light, carefully cut from the gel with a clean,

sharp blade and purified with the Geneclean III kit from Bio101 (USA), following the manufacturer's instructions.

3.7 Cloning of PCR products

Amplified fragments from the final PCRs were cloned into the pGEM-T Easy Vector system, following the manufacturers' instructions and transformed into commercial competent *E. coli* cells (JM101; $1x10^8$ cfu/µg DNA transformed) supplied by Promega (UK). The transformants were plated onto LB agar plates containing ampicillin (0.05 mg/ml) / IPTG (0.25 g/ml) / X-Gal (20 mg/ml). This system is based on blue/white colony screening, where white colonies indicate a plasmid with an insert and blue colonies indicate a plasmid without an insert. The transformation efficiency was determined (x cfu/transformation dilution plated = x cfu/µg DNA).

3.8 Colony PCR

The inserts in the white colonies were amplified via colony PCR (Gussow & Clackson, 1989) using T7 (5'-TAA-TAC-GAC-TCA-CTA-TAG-GG-3') and Sp6 (3'-ATT-TAG-GTG-ACA-CTA-TAG-AA-3') as primers. Each PCR reaction (10 µl) contained 1x PCR buffer (10 mM Tris-HCl pH 9.0, 50 mM KCl, 0.1% (v/v) Triton X-100), 10 µM dNTPs, 1 mM $MgCl_2$, 10 pmol of each primer, 0.5 U Taq polymerase and Sabax water. A white colony was picked off of the LB agar plate with a toothpick and the toothpick was then swirled in the PCR mix. This method provided sufficient DNA for amplification of the insert. The PCR cycle consisted of (94°C for 3 minutes) x 1 cycle; (94°C for 30 seconds, 50°C for 30 seconds, 72°C for 1 minute) x 30 cycles and (72°C for 1.5 minutes, 25°C for 30 seconds) x 1 cycle. PCR reactions were done on a Perkin-Elmer GeneAmp PCR System 9700 DNA thermal cycler (Applied Biosystems). After amplification, the PCR reactions were separated on a 1% (w/v) agarose gel containing ethidium bromide at 100 V for 25 minutes. The resulting bands were visualised under UV light, carefully cut from the gel with a clean, sharp blade and purified with the Geneclean III kit, supplied by Bio101 (USA).

3.9 Sequencing

The PCR amplified bands were sequenced with the Sp6 and T7 primers in both directions. The sequencing reaction was 10 µl total volume and contained 25 ng DNA, 10 pmol Sp6 or T7 primer and 2 µl Big Dye Terminator Sequencing Reaction mix supplied by Perkin-Elmer (UK). The PCR cycle conditions were (96°C for 10 seconds, 50°C for 5 seconds, 60°C for 4 minute) x 25. Sequencing was done on an ABI-3100 Prism Automated sequencer.

3.10 Sequencing analysis and functional anotation

All results were analyzed using the Sequence Navigator programme version 1.0.1 (Applied Biosystems) on the Apple MacIntosh computer. The sequences were submitted to GenBank using the BLASTX algorithm (www.ncbi.nlm.nih.gov/BLAST; Altschul *et al.*, 1997) and sequence identities were obtained after alignment to published sequences. Functions were assigned to the sequences based on the results returned from searches using the BLASTX algorithm. Any sequence that did not produce a BLASTX hit was considered to have an unknown function. The e-value indicated the likelihood of the query sequence being the same as the sequence it showed homology to. The lower the e-value, the more likely it is that the sequences are the same. The lengths of the homologous sequences were also taken into account. If less than 20% of the submitted sequence showed homology to any gene, it was regarded as an unknown sequence. Sequences that produced hits to proteins with e-values greater that 10^{-5} were also considered to have unknown functions (Kruger *et al.*, 2002).

3.11 Hybridization blots

Both DNA and RNA dot blots were done to determine if the isolated sequences were part of the wheat genome and in which tissue were they expressed (Sambrook *et al.*, 1989). The DNA used was extracted from leaves of 'Tugela DN' (AABBDD genome), *Triticum urartu* (AA genome) and *Aegilops speltoides* (BB genome). The RNA used was obtained separately from RWA infested 'Tugela' (susceptible line) using leaf tissue and RWA infested and uninfested 'Tugela DN' (resistant line) using leaf, as well as stem tissue. The

sample concentrations were determined spectrophotometrically at an absorbance of 260 nm. All samples were diluted to a concentration of 200 ng/µl, were denatured for 5 minutes in a boiling water bath and 1 µl of each was volume infiltrated onto separate nylon membranes. The membranes were air-dried and the samples were UV cross-linked to the membranes at 0.15 nm/cm^2. The membranes were not stored, but pre-hybridization was started immediately. Pre-hybridization was done at 60°C for four hours using a hybridization buffer containing 5x SSC (20 x SSC contains 0.3 M sodium citrate, 3M NaCl), 1:20 dilution of liquid block (supplied with Gene Images Random Prime labelling kit, Amersham, UK), 0.1% (w/v) SDS and 5% (w/v) Dextran sulfphate.

3.12 Probe labelling and detection

The Gene Images Random Prime labelling kit from Amersham (USA) was used to label the probes for the dot blots following manufacturers' instructions. Five of the sequences obtained from the SSH experiments as well as a control probe were used. The first two probes (ABO 00010 & ABO 00011) came from the cDNA subtraction of RWA infested 'Tugela' and infested 'Tugela DN' and the last three (ABO 00013, ABO 00027 & ABO 00014) from the second subtraction using cDNA from infested 'Tugela DN' and uninfested 'Tugela DN'. As for the control the ubiquitin gene was used. A final concentration of 50 ng of probe was labelled as advised by the manufacturer. The random primed labelling reaction was done by denaturing 20 µl of each DNA sample for 5 minutes in a boiling water bath and then adding 2x dNTPs, 5 µl random primer (supplied with the Gene Images Random Prime labelling kit from Amersham (USA)), 5 U Klenow enzyme and water to a final reaction volume of 50 µl. After the probes were labelled, 20 µl of each probe was transferred to new tubes and denatured in a boiling water bath for 5 minutes. The probes were then added to the membranes directly into the pre-hybridization buffer, avoiding direct application onto the membrane, and allowed to hybridize overnight at 60°C with gentle agitation.

After hybridization the membranes were subjected to stringency washes. All washes were done while gently shaking the membranes. The first wash was done with 1x SSC and

0.1% (w/v) SDS at 60°C for 15 minutes. The second wash was done at the same temperature and for the same duration as the first, but 0.5x SSC and 0.1% (w/v) SDS was used. The detection of the probes was done with the Gene Images CDP-Star detection kit from Amersham (USA), following the kit protocol. The membranes were blocked with blocking solution (1:10 liquid blocking agent in buffer A (100 mM Tris-HCl, 300 mM NaCl ph 9.5)) for 4 hours. Detection was done by adding anti-fluorescein-AP conjugate (1:5000 in fresh 0.5% (w/v) bovine serum albumin; (BSA)) to the membranes for 1 hour with gentle agitation. The membranes were then washed 3 times in 0.3% (v/v) Tween 20 in buffer A and placed on cling wrap. One ml detection agent was pippeted onto each blot for 5 minutes and then the blots were exposed to X-ray HyperFilm for fluorescence (Amersham Life Sciences, UK) for 1 hour. The X-ray film was developed in a dark room by washing it for 5 minutes in developing solution, 1 minute in a 3% (v/v) acetic acid solution and 5 minutes in fixing solution (Sambrook *et al.*, 1989).

3.13 Real-Time PCR

Primers for the reactions were designed from the sequences obtained in the SSH subtractions. The sequences were submitted to the Primer3 web site (www.genome.wi.mit.edu/cgi-bin/primer/primer3.cgi) and resulting primers that gave a product of no less than 150 bp were selected and synthesized by Inqaba Biotech (South Africa) (Table 3.1).

All primers were tested and reactions optimised on a conventional Perkin-Elmer GeneAmp PCR System 9700 DNA thermal cycler (Applied Biosystems, USA). Each PCR reaction (10 µl) contained 50 ng cDNA, 1x PCR buffer (10 mM Tris-HCl pH 9.0, 50 mM KCl, 0.1% (v/v) Triton X-100), 10 µM dNTPs, 1 mM $MgCl_2$, 10 pmol of each primer, 0.5 U Taq polymerase and Sabax water. The PCR cycle consisted of (94°C for 3 minutes) x 1 cycles; (94°C for 30 seconds, 50°C for 30 seconds, 72°C for 1 minute) x 30 cycles and (72°C for 1.5 minutes, 25°C for 30 seconds) x 1 cycle. Real-Time PCR was done on

the LightCycler Instrument (Roche Molecular Biochemicals, Germany). The FastStart DNA Master SYBR Green I kit from Roche was used for the PCR reaction. The reaction

Table 3.1. Primers synthesized from sequences obtained through SSH.

Primer	Sequence	SSH [a]
Ubiquitin (control for quality of cDNA)	f: 5'-CAT-CCA-CCA-GTC-AAG-GGT-TC-3' r: 5'-CTT-CTC-CTC-TAC-CCG-AAC-CC-3'	-----
ABO 00013	f: 5'-CTA-ACA-CCC-TGC-TGG-AAA-GG-3' r: 5'-CAG-CTC-GTC-TCA-AGT-GGA-CA-3'	SSHb
ABO 00027	f: 5'-TCC-AGC-TGC-TTG-TTT-GCT-TA-3' r: 5'-GCT-ACC-GCT-GGA-CAA-CAA-GT-3'	SSHb
ABO 00014	f: 5'-GGT-CGG-ATA-TCT-CGG-CTC-TC-3' r: 5'-TTG-CGT-TCA-AAG-ACT-CGA-TG-3'	SSHb

a. Subtaction from which the sequence originated i.e. SSHa or SSHb

had a final volume of 10 µl and contained 70 ng cDNA, 3 mM $MgCl_2$, 10 pmol of each primer and 1 µl of the Reaction mix (FastStart Taq DNA polymerase, reaction buffer, dNTP mix, SYBR Green I dye and 10 mM $MgCl_2$) provided in the kit. The program cycle consisted of pre-incubation (10 minutes at 95°C) x1 cycle, amplification (10 seconds at 95°C, 5 seconds at specific annealing temperature, 10 seconds at 72°C) x 45 cycles, melting curve analysis (0 seconds at 95°C, 15 seconds at 65°C, 0 seconds at 95°C) x 1 cycle and cooling (30 seconds at 40°C) x 1 cycle. A minimum of seven reactions had to be done for each SSH fragment tested. Reactions 1-4 consisted of a dilution series of tester cDNA (1, 1/10, 1/100, 1/1000) to set a standard curve that was then used for sample quantification. Reaction 5 and 6 were the test reactions, using a 1/10 dilution of tester cDNA for reaction 5 and a 1/10 dilution of driver cDNA for reaction 6. Reaction 7 was the negative control that contained no DNA. All runs were done in duplicate. The results obtained were analyzed by using the LightCycler Software version 3.5 and quantification analysis and a melting curve was generated. The

quantification results generated a standard curve (Fig. 3.2), an error value and an regression coefficient (r). The error value indicated the tube to tube variations e.g. pipetting errors. Only runs with error values smaller than 0.2 were considered significant. The r-value gives an indication of systematic errors for example accumulated error in the dilution series. All runs had to have an r-value of –1 or the run was repeated. The standard curve depended on the amplification efficiency of the standard samples. To ensure an accurate quantification, the standard curve should form a straight line (Fig. 3.2).

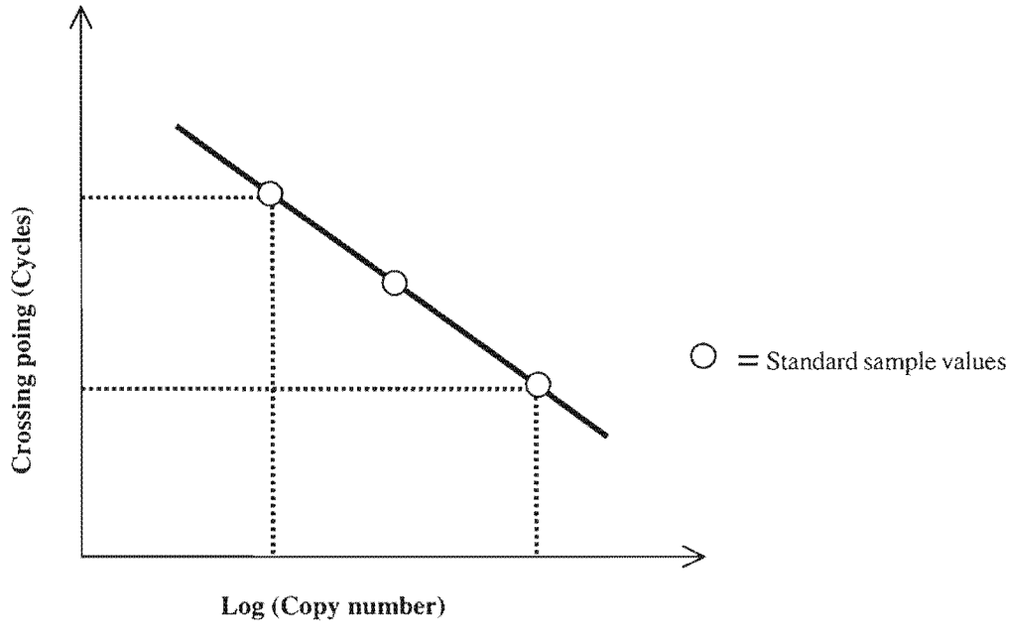


Figure 3.2. An example of a successful Real-Time PCR standard curve.

4. Results

RNA was extracted from the relevant wheat, mRNA was purified and cDNA was synthesized. Adaptors were ligated to the tester cDNA of both hybridizations and the ligations tests were done (Fig. 3.3).

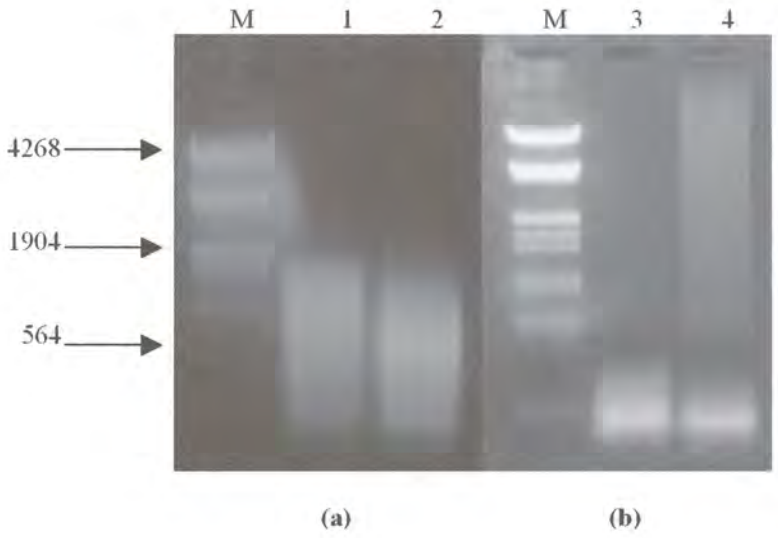


Figure 3.3. Ligation test results of both hybridizations. Lanes M = Marker III (λ DNA digested with *Eco*R1 and *Hind*III), **(a)**: lane 1 = adaptor 1 ligated to tester cDNA from SSHa (cDNA derived from RWA infested 'Tugela DN'), lane 2 = adaptor 2 ligated to tester cDNA from SSHa, **(b)**: lane 3= adaptor 1 ligated to tester cDNA from SSHb (cDNA derived from RWA infested 'Tugela DN'), lane 4 = adaptor 2 ligated to tester cDNA from SSHb.

Two separate subtractive hybridizations were done. The first hybridization (SSHa) was done with RWA infested near isogenic wheat lines, one susceptible to RWA infestation ('Tugela'- driver cDNA) and the other resistant to RWA infestation ('Tugela DN'-tester cDNA). This was followed by a second, independent hybridization (SSHb). RWA infested 'Tugela DN' was used for the tester cDNA and uninfested 'Tugela DN' was used for driver cDNA. Isolated total RNA had equal starting concentrations of 11 µg for both SSHa and SSHb.

Both SSHa and SSHb showed ligation of the adaptors to the driver cDNA (Fig. 3.3). The difference in the ligation reactions were noted and might have been due to insufficient ligation in SSHb. However, the decision was made to continue with the subtraction. The first hybridization was done, by adding driver cDNA to each sample separately. Driver cDNA derived from RWA infested 'Tugela' was added to the two tester samples bound with adapter 1 and adapter 2, respectively in SSHa, while driver cDNA derived from uninfested 'Tugela DN' was added to the two tester samples of SSHb. The second

hybridization was done by combining the two tester samples of each SSH and adding additional driver cDNA. Primary and secondary PCRs were done to amplify any novel sequences originating from the tester cDNA (Fig 3.4).

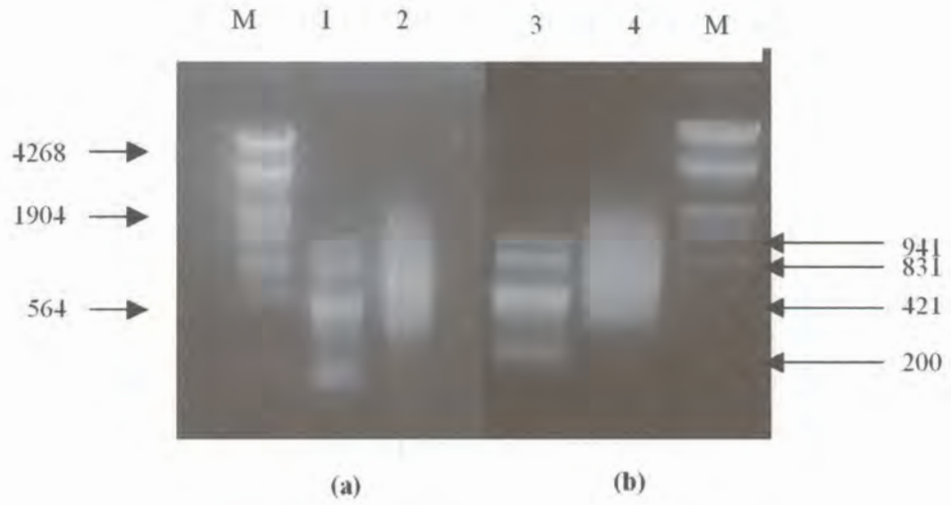


Figure 3.4. Results of both hybridizations after primary and secondary PCR. **(a)**: Three regions were obtained after primary and secondary PCRs were done on the SSHa reaction (infested 'Tugela DN' as tester and infested 'Tugela' as driver). Lane M = Marker III (λ DNA digestd with *Eco*R1 and *Hind*III), lane 1 = results of PCR done on subtracted reaction, lane 2 = results of PCR done on unsubtracted reaction used as control. **(b)**: Four regions were obtained after primary and secondary PCRs were done on the SSHb reaction (infested 'Tugela DN' as tester and uninfested 'Tugela DN' as driver). Lane 3 = results of PCR done on subtracted reaction, lane 4 = results of PCR done on unsubtracted reaction used as control.

Seven regions of amplification were obtained from both hybridizations (SSHa & SSHb) in total (Fig. 3.4). The regions obtained from SSHa were sized 941 bp (e.g. clone ABO 00010), 831 bp (e.g. clone ABO 00011) and 421 bp (e.g. clone ABO 00012) and the regions from SSHb were 941 bp (e.g. clone ABO 00013), 831 bp (e.g. clone ABO 00014), 421 bp (e.g. clone ABO 00026) and 200 bp (e.g. clone ABO 00027) in size. All the regions were isolated from the gel, cleaned and cloned to the pGEM-T Easy Vector cloning system. Ten clones of each region were sequenced (Table 3.2).

Table 3.2. A typical example of the results obtained after sequences resulting from SSHa and SSHb were submitted to GenBank. All sequences gave an "unknown" annotation.

BLASTX annotation	e-value	NCBI accession no [a]	Seq. accession no [b]
Unknown protein *(Mesorhizobium loti)*	5e-05	AP003017	BU808657
Integrase/recombinase *(Brucell melitensis)*	0.001	AE009541	BU808658
Integrase-like protein (bacteriophage H19J)	0.61	AF236875	BU808656
Putative protein *(Arabidopsis thaliana)*	7e-05	NP191007.1	BU808660
3-methylcrotonyl CoA carboxylase ciotin-containing subunit *(Oryza sativa)*	4e-06	AAL65397.1	BU808659
Copia-like polyprotein *(Arabidopsis thaliana)*	0.49	CAC7623.1	CA407985
Hypothetical protein *(Ralstonia metallidurans)*	0.49	ZP00026480.1	CA407984
Unknown (*Arabidopsis thaliana)*	7e-05	NP191007.1	BU808659
Photosystem I P700 apoprotein A1 *(Anthoceros punctatus)*	0.11	BAA83440.1	CA407985

a. Accession number obtained from GenBank identifying the sequence to which the submitted query showed homology to.

b. Accession number allocated to sequence when it was submitted to dbEST/GenBank. (Van Niekerk C. & Botha A.M. *Isolation and Characterization of cDNA Sequences from Russian wheat aphid induced 'Tugela DN' (Dn1) libraries*; www.ncbi.nlm.nih.gov/BLAST).

The sequences were submitted to GenBank and the nucleotide sequences, as well as their translations to amino acids were analyzed (Altschul *et al.*, 1997). Any sequences with an e-value higher than 10^{-5} were considered to have unknown functions. All results obtained from GenBank (Altschul *et al.*, 1997) showed no homology to any proteins with known function on either nucleotide or amino acid level (Table 3.2). The e-values showed no significant homology to any of the published proteins in the database.

Five sequences, ABO 00010, ABO 00011, ABO 00013, ABO 00014 & ABO 00027 were selected and labeled as probes for Northern and Southern blots (Table 3.3). These

samples were randomly selected since no sequences showed homology to any published sequences in GenBank.

Table 3.3. Designated probes for the DNA and RNA blots selected from sequences obtained from SSHa (tester cDNA = infested 'Tugela DN' and driver cDNA = infested 'Tugela') and SSHb (tester cDNA = infested 'Tugela DN' and driver cDNA = uninfested 'Tugela DN').

BLASTX annotation	e-value	NCBI Accession no.	Seq. accession no [a]	SSH [b]	Probe no. [c]
Unknown protein *(Mesorhizobium loti)*	5e-05	AP003017	BU808657	SSHa	ABO 00010
Integrase/recombinase *(Brucell melitensis)*	0.001	AE009541	BU808658	SSHa	ABO 00011
Putative protein *(Arabidopsis thaliana)*	7e-05	NP191007.1	BU808660	SSHb	ABO 00013
Hypothetical protein *(Ralstonia metallidurans)*	0.49	ZP00026480.1	CA407984	SSHb	ABO 00027
Unknown *(Arabidopsis thaliana)*	7e-05	NP191007.1	BU808659	SSHb	ABO 00014

a. Accession number allocated to sequences when it was submitted to dbEST/GenBank. (Van Niekerk C. & Botha A.M. *Isolation and Characterisation of cDNA Sequences from Russian wheat aphid induced 'Tugela DN' (DnI) libraries*; www.ncbi.nlm.nih.gov/BLAST).

b. The subtraction from which the sequence was obtained i.e SSHa or SSHb.

c. Clone name of submitted sequence.

The Southern blots were done to determine if the selected sequences (probes) originated from wheat (i.e. to exclude contamination) and from which genome they originally derived from. The Northern blots were done to show any up- or down-regulation of the sequences (probes) in different tissue from infested and uninfested wheat. The sequences were labeled with fluorescence. The Northern and Southern hybridizations were done and the membranes were exposed to x-ray film (Fig. 3.5).

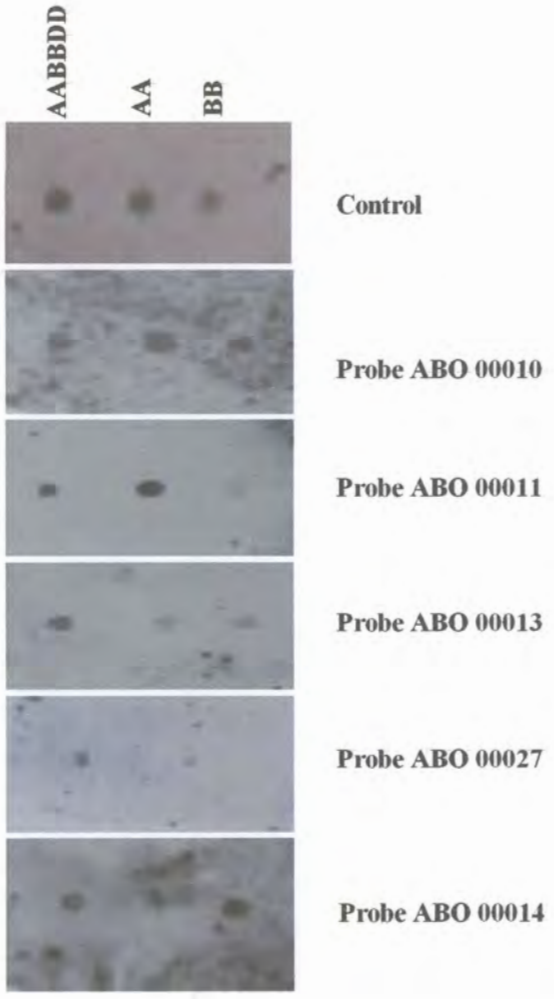


Figure 3.5. DNA dot blot analysis of cross hybridization of the probes to different wheat genomes. Control = ubiquitin probe; Probe ABO 00010 & ABO 00011 = sequences identified during SSHa; Probe ABO 00013, ABO 00027 & ABO 00014 = sequences identified during SSHb. Sample AABBDD = genomic DNA of 'Tugela DN', sample AA = genomic DNA of *Triticum urartu* and sample BB = genomic DNA of *Aegilops speltoides* .

From the results obtained it can be determined that all the probes are part of the wheat genome and not foreign DNA that contaminated the tester samples (Fig. 3.5). Probe ABO 00010 and probe ABO 00011 showed higher cross hybridization in the AA genome (*Triticum urartu*), probe ABO 00013 and probe ABO 00027 higher cross hybridization in

the AABBDD genome ('Tugela DN') and probe ABO 00014 higher cross hybridization in the BB genome (*Aegilops speltoides*).

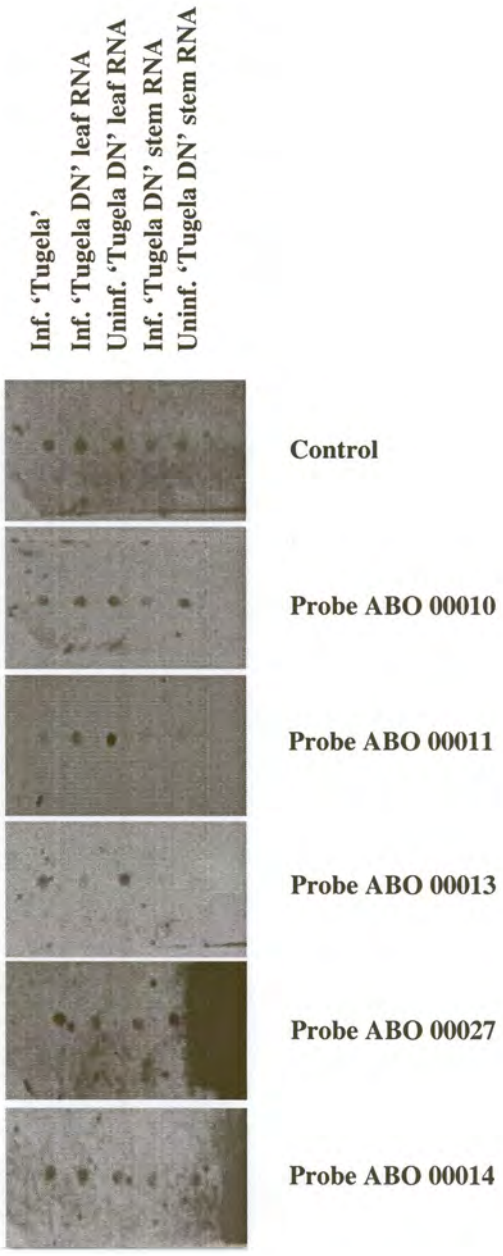


Figure 3.6. RNA dot blot analysis of cross hybridization of the probes to different wheat tissue. Control = ubiquitin probe, probe ABO 00010 & ABO 00011 = sequences identified during the first SSH subtraction, probe ABO 00013, ABO 00027 & ABO 00014 = sequences identified during the second SSH subtraction. Samples were obtained from the 'Tugela' (susceptible) and 'Tugela DN' (resistant) wheat lines. Leaf and stem tissue from RWA infested and uninfested plants were used for RNA extraction.

The results of the Northern blots showed up- and down-regulation that varied between the different probes (Fig. 3.6). The control reaction showed equal loading of all the samples. Probes ABO 00010 and ABO 00011 showed specific hybridization to the leaf RNA compared to the stem RNA, but no preference between RWA infested and uninfested material was observed. Probe ABO 00013 showed more hybridization to uninfested leaf RNA than to infested leaf RNA and almost no hybridization to infested or uninfested stem RNA. From the hybridizations of probes ABO 00027 and ABO 00014 no differences could be observed between the different tissue RNA samples or the RWA infested and uninfested samples.

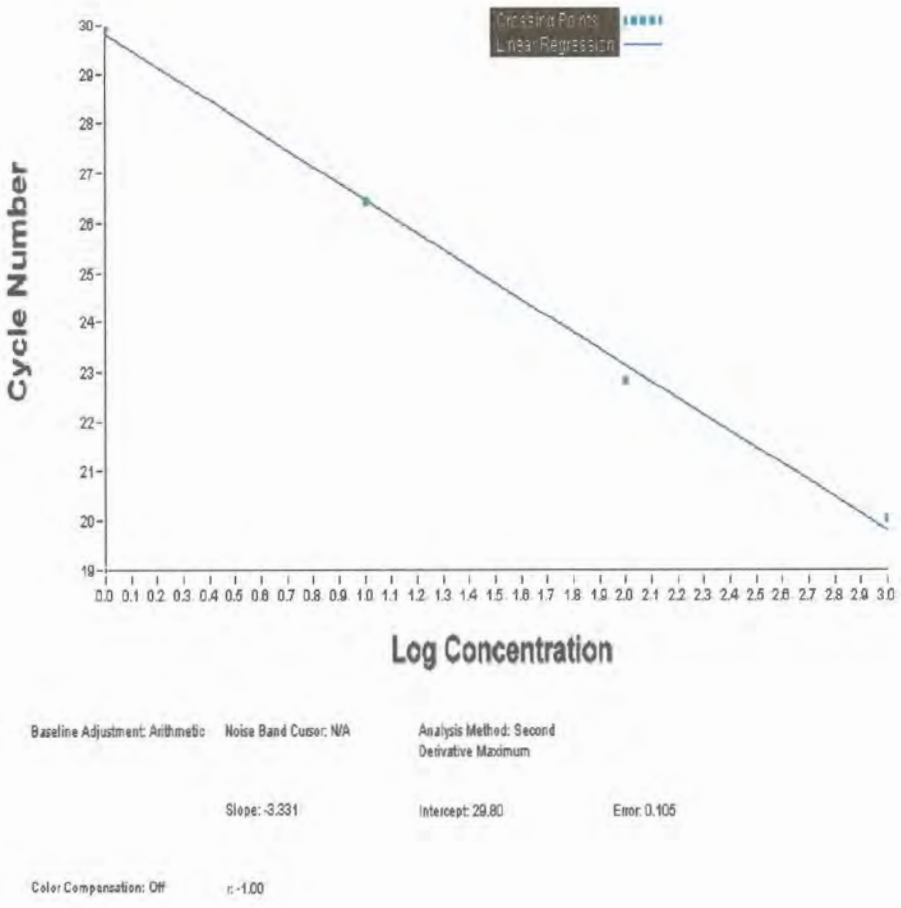


Figure 3.7. Standard curve generated during Real-Time PCR with primer ABO 00013. The r-value = -1 and the error value = 0.105.

Real-Time PCR was done to determine a relative quantification of the up-and down-regulation of three of the SSH obtained sequences, namely ABO 00013, ABO 00014 & ABO 00027. A standard curve was generated for each reaction. All reactions had an r-value of –1 and the error value was less than 0.2 (Fig.3.7). A melting curve was obtained after each Real-Time PCR (Fig. 3.8). This showed the products generated from the PCR as well as any primer dimer formation.

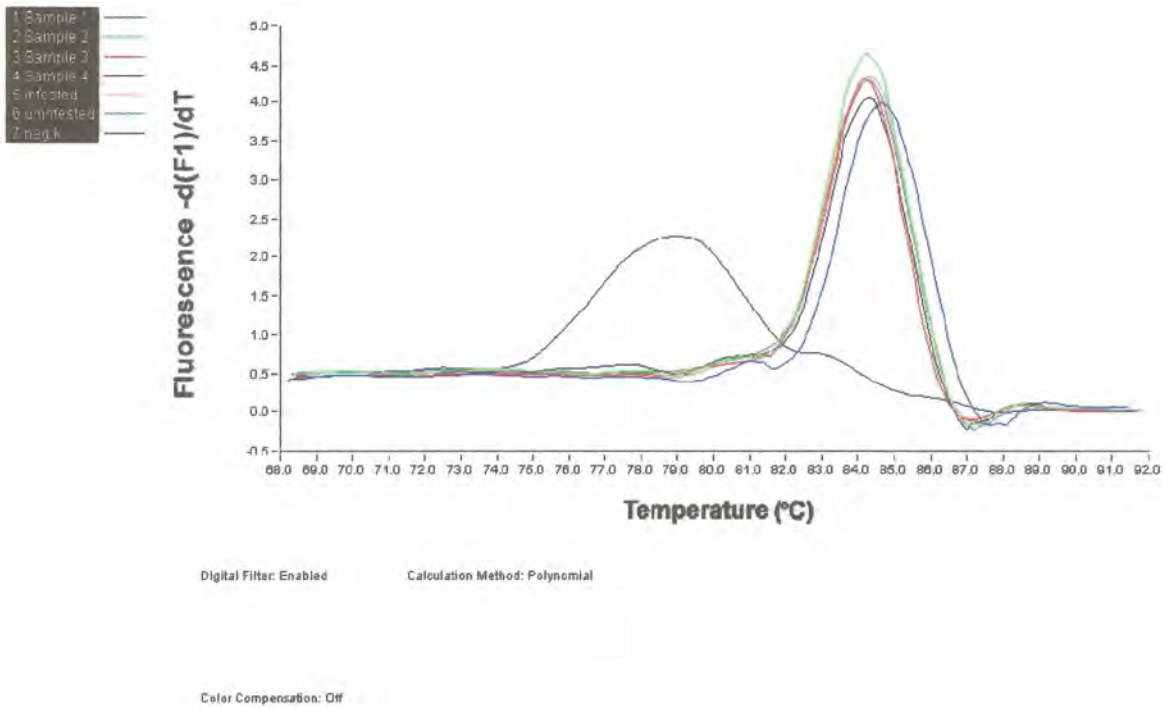


Figure 3.8. Melting curve generated during Real-Time PCR with primer ABO 00013. The smaller peak situated at 79 °C indicated the primer dimers. The larger peaks present at 85 °C represent the product.

Real-Time PCR was done on cDNA synthesized from RNA obtained from uninfested as well as infested 'Tugela DN' leaf and stem tissue. A time trail was also conducted by using leaf tissue infested for two and five days for RNA extraction and cDNA synthesis. The results were analyzed (Fig. 3.9).

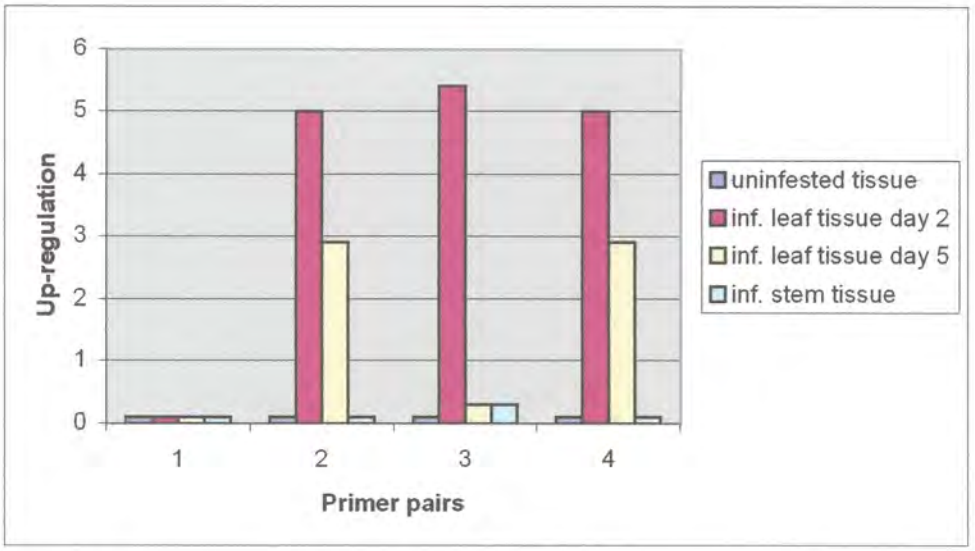


Figure 3.9. Histogram illustrating up-regulation of SSH sequences obtained through Real-Time PCR. Primer pair 1 = ubiquitin control; Primer pair 2 = primer ABO 00013 (SSHb); Primer pair 3 = primer ABO 00027 (SSHb) ; Primer pair 4 = primer ABO 00014 (SSHb). Leaf and stem tissue from RWA infested material was used. Leaf tissue was isolated on day 2 and day 5 of infested wheat.

From the results obtained, no up- or down-regulation was observed in the cDNA synthesized from RWA infested stem tissue with all the primer pairs, while infested leaf tissue on day 2 showed up-regulation of 5 times in primer pairs ABO 00013 and ABO 00014 and 5.4 times up-regulation in primer pair ABO 00027. Up-regulation of 2.9 times was still observed in tissue isolated from RWA infested leaf material on day 5 in primer pairs ABO 00013 and ABO 00014, but primer pair ABO 00027 showed only 0.3 times up-regulation.

5. Discussion

The study of an organism with as complex a genome as wheat is very complicated (Devos & Gale, 1993). By utilising SSH, novel sequences can be identified which might be differentially regulated (Birch *et al.*, 2000; Desai *et al.*, 2000). In SSH, two cDNA populations are compared with one another. The tester population might contain a novel sequences while the driver population might not and thus drives the reaction by

hybridizing to all homologous sequences between the two populations. This brings about that the novel sequences will be more likely to amplify during PCR.

Two approaches were taken in this study. The first (SSHa) used cDNA obtained from RWA infested leaf tissue from the resistant wheat line 'Tugela DN' as tester and compared it with cDNA synthesized from RWA infested leaf tissue from the susceptible wheat line 'Tugela' as driver. Thus, two near isogenic wheat lines were compared; the distinction being one wheat line is resistant to RWA infestation and the other not. The second approach (SSHb) compared tester cDNA from RWA infested leaf tissue from 'Tugela DN' to cDNA derived from uninfested leaf tissue from 'Tugela DN'. This approach compared cDNA from the same resistant wheat line, but the tester populations was subjected to RWA infestation and the driver population not. These subtractions were done independently of one another.

After the suppression subtractions were done, seven DNA regions were isolated after PCR amplification and sequenced. Randomly selected sequences were submitted to GenBank for identification. No homology was found with any of the submitted sequences on either nucleotide or amino acid level. Many SSH studies show novel genes being isolated. In a study done on breast cancer cell lines, 10 clones were identified, of which two were novel sequences (Yang *et al.*, 1999). Another study on breast cancer showed 18 clones from 29 that were novel sequences (Kuang *et al.*, 1998). When two *E. coli* lines were compared, 22 fragments were isolated of which seven did not show homology to any known genes (Janke *et al.*, 2001)

Southern hybridizations showed that the sequences identified during SSHa and SSHb were indeed part of the wheat genome and not a contamination (Wang & Feuerstein, 2000; Janke *et al.*, 2001). Probes ABO 00010 and ABO 00011 obtained from SSHa showed higher hybridization to the AA genome (*Triticum urartu*; Feldman, 1976) and probes ABO 00013 and ABO 00027 obtained from SSHb had higher cross hybridization to the AABBDD genome ('Tugela DN'). Probe ABO 00014 resultant from SSHb hybridized in higher levels to the BB genome (*Aegilops speltoides*; Feldman, 1976). The

DD genome contributes the *Dn1* resistance gene, which is a single dominant gene (Du Toit, 1989b; Nkongolo *et al.*, 1991). There was a high amount of hybridization between the AABBDD genome and both probe ABO 00013 and probe ABO 00027, thus implying hybridization potential to the DD genome.

The northern hybridizations were done to determine expression of the identified sequences (Wang & Feuerstein, 2000; Kuang *et al.*, 1998; Konietzko & Kuhl, 1998). Probes ABO 00010 and ABO 00011 (SSHa) hybridized to leaf RNA and stem RNA. No preference to infested or uninfested RNA could be determined. Probe ABO 00013 (SSHb) showed a higher hybridization to uninfested leaf RNA and almost no hybridization to stem RNA. Probes ABO 00027 and ABO 00014 (SSHb) hybridized equally to all samples. This indicates that the isolated SSH products are among the transcripts, although a simple dot blot cannot exclude the possibility of hybridisation with other homologous DNA sequences. Since the dotblot data is not specific enough, Real-Time PCR was done on the same samples to confirm results.

Probes ABO 00013, ABO 00027 and ABO 00014 were submitted to Real-Time PCR for relative quantification and comparison with hybridization results for the determination of up- or down-regulation in wheat on day 2 and day 5 of RWA infestation (Schnerr *et al.*, 2001). The time trail was done because previous research has shown that inter- and intracellular activities in the wheat plants are induced to much higher levels in resistant plants after 48 hours of RWA infestation (Van der Westhuizen *et al.*, 1998). The results indicated that cDNA isolated from stem tissue showed no up- or down-regulation with all three probes (ABO 00013, ABO 00027 and ABO 00014), while cDNA synthesized from infested leaf tissue showed up-regulation of 5 times on day 2 (probes ABO 00013 and ABO 00014) and 5.4 times with probe ABO 00027. Day 5 also showed up-regulation, but less than on day two, with probes ABO 00013 and ABO 00014 2.8 times up-regulated and probe ABO 00027 0.3 times up-regulated. The results obtained from the Northern blots and Real-Time PCR for ABO 00013 are contradictory. The Northerns show no hybridization to infested leaf tissue, while the Real-Time PCR shows up-regulation of 5 times on day 2. This anomaly can be due to less sensitive hybridization in

the Northern blots. Thus, doing Real-Time PCR as well is a good way of determining up- or down-regulation, rather than just relying on Northern blots, due to low transcript levels preventing detection in a dot blot system

The objective of this study was to isolate novel sequences from the RWA resistant wheat line 'Tugela DN' that are involved in the response to RWA infestation. By utilising SSH novel sequences could be isolated and identified, of which three showed up-regulation on day 2 and day 5 after RWA infestation. This study showed that SSH is a useful method to isolate novel DNA sequences. However, future experiments have to prove if these sequences are unique to the tester DNA. This can only be determined by hybridizing these sequences back to the tester and driver. Real-Time PCR evidence indicates up-and down-regulation of some of these sequences, but this cannot be taken as conclusive evidence. Although mRNA isolation has been applied to isolate gene sequences, future studies have to show if these novel sequences indeed represent gene sequences. To determine that, RACE (rapid amplification of 5' and 3' cDNA ends) should be used to identify full-length cDNA sequences. The obtained sequences were all submitted to GenBank, forming a database of sequences involved in response to RWA infestation.

6. References

Altschul S.F., Madden T.L., Schaffer A.A., Zhang J., Zhang Z., Miller W. & Lipman D.J. 1997. *Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.* Nucleic Acids Research 25: 3389-3402. www.ncbi.nlm.nih.gov/BLAST/

Birch P.R.J., Toth I.K., Avrova A.O., Dellagi A., Lyon G.D. & Pillay D. 2000. *Recent advances in the identification of differentially expressed genes.* South African Journal of Science 96: 83-85.

Caturla M., Chaparro C., Schroeyers K. & Holsters M. 2002. *Suppression subtractive hybridization to enrich low-abundance and submergence-enhanced transcripts of adventitious root primordia of Sesbania rostrata.* Plant Science 162: 915-921.

Chomczynski P. & Sacchi N. 1987. *Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction.* Anal Biochem 162: 156-159.

Clontech Laboratories. 1999. *Clontech PCR-select cDNA subraction kit user manual.* Protocol # PT1117-1, version # PR99446.

Desai S., Hill J., Trelogan S., Diachenko L. & Siebert P.D. 2000. *Identification of differentially expressed genes by suppression subtractive hybridisation.* In Functional Genomics (Edited by S.P. Hunt & F.J. Livesey) Oxford University Press pp 81-111.

Devos K. & Galle M. 1993. *The genetic maps in wheat and their potential in plant breeding.* Outlook in Agriculture 22: 93-99.

Diachenko L., Lau Y-F. C., Campbell A.P., Chenchik A., Moqadam F., Huang B., Lukyanov S., Lukyanov K., Gurskaya N., Sverdlov E.D & Shebert P.D. 1996. *Suppression subtractive hybridization: A method for generating differentially regulated or tissue-specific cDNA probes and libraries.* Proclamations of the National Acadamy of Science, USA 93: 6025-6030.

Du Toit F. 1989a. *Inheritance of resistance in two Triticum aestivum lines to Russian wheat aphid (Homoptera: Aphididae).* Journal of Economic Entomology 82 (4): 1251-1253.

Du Toit F. 1989b. *Components of resistance in 3 bread lines to Russian wheat aphid (Homoptera: Aphididae) in South Africa.* Journal of Economic Entomology 82 (6): 1779 – 1781.

Feldman M. 1976. *Wheats, Triticum spp (Gramineae - Triticinae).* In Evolution of Crop Plants (Ed. N.W. Simmonds). Longman. London.

Gehrig H.H., Winter K., Cushman J., Borland A. & Taybi T. 2000. *An improved RNA isolation method for succulent plant species rich in polyphenols and polysaccharids.* Plant Molecular Biology Reporter 18: 369-376.

Gussow D. & Clackson T. 1989. *Direct clone characterization from plaques and colonies by the Polymerase chain reaction.* Nucleic Acid Research 17: 4000.

Janke B., Dobrindt U., Hacker J. & Blum-Oehler G. 2001. *A subtractive hybridization analysis of genomic differences between the uropathogenic E. coli strain 536 and the E. coli K-12 strain MG 1655.* Federation of European Microbiological Societies Microbioloby Letter 199: 61-66.

Konietzko U & Kuhl D. 1998. *A subtractive hybridization method for the enrichment of moderately induced sequences.* Nucleic Acid Research 26(5): 1359-1361.

Kruger W.M., Pritsch C., Chao S. & Muehlbauer G.J. 2002. *Functional and comparative analysis of expressed genes from wheat spikes infected with Fusarium graminearum.* MPMI 15 (5): 445-455.

Kuang W.W., Thompson D.A., Hoch R.V. & Weigel R.J. 1998. *Differentail screening and suppression subtractive hybridization identified genes differentially expressed in an estrogen receptor-positive breast carcinoma cell line.* Nucleic Acid Research 26(4): 1116-1123.

Liang P. & Pardee A.B. 1992. *Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction.* Science 257: 967-971

Lisitsyn N., Lisitsyn L. & Wigler M. 1993. *Cloning the differences between two complex genomes.* Science 259: 946-951.

Money T., Reader S., Li J.Q., Dunford B.P & Moore G. 1996. *AFLP-based mRNA fingerprinting.* Nucleic Acid Research 24: 2119-2124.

Nkongolo K.K., Quick J.S., Limin A.E. & Fowler D.B. 1991. *Sources and inheritance of resistance to Russian wheat aphid in Triticum species amphiploids and Triticum tauschii.* Canadian Journal of Plant Science 71: 703 - 708.

Raeder U. & Broda P. 1985. *Rapid preparation of DNA from filamentous fungi.* Letter in Applied Microbiology 1: 17-20.

Sambrook J., Fritsch E.F. & Maniatis T. 1989. *Molecular Cloning: A Laboratory Manual.* Second Edition. Cold Spring Harbour Laboratory, New York.

Schnerr H., Niessen L. & Vogel R.F. 2001. *Real time detection of the tri5 gene in Fusarium species by LightCycler-PCR using SYBR Green I for continuous fluorescence monitoring.* International Journal of Food Microbiology 71: 53-61.

Van der Westhuizen A.J., Qian X-M & Botha A-M. 1998. *β-1,3-Glucanases in wheat and resistance to the Russian wheat aphid.* Physiologia Plantarum 103: 125-131.

Wang X. & Feuerstein G.Z. 2000. *Suppression subtractive hybridization: application in the discovery of novel pharmacological targets.* Pharmacogenomics 1 (1): 101-108.

Yang G.P., Ross D.T., Kuang W.W., Brown P.O. & Weigel R.J. 1999. *Combining SSH and cDNA microarrays for rapid identification of differentially expressed genes.* Nucleic Acid Research 27(6): 1517-1523.

CHAPTER 4

APPLYING MICROARRAY TECHNOLOGY TO DETECT EXPRESSED SEQUENCES IN WHEAT AFTER RUSSIAN WHEAT APHID INFESTATION.

Chapter 4:

Applying microarray technology to detect expressed sequences in wheat after Russian wheat aphid infestation.

1. Abstract

Diuraphis noxia (Russian wheat aphid, RWA) is a major pest on wheat in South Africa and many other wheat-growing countries. Many R genes from various plant species have conserved amino acid domains, particularly the nucleotide binding sites (NBS) and leucine repeat regions (LRR), which is consistent with their putative roles in signal transduction and protein-protein interactions. Previous studies on the intercellular washing fluid (IWF) of wheat cultivars resistant ('Tugela DN') and susceptible ('Tugela') to RWA showed alteration of some protein complexes within the first 12h after RWA infestation in the resistant cultivar, but not in the susceptible near isogenic line. Two responses, an initial hypersensitive response (HR) that decreases after approximately 24h, which is followed by systemic acquired resistance (SAR) that prevails in the tissue for an extended period of time, were observed. Two hundred and fifty-six wheat sequences were obtained using degenerate primer sets designed from the consensus NBS motif from other genome studies (e.g. *Arabidopsis* and rice), subtraction suppression hybridization (SSH) and cDNA libraries. Selected wheat cDNA clones were spotted onto microarrayer slides. Purified mRNA from infested material, containing the RWA resistance gene *Dn1*, was isolated 0, 2, 5 and 8 days after infestation, post-labelled with Cy3- or Cy5-fluorescent dyes and hybridized to the arrays. Statistical analysis of the expression data revealed the regulation of 5% of all the spotted gene fragments at a threshold log-2 expression ratio of 1.5 and $P \leq 0.05$. Wheat homologs to RGA-2 are regulated in response to RWA feeding. The expression levels of a subset of clones were verified by Real-time PCR and Northern blot analysis.

2. Introduction

Microarray is an ideal method to generate data in a systematic and comprehensive way (Brown & Botstein, 1999). Microarrays were first invented in Stanford, where they used *Arabidopsis* as the model organism (Schena *et al.*, 1995).

Microarray technology is based on the complementary binding of single-stranded nucleic acid sequences. Single-stranded DNA fragments are spotted onto a glass slide, each representing a single gene (Brazma *et al.*, 2000). RNA is used as a probe – a sample from the control is labeled with a fluorescent dye (i.e. green) and a sample from the experiment is labeled with another fluorescent dye (i.e. red). Both are hybridized to the slide simultaneously. Complementary sequences will hybridize. The slide is scanned with a laser scanner that reads the fluorescence being emitted. A green spot will indicate binding to the experimental sample and a red spot will indicate binding to the control sample. Equal binding will show a yellow spot and no binding will be indicated by a black spot. In this way the relative expression levels of genes can be determined (Brazma *et al.*, 2000).

Planning a microarray experiment correctly is crucial to the success of the venture. Various experimental designs have been described, for example incomplete block design, robust design and classic block design to name a few (Kerr & Churchill, 2001). Extensive studies have been done to establish which design is the most effective. An experiment should be designed according to the question of the study, in other words, what do you want to prove with your study.

The amount of data generated with a microarray experiment is enormous. Adequate software is essential to process the data and to analyze it sensibly. A large number of programs are available and each method used could have a large impact on the interpretation of the results. Thus, a basic knowledge of these analytic tools is essential before planning, executing and analyzing a microarray experiment (Quackenbush, 2001).

Numerous studies have already been done using microarrays, mostly in the medical field. A study was done on genes involved in multiple sclerosis (MS) (Whitney *et al.*, 1999). The expression patterns of over 5 000 genes from MS tissue was monitored and 62 differentially expressed genes were identified, some not previously connected

to MS development. Another study was done on plant defense responses in *Arabidopsis* (Schenk *et al.*, 2000). Over 2 000 genes were examined to determine their expression patterns after inoculation with the fungal pathogen *Alternaria brassicicola*. Of these genes, 705 showed up- or down-regulation, some with a known defense-related function and others with an unknown function. A plant study, using *Arabidopsis*, was done to determine the influence of vitamin C deficiency on plants (Pastori *et al.*, 2003). They discovered 171 genes that were differentially expressed in plants deficient in vitamin C when compared to the wild type. These genes included defense genes. They concluded that vitamin C plays a role in plant defense, survival and growth. As can be seen from these examples, microarrays can be used to study various subjects successfully.

In this chapter a comparative analysis was done, using infested material containing the RWA resistance gene *Dn1* and utilizing degenerate primer sets designed from the consensus NBS motif from other genome studies (e.g. *Arabidopsis* and rice). From the previous study (Chapter 2), numerous ESTs / clones were isolated with no discernable function when compared to existing published data in GenBank. Wheat homologs to NBS-LRR putative resistance genes were also isolated. There were two objectives in this part of the work, namely to determine the up- and down-regulation of unknown ESTs / clones during RWA infestation, and secondly to established the strength of using a microarray approach for such a study. To follow the expression profiles of these gene sequences, the microarray was hybridized against cDNA synthesized from leaves of the RWA resistant cultivar 'Tugela DN' pre- (day 0) and post-infestation (days 2, 5 and 8), in an effort to identify possible gene sequences involved in the RWA defense response.

3. Material and Methods

3.1 Experimental outline

An outline of the major steps for the microarray preparation is given (Fig. 4.1).

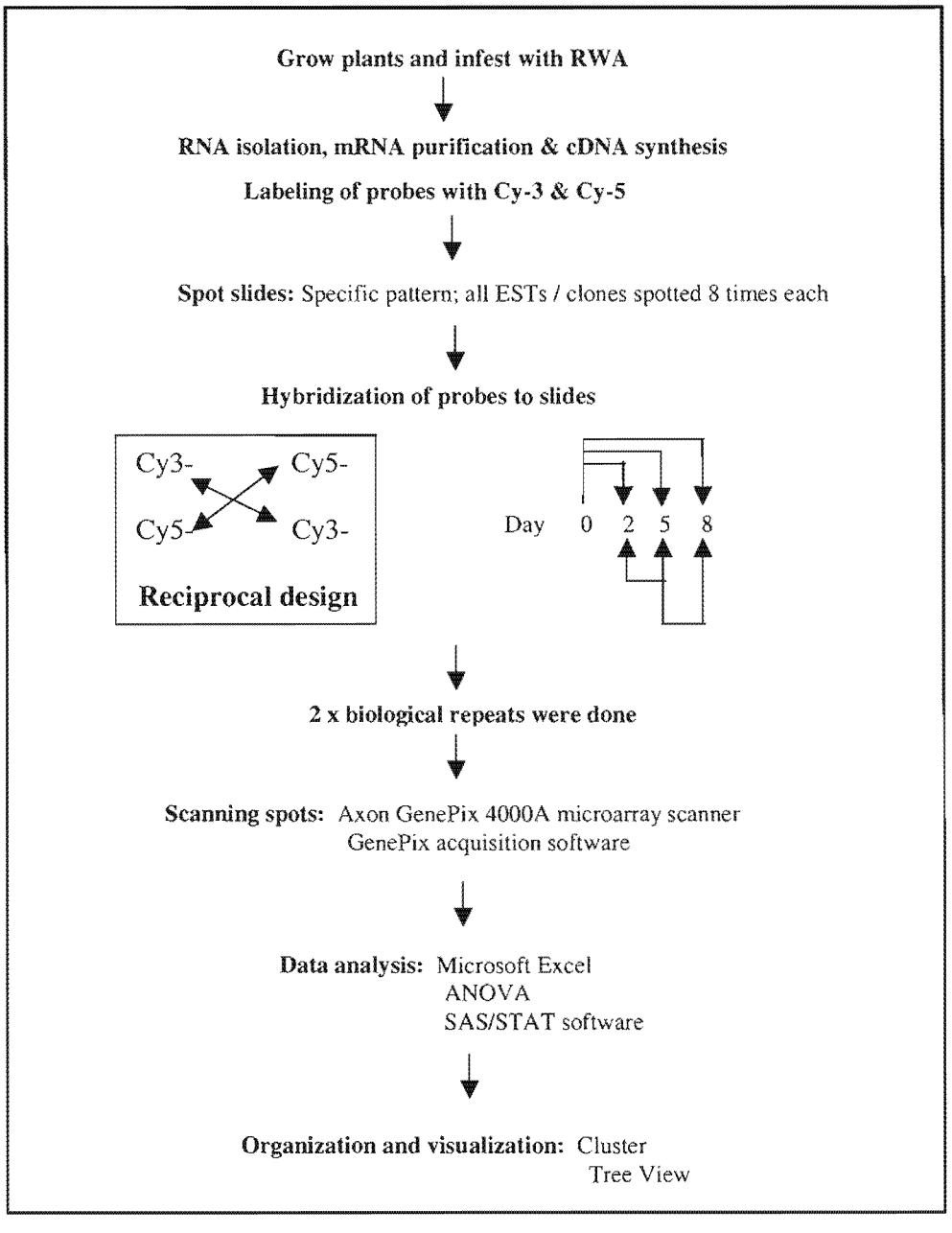


Figure 4.1 Experimental outline of microarray experiment.

3.2 Plant material

Russian wheat aphid resistant cultivar 'Tugela DN' (Tugela*5/SA1684, *Dn1*) was grown in pots under greenhouse conditions with prevailing day and night cycles at the Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria. The temperature was maintained at about 24°C, and the plants were watered daily. Half of the 20 wheat seedlings were infested with RWA (10 aphids per plant) at the three to four leaf growth stage (Botha *et al.*, 1998). The second leaves from uninfested and infested plants were removed on day 0, the third leaves on day two, fourth leaves on day five and sixth leaves on day eight post-infestation (PI), for analysis. The aphids were removed from the infested leaves prior to RNA isolation. Only leaves showing signs of aphid feeding i.e. yellow spots were used.

3.3 Preparation of ESTs / clones for spotting on the Microarray

RNA was extracted from 10 different plants and pooled. Day 0,2,5 and 8 were kept separately and the extractions were repeated over biological material (i.e. once more at a later date). Total RNA isolation, purification of mRNA, cDNA synthesis, cDNA library construction and sequencing was performed as previously described (Chapter 2). After sequencing of randomly selected clones, sequence identities were annotated through BLAST searching and alignment to other published sequences in GenBank (Altschul *et al.*, 1997). Functions were assigned to the ESTs / clones based on the results (e value $< 10^{-5}$; Kruger *et al.*, 2002) returned from searches using the BLASTX algorithm.

3.4 Microarray preparation

Target cDNA (192 wheat ESTs / clones derived by using NBS primers (Chapter 2), 55 flax and 33 banana ESTs / clones (obtained from Prof. C.Cullis, USA) for spotting on the microarray was amplified using standard PCR procedures (40 cycles; annealing at 64°C; 2 ng plasmid template). PCR products were purified using Multiscreen purification plates as prescribed by the manufacturer (Millipore, USA). PCR products were quantified by electrophoresis on 0.8% agarose gels (w/v) and visualized by

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
A	1	33	65	97	129								17	49	81	113								
B	2	34	66	98	130								18	50	82	114								
C	3	35	67	99	131								19	51	83	115								
D	4	36	68	100	132								20	52	84	116								
E	5	37	69	101	133								21	53	85	117								
F	6	38	70	102	134								22	54	86	118								
G	7	39	71	103	135								23	55	87	119								
H	8	40	72	104	136								24	56	88	120								
I	9	41	73	105	137								25	57	89	121								
J	10	42	74	106	138								26	58	90	122								
K	11	43	75	107	139								27	59	91	123								
L	12	44	76	108	140								28	60	92	124								
M	13	45	77	109	141								29	61	93	125								
N	14	46	78	110	142								30	62	94	126								
O	15	47	79	111	143								31	63	95	127								
P	16	48	80	112	144								32	64	96	128								

Fig 4.2(a)

Fig. 4.2(b)

Fig. 4.2. Outlay of microarray slides. Fig. 4.2(a) shows the different ESTs / clones spotted onto each microarray slide. Fig. 4.2(b) shows the pattern in which the ESTs / clones were spotted onto the slide. All ESTs / clones were spotted 4 times underneath each other, and the whole pattern was repeated twice.

ethidium bromide staining. Microarray slides were printed by using a BioRobotics Generation II Arrayer, according to the manufacturer's instructions. Arrays were printed on aminosaline slides and each target DNA was spotted 8 times (Fig. 4.2). Thirty-six slides were made, all identical in pattern and spots. Negative controls on the array included blank spots, Lucidea Universal Scorecard (ratios 1:1, 1:3, 1:10, 3:1, 10:1), constitutively expressed and stress responsive genes (i.e. actin).

3.5 Fluorescent probe preparation

cDNA labelled with Cy3- and Cy5 was used as probes. Total RNA was isolated from wheat leaves on days 0, 2, 5 and 8 post-infestation by the RWA as previously described, and pooled. Poly-A^+ RNA was purified from total RNA using the Oligotex mRNA spin-column protocol (Oligotex™ Handbook 07/99, Qiagen). Purified mRNA (100 ng) was used for the preparation of Cy3- and Cy5-labelled cDNA for microarray hybridization using the Cyscribe Post-labelling kit according to the manufacturers instructions (Amersham Biosciences, UK). Unincorporated label and single stranded nucleotides were removed from the prepared labelled cDNA using the MinElute cleanup kit according to the manufacturers protocol (MinElute™ Handbook 04/2001, Qiagen).

3.6 Hybridization of probe to slide

One microarray slide was pre-hybridized per probe pair by adding 35 µl pre-hybridization solution (3.5 x SSC, 0.2% (w/v) SDS, 1% (w/v) BSA) for 20 min at 60°C using a humidified hybridization-cassette. Slides were washed in ddH_2O for 1 min and air-dried using nitrogen gas. For hybridization, 30 pmol of each probe per slide was dried in a 0.5 ml tube, resuspended in 35 µl hybridization solution (50% formamide, 25% hybridization buffer; 25% mQ) and denatured (98°C for 2 min). The slides were hybridized overnight for 12-18 h at 42°C. Slides were washed three times at 42°C for 4 min (once in 1x SSC / 0.2% (w/v) SDS, twice in 0.1 x SSC / 0.2% (w/v) SDS). This was followed by three washes at room temperature for 1 min each in 0.1 x SSC and dried with nitrogen gas. One slide was used per probe pair, with each EST spotted eight times. This enabled the direct comparisons of identical spots to determine the success of the particular hybridization. A 2x2 factorial design was used

to design the experiment (Wang and Speed, 2002). The experiment was repeated twice over biological material.

3.7 Scanning and data analysis

An Axon GenePix 4000A Microarray scanner and GenePix acquisition software (Axon Instruments, Inc., USA) were used according to the manufacturer's instructions regarding dye emission, to capture the data. Normalization between Cy3 and Cy5 fluorescent dye emission intensities was achieved by adjusting the level of the photomultiplier gains ('global normalization'). After scanning and capturing of data using the GenePix 3.0 software, the raw data was imported into Microsoft Excel for further analysis. Background fluorescence values were automatically calculated by the GenePix program and subtracted before further calculations were performed. Genes of interest were identified by computational analysis using ANOVA as proposed by Dudoit and coworkers (2001), the mixed model approach (Wolfinger *et al.*, 2001; Chu *et al.*, 2002) and SAS/STAT software version 8 (SAS Institute Inc., 1999). Genes were also organized and visualized by Cluster and Tree View (Eisen *et al.*, 1998). Duplicate spots were compared with one another to determine spot-to-spot variation. The data obtained from the RWA infested plant material was compared to the data obtained from uninfested plant material, and to each other.

3.8 Northern blots analysis and Real-Time PCR

Northern blot analysis was performed using total RNA extracted from uninfested (day 0) and infested (day 2) wheat leaves and stems, as well as leaves infested at day 0, 2, 5 and 8 post-infestation with the RWA. The RNA was pooled, 200 ng RNA was transferred onto a nylon membrane (Roche Diagnostic Corporation, Germany) and the RNA was UV-cross linked to the membranes (Sambrook *et al.*, 1989).

Probe labeling of 50 ng fragments each was done using the Gene Images Random Prime Labeling module (Amersham Pharmacia Biotechnology, USA) according to manufacturers instructions. Pre-hybridization of RNA was preformed at 60°C using the hybridization buffer (5x SSC (75 mM NaOAc and 0.75 M NaCl), 0.1% (w/v) SDS, 5% Denhardt's solution) and a 20x dilution of Liquid Block (Gene Images

Random Prime Labeling module, Amersham, USA) for 3.5 hours. 15 μl of each probe was heat-denatured for 5 min and added to the respective pre-hybridized membranes. Hybridization was done overnight at 65°C in a HB-1D Hybridizer (TECHNE, Cambridge, UK). Two stringency washes followed hybridization. The membranes were washed once in 1 x SSC (15 mM NaOAc and 0.15 M NaCl) and 0.1% (w/v) SDS, followed by 0.1 x SSC (1.5 mM NaOAc and 15 mM NaCl) and 0.1% (w/v) SDS. The blots were then incubated in buffer A (100 mM Tris-HCl and 300 mM NaCl) containing a 10x dilution of Liquid Block (Amersham Pharmacia Biotechnology, USA) for 1 hour at room temperature. The blots were then incubated in buffer A containing 0.5% (w/v) BSA and a 1:5000 dilution of anti-fluorescein-AP conjugate for 1 hour. This was followed by three wash steps of 10 min each in buffer A and 0.1% (v/v) Tween-20. CDP-Star (500 μl) detection reagent was added to the blots for 5 min, before exposure to HyperFlim (Amersham Pharmacia Biotechnology, USA) for 30 min, and developed (Sambrook *et al.*, 1989).

Quantitative PCR was performed using 70 ng first strand cDNA from selected total RNA as required, 10 pmol forward and reverse primers, 3 mM $MgCl_2$ and the LightCycler–FastStart DNA Master SYBR Green I Mix (FastStart Taq DNA polymerase, reaction buffer, dNTPs, SYBR Green I Dye and 10 mM $MgCl_2$) in a 20 μl reaction, as advised by the manufacturers (LightCycler–FastStart DNA Master SYBR Green Manual, Roche Applied Science, Germany). The cycling parameters consisted of (95°C for 10 min) x1 cycles; (95°C for 10s, primer specific annealing T_m for 5s, 72°C for 10s) x40 cycles, followed by the melting curve analysis (95°C for 0s, 65°C for 15s, 95°C for 0s) x1 cycle, and cooling (40°C for 30s) x1 cycle. A minimum of seven reactions was done for each fragment analyzed, which included a sample with uninfested wheat cDNA to set a standard value from which to determine the up- or down-regulation. Standard curves were generated using dilution series (1:1, 1:10, 1:100, 1:1000) and repeated. Results obtained were analyzed using LightCycler Software version 3.5.

4. Results

To monitor the expression of leaf tissue in post-RWA infested tissue, we used cDNA microarray slides containing 384 spots, including 192 selected expressed sequence tags (ESTs) previously isolated from RWA induced cDNA libraries (as described in Chapter 2). To this collection, we added 55 *Linum usitatissimum* (flax) and 33 banana clones (obtained from Prof C. Cullis, Ohio, USA). These consisted of 27 clones with no discernable function, and nine flax clones with significant homology to published sequences in GenBank. Amongst these, we included a flax homolog to APC/C ubiquitin-protein ligase with known function in cell cycle regulation, as well as five different clones with significant homology to the *Linum usitatissimum* LIS-1 insertion sequences. The latter has previously been shown to be induced in genotrophs by the environment (AF104351). The flax and banana clones were incorporated to serve as internal controls, since most of these clones had a known function. Controls (54 spots) were also incorporated, including genes known to be regulated under stress i.e. actin genes (eight spots), and the Lucidea Universal Scorecard (46 spots). Our focus was on a comparison between pre-and post-infestation events. The ESTs / clones were spotted onto aminosaline slides and hybridized against Cy3- and Cy5-labeled cDNA prepared from RWA infested leaf material pre-(day 0) and post-infestation (days 2, 5 and 8) in a time trial spanning from day 0 to day 8.

4.1 Statistical analysis of data

We analyzed the fluorescence data from the microarray slides using a general analysis of variance (ANOVA) as suggested by Dudoit *et al.* (2001). The analysis indicated that 27% ESTs / clones were down-regulated, and 28% up-regulated. This model is based on the normalization of log ratios, then permutation-based t-statistics for testing the significance of each gene, and *p*-values, which are suitably adjusted for multiplicity. We argued that the obtained significance data were statistically too many, and thus allowed for false positives. Thus, we subjected the fluorescence data to the statistically rigorous mixed model approach (Wolfinger *et al.*, 2001) that allows for the identification of false positives, as well as the selection of genes with significant expression (Figure 4.3, Table 4.1, data points indicated in red and green).

Using this approach, our obtained data revealed the regulation of 12% of all the spotted gene fragments at a threshold log-2 expression ratio of 1.5 and $P \leq 0.05$ (Fig. 4.3). All spots above the cut-off line of 1.5 were considered significantly up- or down-regulated. This amounted to 42 spots that showed significant homology to 120 sequences (Table 4.1).

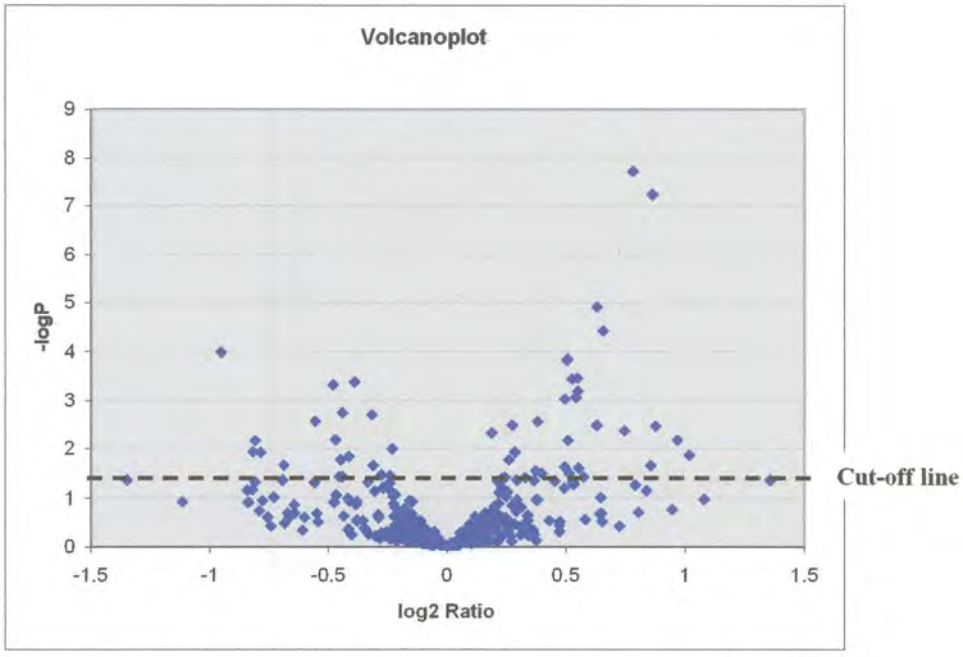


Figure 4.3. Volcano significance plot for a subset of wheat ESTs / clones, indicating up- and down-regulation. Plotted on the vertical axis is $-\log_{10}(p\text{-value})$ for contrast between treatments. Horizontal axis is $\log_2$ of the estimated fold change, suitably adjusted for other systematic and random effects in the experiment. All spots above the cut-off line were considered significantly up- or down-regulated.

4.2 Expression profiling of transcripts

The 120 ESTs / clones with significant changes in expression were divided into four broad categories (Table 4.1), namely (1) sequences involved in cell division, growth and organization, (2) chloroplast structure and function, (3) resistance gene analogs (defense related) and (4) unclassified or unknown. The data from ESTs / clones with no significant change in expression were discarded, and these were thus not included in the table. Six ESTs / clones from the first group changed significantly upon RWA infestation, and include wheat homologs to a *Beta nana* Ty-1-copia-like

retrotransposon, an *Oryza sativa* T-DNA intergration factor, *Hordeum vulgare* BARE-1 long terminal repeat, and a ribosomal RNA gene. Three ESTs / clones listed under the chloroplast structure and function group were significantly regulated, and include an unknown chloroplast sequence and a chloroplast gene for a chloroplast product, as well as a chloroplast ATP synthase. Two wheat homologs to RGA-2 (*T. monococcum* putative resistance protein) resulted in significant down- (day 2/5) and up-regulation (day 5/8) in response to RWA infestation. Seven of the ESTs / clones classified as unknown or without any discernable function resulted in a significant regulation due to RWA infestation. These include all types of regulation, e.g. mainly down-regulated, mainly up-regulated, only down-regulated, only up-regulated, as well as a combination of down-/up-regulation in response to RWA feeding.

Table 1. Genes differentially expressed in response to RWA feeding. The expression of mRNA extracted from leaves was determined through cDNA microarray hybridization. Relative fluorescence signals are shown for those genes whose ratios are significantly up- or down-regulated in response to RWA feeding, respectively. Values are highlighted in red (+) or green (-) if expression differs at a significance level of $P \leq 0.05$ (Mixed Model; Wolfinger *et al.*, 2001); values in black differ at a significance level of $P \leq 0.05$ (ANOVA, Dudoit *et al.*, 2001).

GenBank Accession#	Annotation [b]	Expression ratios (Fluorescence Units) Day 2/0	Day 5/0	Day 8/0
	Cell division, growth and organization			
	Xenopus laevis mRNA for KREMEN, complete cds (e-value = 9e-24)	-0.586	2.833	5.650
	Uncultured bacterium partial 16 rRNA gene(AF327894, e-value=4e-24)	-0.889	2.403	-0.531
	APC/C ubiquitin-protein ligase (cell cycle regulation)	-1.661	-0.351	1.812
	Beta nana Ty1-copia-like retrotransposon for putative reverse transcriptase (AJ489197.1; e-value = 3e-28)	-1.445	-0.332	2.753
	Beta nana Ty1-copia-like retrotransposon for putative reverse transcriptase-3 (AJ489202.1; e-value = 3e-13)	9.541	0.842	3.895
	Hordeum vulgare BARE-1 long terminal repeat (Z84562.1, e-value = 5e-09)	-1.652	-0.115	4.156
	Cicer arietinum Ty1-copia retrotransposon for putative reverse transcriptase-3 (AJ535884.1, e-value = 8e-43)	6.125	-8.058	-7.711
	Gene for Ribosomal DNA CON-2	-3.337	-0.871	5.853
	Gene for Ribosomal DNA CON-5	3.236	-0.341	1.040
	Atropa belladonna partial mRNA 3'URT (AJ309392.1, e-value = 1e-19)	1.559	-0.285	0.048
	Oryza sativa T-DNA integration factor-2 (U40841.1; e-value = 2e-23)	-2.506	1.372	2.382
	Cicer arietinum Ty1-copia retrotransposon for putative reverse transcriptase-2 (AB086192.1, e-value = 2e – 42)	5.155	-0.347	-10.463
	Zea mays serine hydroxymethyl-transferase mRNA (AF439728.1, e-value = 6e-52)	-15.730	2.286	35.611
	Atropa belladonna partial mRNA 3'URT (AJ309392.1, e-value = 1e-16)	5.547	-1.526	-10.674
	Linum usitatissimum LIS-1 insertion sequence in genotrophs induced by the environment (AF104351, e-value = 0)	-6.398	1.239	9.017

	Linum usitatissimum LIS-1 insertion sequence (AF104351, e-value = 0)	1.130	-0.171	0.988
	Linum usitatissimum LIS-1 insertion sequence (AJ131994.1, e-value = 1e-149)	1.593	1.681	-3.808
	T-DNA integration factor (U40814.1; e-value = 3e-23)	1.101	0.014	0.540
	Beta nana Ty1 copia-like retrotransposon (AJ489200.1, e-value = 4e-13)	-0.969	-2.655	0.558
	Oryza sativa T-DNA integration factor-3 (U40841.1; e-value = 2e-23)	-2.256	0.334	-1.191
	Linum usitatissimum LIS-1 insertion sequence-2 (AJ131994.1, e-value = 0)	2.025	0.511	-2.840
	Linum usitatissimum LIS-1 insertion sequence-3 (AJ131994.1, e-value = 1e-102)	1.215	0.862	2.350
	Oryza sativa T-DNA integration factor (U40814.1; e-value = 6e-23)	-1.293	0.979	6.887
	Hordeum vulgare BARE-1 long terminal repeat-2 (Z84562.1; e-value = 9e-25)	2.433	0.179	-2.141
	Beta nana Ty1-copia-like retrotransposon for putative reverse transcriptase-2; (AJ489197.1; e-value = 7e-26)	-15.961	0.865	12.369
	Gene for Ribosomal DNA CON-6	3.848	1.481	2.944
	Gene for Ribosomal DNA CON-3	2.222	0.703	-1.957
	Genes for Ribosomal DNA CON-1	-14.251	-2.334	28.515
	Genes for Ribosomal DNA CON-4	2.664	0.432	0.838
	Hordeum vulgare BARE-1 long terminal repeat-3 (Z84562.1; 3-value= 1e-19)	-0.268	-1.124	0.243
	Cicer arietinum Ty1-copia retrotransposon for putative reverse transcriptase (AJ535884.1, e-value = 9e-49)	-1.133	-0.997	-5.207
	Zea mays Serine hydroxymethyl transferase mRNA (AF439728.1, e-value = 6e-52)	-1.328	0.467	1.627
	Chloroplast structure / function			
	Banana unknown chloroplast sequence(e-value = 1e-76) CC18	2.361	-0.107	-2.887
	Dendrobium chrysotoxum trnK, matK pseudogene, chloroplast genes (AF448862.1, e-value = 2e-14)	-2.872	-1.468	14.494
CB412206	Wheat chloroplast – ATP synthase-7 (M16843.1, e-value = 0) [a]	1.165	0.024	-1.834
	Banana chloroplast gene for chloroplast product (e=value =1e-76) CC17	2.007	0.905	-0.120
CB412223	Triticum aestivum chloroplast (AB042240.3, e-value = 0)	-1.055	-0.269	3.342
	Anthoceros punctatus chloroplast gene for photosystem I P700 apoprotein A1 (AB013664.1, e-value = 4e-16)	11.163	-1.840	14.849
CB412239	Wheat chloroplast – ATP synthase (M16843.1, e-value = 0) [a]	1.796	0.209	4.024
CB412251	Aegilops crassa chloroplast genes –ATP synthase (AEGATPS2, e-value = 0) [a]	224.264	-63.190	463.415
	Maize chloroplast DNA for 4.5S rRNA (X01365.1, e-value = 8e-08)	2.153	1.134	-2.983
CB412240	Triticum aestivum chloroplast gene (AB042240.3, e-value = 0)	-0.933	3.773	-3.479
CB412217	Wheat chloroplast – ATP synthase-2 (M16843.1, e-value = 0[a]	-1.209	-1.085	-0.653
CB412200	Wheat chloroplast – ATP synthase-6 (M16843.1, e-value = 0) [a]	0.719	-1.015	-1.423
CB412218	TugelaDN chloroplast – ATP synthase-5 (M16843.1, e-value = 0) [a]	9.171	-18.724	20.827
CB412238	Wheat chloroplast – ATP synthase-4 (M16843.1, e-value = 0) [a]	-0.156	-1.276	5.234
CB412237	Wheat chloroplast – ATP synthase-6 (M16843.1, e-value = 0) [a]	-1.177	0.330	-6.795
CB412222	Wheat chloroplast – ATP synthase-3 (M16843.1, e-value = 0)[a]	6.959	0.834	-10.864
	Resistance gene analogs (Defence related)			
CB412258	T monococcum putative resistance protein (RGA-2) (AF326781, e-value = 4e-93)	-46.583	25.112	132.091
CB412215	Aegilops tauschii leucine-rich-like protein gene-3 (AF497474.1, e-value = 2e-44)	2.828	1.043	96.983
CB412254	Aegilops tauchii leucine-rich-like protein gene (AF497474.1, e-value = 1e-89)	-2.687	1.827	-0.097
CB412247	Aegilops tauschii putative resistance protein (RGA-2)(e-value = 2e-57, 2e-45)	22.777	2.133	-45.604
	Aegilops tauschii leucine-rich-like protein gene-4 (AF497474.1, e-value = 1e-89)	14.886	0.063	9.260
	Putative resistance protein (RGA-2) (AF326781, e-value = e-35)	-9.721	-5.733	1.050
	Aegilops tauschii leucine-rich-like protein gene	1.315	-0.095	-1.060
	Unclassified or unknown			

	Linum usitatissimum clone 5-2 (AF074884.1, e-value = 1e-177)	-1.461	0.764	9.264
	Banana unknown protein-CC92	1.473	-0.103	-1.804
	Banana unknown protein-CC14	0.387	-2.235	2.700
	Banana unknown protein-CC51	-7.0775	4.238	14.461
	Flax unknown protein-CC5-9	9.155	-4.915	-15.889
	Banana unknown protein-CC91	-1.490	-2.152	-0.408
	Flax unknown protein-CC6-5	14.349	7.279	-28.731
	Flax unknown protein-CC266L1	-18.761	-4.906	23.637
	Flax unknown protein-CC271_L6	1.172	-0.546	-12.219
	Flax unknown protein-CC261_L1	-4.597	0.739	-36.265
	Banana unknown protein-CC11	4.138	-0.408	-5.355
	Banana unknown protein-CC12	-1.188	0.851	0.716
	Flax unknown protein-CC31	-0.017	-2.694	6.121
	Flax unknown protein-CC14	-0.970	1.686	0.806
	Flax unknown protein-CC295_L1	1.111	0.526	-2.029
	Flax unknown protein-CC9	-3.193	-1.519	2.207
	Flax unknown protein-CC11	2.695	-0.431	19.00
	Banana unknown protein-CC23	1.155	-0.269	-0.231
	Flax unknown protein-CC30	2.811	0.315	-3.250
	Banana unknown protein-CC7	-2.998	1.056	-7.153
	Banana unknown protein-CCW74	-49.798	105.717	743.465
	Banana unknown protein-CCR28	2.058	-0.617	2.766
	Banana unknown protein-CCR5	-2.059	-0.506	3.169
	Banana unknown protein-CC90	-2.261	-0.266	-5.772
	Banana Unknown protein-CCY28	-13.544	-5.936	17.332
	Banana unknown protein-CC92	1.642	-0.101	-1.832
	Flax unknown protein-CC28	-3.470	2.197	51.017
CB412174	Unknown protein#307	1.022	1.210	0.186
CB412178	Unknown protein#309	-1.264	-0.565	0.668
CB412183	Unknown protein#311	16.472	14.401	5.099
CB412184	Unknownprotein#312	-1.111	-0.038	2.006
CB412192	Unknown protein#212	-1.316	0.359	0.941
CB412194	Unknown protein#277	-1.271	-0.661	0.813
CB412200	Unknown protein#285	3.691	-0.254	8.020
CB412207	PI137739 unknown protein-1	5.470	4.418	-4.709
CB412216	PI137739 unknown protein-2	8.628	2.272	-23.059
CB412238	PI137739 unknown protein-2	2.853	-1.039	-0.230
CB412239	PI137739 unknown protein-10	3.895	-0.050	-5.139
CB412254	PI137739 unknown protein-3	1.649	-0.451	-1.092
CB412257	PI137739 unknown protein-4	-1.478	-0.531	-0.237
CB412258	T aestivum line PI137739 unknown protein	-7.879	-1.383	16.725
	TugelaDN unknown protein-4	-3.059	0.812	-3.621
CB412201	'Tugela DN' unknown protein-22	2.479	0.128	-2.063
CB412225	'TugelaDN' unknown protein-12	-4.0678	-2.206	-4.4274
CB412232	'TugelaDN' unknown protein-15	-15.641	-2.161	27.238
CB412197	'TugelaDN' unknown protein-21	-23.965	-8.792	88.234
CB412255	'TugelaDN' unknown protein-2	1.388	0.119	-0.311
CB412203	TugelaDN unknown protein-23	1.030	-0.384	-0.259
CB412214	'Tugela DN' unknown protein-7	1.891	0.217	-2.029

CB412208	TugelaDN unknown protein-3	8.740	2.903	25.424
CB412223	TugelaDN unknown protein-11	4.760	-0.223	-6.237
CB412209	TugelaDN unknown protein-21	1.993	0.220	-1.255
CB412196	TugelaDN unknown protein-24	3.835	-1.010	-7.001
	'TugelaDN' unknown protein-2	-1.708	0.243	1.962
CB412243	TugelaDN unknown protein-5	2.455	0.572	-0.830
CB412211	TugelaDN unknown protein-6	-3.795	1.370	5.640
CB412221	'Tugela DN' unknown protein-10	-0.560	-1.389	2.629
CB412227	'Tugela DN' unknown protein-14	-0.218	1.351	3.505
CB412235	'Tugela DN' unknown protein-18	-0.176	1.117	-0.165
CB412219	'TugelaDN' unknown protein-7	1.786	-2.084	17.748
CB412245	'Tugela DN' Unknown protein-1	-212.293	-10.013	1.697
CB412236	'Tugela DN' Unknown protein-13	9.804	5.988	-12.688
CB412246	'Tugela DN' Unknown protein-17	1.337	0.367	-4.082
CB412234	'Tugela DN' Unknown protein-16	-2.140	-0.505	-0.331
CB412241	'Tugela DN' Unknown protein#238	-10.936	-19.501	-15.971
CB412256	'Tugela DN' Unknown protein-3	-9.173	1.737	-1.887

a Wheat chloroplast –ATP synthase CF-1 gene
b Numerical numbers indicate the number of clone for differentiation purposes

A total of 120 sequences were identified as significantly up- or down-regulated and classified into four categories. Fig. 4.4 shows the distribution between the functional classes.

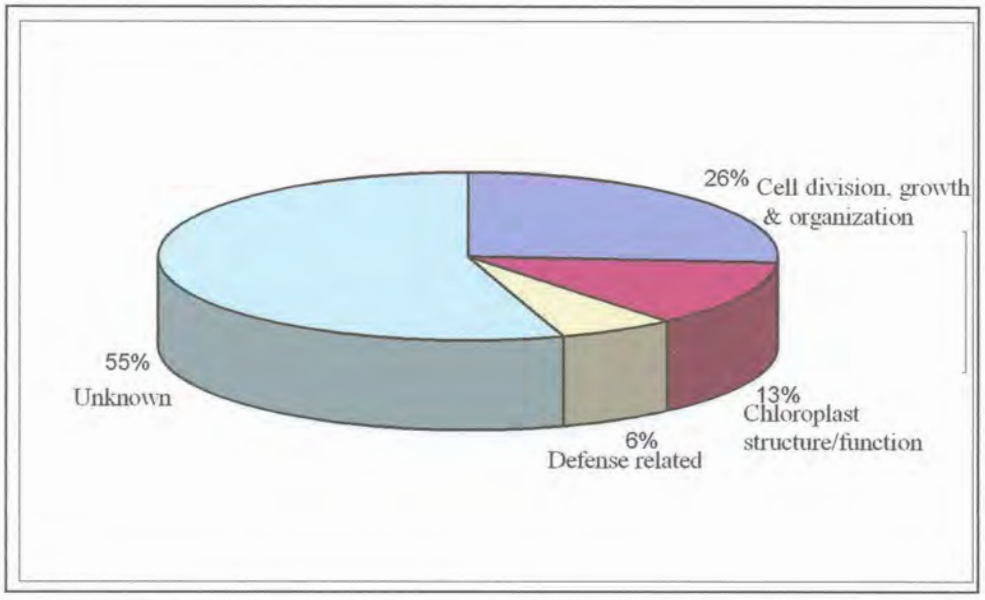


Fig. 4.4. Distribution of sequences into functional classes. Statistical representation of results obtained from microarray analysis.

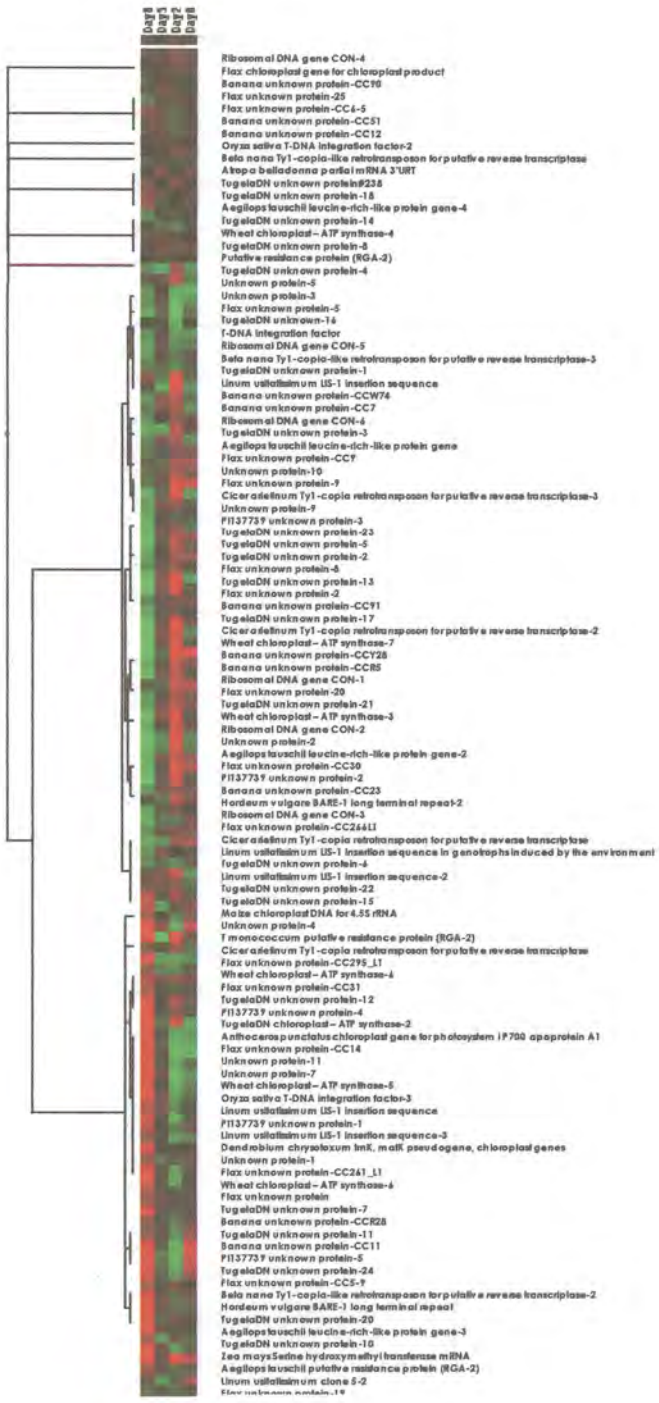


Figure 4.5. Cluster analysis of genes found significantly affected in response to Russian wheat aphid feeding. Genes were clustered using hierarchical clustering, and expression ratios in columns are shown for Day 2/Day 0; Day 5/Day 0; Day 8/Day 0. Red blocks = up-regulation, green blocks = down-regulation.

Hierarchical clustering (Eisen *et al.*, 1998) was performed in order to group the ESTs / clones with similar expression profiles in clusters (Figure 4.5). The hierarchical cluster revealed two groups consisting of a group of genes with little or no regulation, and a group of genes that are significantly up/down regulated. This last group contained two major clusters which represents the genes that are initially either significantly up- regulated and then down-regulated, or *vice versa*. The expression profiles of the ESTs / clones fall into clusters that could be categorized as (1) up-regulated through the time- trial; (2) initially up-regulated (day 0 and 2) and then down-regulated (day 5 and/or 8); (3) initially down-regulated (day 0 and 2) and then up-regulated (day 5 and/or 8); (4) down-regulated through the time trial; and (5) up-regulated (day 0 and 2), down- regulated (day 5) and then up-regulated (day 8). So far, similar data are not available concerning the up- or down-regulation of RGAs and NBS-LRRs due to stress in wheat.

4.3 RNA blot analysis and transcript quantification by Real-Time PCR

To confirm the microarray quantification, we selected five ESTs / clones that showed the most up-or down-regulation as probes for RNA gel-blot analysis. Microarray analysis indicated that one EST was unchanged, two were up-regulated, then down-regulated, and again up-regulated, and two ESTs / clones were up-regulated in response to RWA feeding. Included were two BARE-1 long terminal repeats (clone #317 and clone #182), a Ty1-copia-like retrotransposon (clone #215), and two unknown ESTs / clones. Figure 4.6 confirms that the data obtained with the microarray, the blots and Real-Time PCR were in good agreement, with a slight variation between the methods. For example, with both BARE-1 long terminal repeat ESTs / clones probes, the cross hybridization on the RNA blot indicated up-regulation also on day 2. These differences may be due to the fact that Real-Time PCR constitutes direct quantification whereas microarray and Northern blots rely on probe binding that can be influenced by numerous factors, such as unspecific binding. The Northern blot analysis is also a visual quantification whereas Real-Time PCR relies on calculated fluorescence emmitance that is much more sensitive.

	DAY 0	DAY 2	DAY 5	
BARE-1 long terminal repeat #317:				**Northern Blot**
	0.420	0.906	4.773	**Microarray**
	0	-1.5	8.2	**Real- Time PCR**
Unknown protein #314:				**Northern Blot**
	0.825	2.173	4.613	**Microarray**
	0	77.7	1.4	**Real-Time PCR**
BARE-1 long terminal repeat # 182:				**Northern Blot**
	-2.842	-0.654	5.415	**Microarray**
	0	-3.5	$5.3x10^3$	**Real-Time PCR**
Unknown protein #310:				**Northern Blot**
	-0.431	-0.225	10.211	**Microarray**
	0	$2.2x10^3$	$1.1x10^2$	**Real-Time PCR**
Ty1-copia-like retrotransposon #215:				**Northern Blot**
	-0.974	1.231	11.465	**Microarray**
	0	32	$2x10^5$	**Real-Time PCR**

Figure 4.6. *Comparison of transcript quantified by conventional RNA Northern blot analysis versus* cDNA microarray versus Real- Time PCR methods. Numerical data = fluorescence values.

4.4 Tissue specificity and response in RWA susceptible and resistance cultivars

Tissue specificity and response to feeding in RWA susceptible ('Tugela') and RWA resistant ('Tugela DN') cultivars were also tested using Northern blots (Figure 4.7). In order to confirm the differential expression of ESTs / clones in response to RWA, and in tissue mostly targeted by the insects for feeding, we selected four different probes for this purpose, e.g. chloroplast-ATP synthase (wheat chloroplast-ATP synthase-6 and 'Tugela DN' chloroplast-ATP synthase-5) and RGA-2 (wheat RGA-2 homologs from *T. monococcum* and a putative resistance protein). These probes showed significant regulation in the microarray data. The obtained results indicated significantly higher expression in resistant 'Tugela DN' when compared to the susceptible 'Tugela' cultivar. Also, expression is significantly higher in leaf tissue when compared to stem tissue. However, RWA feeding did not enhance the expression of these ESTs / clones, but rather down-regulated the expression in 'Tugela DN' at day 2 after infestation, when compared to uninfested leaf material.

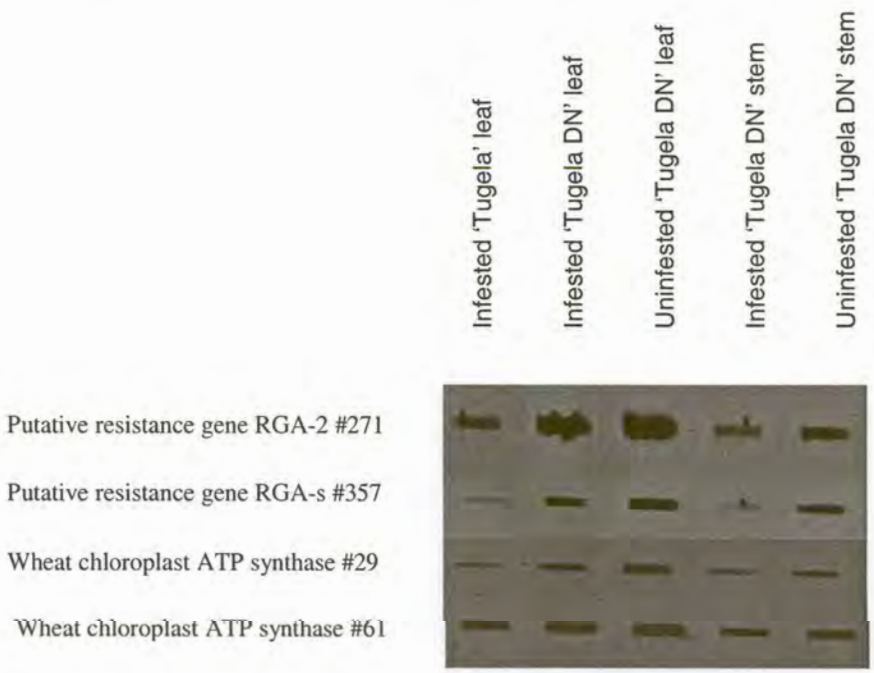


Figure 4.7. Tissue and differential expression. Tissue specific expression, as well as differential expression between RWA resistant ('Tugela DN') and susceptible ('Tugela') wheat cultivars was also verified using conventional RNA Northern blot analysis. RNA was extracted on day 0 (uninfested) and day 2 (infested) leave and stem material.

4.5 Microarray data vs Real-Time PCR data

The four genes RGA-2 #271 & # 357, ATP synthase #29 & #61 were used in a comparative study between microarray data and Real-Time PCR data concerning the up- and down-regulation of these genes on day 2 and 5 due to RWA infestation (Table 4.2). This was done to compare the two systems.

Table 4.2. Comparison between microarray values and Real-Time PCR values obtained for the same ESTs / clones. Numerical values indicate the fold up- or down-regulation.

EST	Method	Day 2	Day 5
Putative resistance gene RGA-2 #271	Microarray	22.7	2.1
	Real-Time PCR	-2	-8
Putative resistance gene RGA-2 #357	Microarray	-9.712	-5.733
	Real-Time PCR	-2.5	-7.8
Wheat chloroplast ATP synthase #29	Microarray	-1.177	0.330
	Real-Time PCR	1.5	3.7
Wheat chloroplast ATP synthase #61	Microarray	9.171	-18.724
	Real-Time PCR	1	-8

The data obtained shows some differences between the microarray and Real-Time PCR results. However, both sets of data mostly indicate the same type of regulation, whether it is up- or down-regulation. The difference may once again be due to the higher sensitivity of Real-Time PCR compared to Northern dotblots.

5. Discussion

5.1 Transcripts for analysis

Wheat transcripts were cloned using degenerate primer sets designed from the consensus NBS motif from other genome studies (e.g. *Arabidopsis* and rice). Bioinformatic analysis of the obtained transcripts revealed that 38% of the sequences were involved in metabolism, 17% were structural in function, 1% was involved in protein synthesis and modification, 9% were regulatory elements, 19% of these were

defense related and 16% had no significant homology to any published sequence in GenBank (E-value < 10^{-5}; Kruger *et al.*, 2002) (Chapter 2).

5.2 Microarray hybridization experiments and expression analysis.

Statistical analysis using the ANOVA method as suggested by Dudoit *et al.* (2001), indicated that 27% of the ESTs / clones were down-regulated, and 28% up-regulated, which is statistically too high, and thus allowed for false positives. After the fluorescence data was subjected to the statistically rigorous Mixed Model approach (Wolfinger *et al.*, 2001), only 5% of the transcripts were significantly regulated. These included the ESTs / clones previously identified as significantly regulated (Figure 4.1, Table 4.1, data points indicated in red and green). The latter model centers around two interconnected ANOVA models, namely the normalization model and the gene model. The analysis corrects amongst others, for spot position, pen position, fluorescence bias, and differences due to experiment design and biological repeats. It was found to be very rigorous, but it also excludes all false positives. The flax homolog to APC/C ubiquitin-protein ligase with known function in cell cycle regulation proved to be regulated if statistically analyzed via ANOVA, but not if analyzed via the Mixed Model approach. This was also true for the clones with significant homology to the *Linum usitatissimum* LIS-1 insertion sequences, even though it was previously shown to be induced in genotrophs by the environment (AF104351). Transcripts that were significantly regulated using both statistical approaches include wheat homologs to chloroplast-ATP synthase, *Beta nana* Ty-1-copia-like retrotransposon, an *Oryza sativa* T-DNA intergration factor, *Hordeum vulgare* BARE-1 long terminal repeat, a gene for ribosomal RNA, RGA-2 (*T. monococcum* putative resistance protein), unknown chloroplast sequence and a chloroplast gene for chloroplast product (Table 4.1).

The obtained clones were subjected to transcript quantitation using RNA blot and Real-time PCR analysis. The selected ESTs / clones changed significantly upon RWA infestation, and include all types of regulation, e.g. mainly down-regulated, mainly up-regulated, only down-regulated, only up-regulated, as well as a combination of down-/up-regulation in response to RWA feeding.

5.3 The comparison between the techniques used

We employed two other methods to verify some of the data obtained from the microarray. We used Real-Time PCR and Northern Blots and comparisons were made between some of the results obtained. It seemed that most of the time the blot data and the Real-Time PCR data concurred. Some anomalies did occur between the Real-Time PCR data and that obtained from the microarray, however, the data was comparable on the whole. The differences may be due to higher sensitivity in the Real-Time PCR that uses direct quantification compared to probe hybridization. Microarray technology is still a relatively new field of study especially the data analysis part of the experiments. From this study, however, it could be said that it is a useful method to gather large amounts of data. Verification of the data using other methods is advisable to further ensure the exclusion of false positives.

5.4 Role of expressed transcripts in cell maintenance

Little is known regarding the roles of BARE-1 long terminal repeats and Ty-1-copia-like retrotransposons, apart from its evolutionary significance and that they are important elements in genome organization (Hanson *et al.*, 2000; Katsiotis *et al.*, 2002). Chloroplast-ATP synthases (encoded in the nucleus) are key enzymes in plant metabolism, providing cells with energy in the form of ATP. The enzyme is located in the thylakoid membrane, and synthesizes ATP from adenosine diphosphate and inorganic phosphate at the expense of the electrochemical proton gradient formed by light-dependent electron flow (McCarty *et al.*, 2000; Mellwig and Bottcher, 2003). RGA-2 is an NBS-LRR-like protein with a putative receptor-like function, and was suggested to be involved in signal transduction (Whitham *et al.*, 1994; Jackson and Taylor, 1996; Pan *et al.*, 2000; Cannon *et al.*, 2002) and pathogen/pest recognition (Dangl and Jones, 2001). However, it has also been shown with tomato I2 and Mi R proteins to bind ATP and GTP (Tameling *et al.*, 2002), making it a potential energy carrier.

It is interesting to note that these ESTs / clones showed the highest levels of up- and down-regulation during this experiment. However, their involvement in the RWA resistance response can only be speculated on.

Also interesting was the fact that so few resistance genes were detected, probably due to their low abundance. In model plant species such as *Arabidopsis* and rice, a number of disease resistance genes have been isolated, but it has been difficult in large and repetitive genomes such as barley and wheat (Feuillet *et al.*, 1993). Those few that were, seemed to be only slightly up- or down- regulated. Previous studies on RWA feeding induced responses indicated the induction of PR-proteins and other defense related proteins, for example chitinases, peroxidases, β-1,3-glucanases (Botha *et al.*, 1998; Van der Westhuizen *et al.*, 1998a,b), lipoxigenases and ROS (Van der Westhuizen *et al.*, 2002) 3 to 12 days post-infestation. Since our study showed high regulation of ATP-synthase and lower regulation of the RGAs, it could be postulated that RWA infestation leads to a resistance response linked more to cell maintenance than a direct defense. This hypothesis will however be the subject of future studies.

6. References

Altschul S.F., Madden T.L., Schäffer A.A., Zhang J., Zhang Z., Miller W. & Lipman D.J. 1997. *Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.* Nucleic Acids Research 25: 3389-3402.

Botha A-M., Nagel M.A.C., van der Westhuizen A.J. & Botha, F.C. 1998. *Chitinase isoenzymes in near-isogenic wheat lines challenged with Russian wheat aphid, exogenous ethylene and mechanical wounding.* Botanicla Bulletin of Academia Sinica **39:** 99-106.

Brazma A., Robinson A., Cameron G. & Ashburner M. 2000. *One-stop shop for microarray data.* Nature 403: 699-700.

Brown P.O. & Botstein D. 1999. *Exploring the new world of the genome with DNA microarrays.* Nature Genetics Supplement 21: 33-37.

Chu T.-M., Weir B. & Wolfinger R.D. 2002. *A systematic statistical linear modeling approach to oligonucleotide array experiments.* Mathematical Biosciences 176: 35-51.

Cannon S.B., Zhu H., Baumgarten A.M., Spangler R., May G., Cook D.R. & Young N.D. 2002. *Diversity, distribution, and ancient taxonomic relationships within the TIR and non-TIR NBS-LRR resistance gene subfamilies.* Journal of Molecular Evolution 54: 548-562.

Dangl J.L. & Jones J.D. 2001. *Plant pathogens and integrated defense responses to infection.* Nature 411: 826-833.

Dudoit S., Yang Y.H., Callow M.J. & Speed, T.P. 2001. *Statistical methods for identifying differentially expressed gene in replicated cDNA microarray experiments.* Submitted to JASA.
http://www.stat.berkeley.edu/users/terry/zarray/Html/papersindexhtml.

Eisen M.B, Spellman P.T., Brown P.O. & Botstein D. 1998. *Cluster analysis and display of genome-wide expression patterns*. Proclamations of the National Academy of Science USA 95: 14863-14868.

Feuilet C., Travella S., Stein N., Albar L., Nublat A. & Keller B. *Map-based isolation of the leaf rust disease resistance gene Lr10 from the hexaploid wheat (Triticum aestivum L.) genome*. Proceedings of the 10th international wheat genetics symposium. Parstum, Italy (1-6 September 2003) volume 1: 357-360.

Hanson R.E., Islam-Faridi M.N., Crane C.F. Zwick M.S., Czeschin D.G., Wendel J.F., McKnight T.D., Price H.J. & Stelly D.M. 2000. *Ty-1-copia-retrotransposon behaviour in a polyploid cotton*. Chromosome Research 8: 73-76.

Jackson A.O. & Taylor C.B. 1996. *Plant-microbe interactions: Life and death at the interface*. The Plant Cell 8: 1651-1668.

Katsiotis S.G., Hagidimitrriou M. & Loukas M. 2002. *Genomic and chromosomal organization of Ty1-copia-like sequences in Olea europaea and evolutionary relationships of Olea retroelements*. Theoretically Applied Genetics 104: 926-933.

Kerr M.K & Churchill G.A. 2001. *Experimental design for gene expression microarrays*. Biostatistics. www.jax.org/research/churchill/pubs/index.htlm.

Kruger W.M., Pritsch C., Chao S. & Muehlbauer G.J. 2002. *Functional and comparative analysis of expressed genes from wheat spikes infected with Fusarium graminearum*. MPMI 15 (5): 445-455.

McCarty R.E., Evron Y. & Johnson E.A. 2000. *The chloroplast ATP aynthase: A rotary enzyme?* Annual Review of Plant Physiology Plant Molecular Biology 51: 83-109.

Mellwig C. & Bottcher B. 2003. *A unique rsting position of the ATP-synthase from chloroplasts*. Journal of Biological Chemistry 278: 18544-18549.

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

Pastori G. M., Kiddle G., Antoniw J., Bernard S., Veljovic-Jovanovic S., Verrier P.J., Noctor G. & Foyer C.H. 2003. *Leaf vitamin C contents modulate plant defense transcripts and regulate genes that control development through hormone signaling.* Plant Cell 15: 939-951.

Quackenbush J. 2001. *Computational analysis of microarray data.* Nature Genetics Reviews 2: 418-427.

SAS Institute Inc. (1999). SAS/STAT Software version 8 (SAS Institute, Cary, NC, 1999).

Sambrook J., Fritsch E.F. & Maniatis, T. 1989. *Molecular cloning: A laboratory manual, 2nd edition, Vol. 1, 2, 3.* Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press.

Schena M., Shalon D., Davis R.W. & Brown P.O. 1995. *Quantitative monitoring of gene expression patterns with a complementary DNA microarray.* Science 270: 467-470.

Schenk P.M., Kazan K., Wilson I., Anderson J.P., Richmond T., Somerville S.C. & Manners J.M. 2000. *Coordinated plant defense responses in Arabidopsis revealed by microarray analysis.* PNAS 97 (21): 11655-11660.

Tameling W.I., Elzinga S.D., Darmin P.S., Vossen J.H., Takken F.L., Haring M.A. & Cornelissen B.J. 2002. *The tomato R gene products I-2 and MI-1 are functional ATP binding proteins with ATPase activity.* Plant Cell 14: 2541-2554.

Van der Westhuizen A.J., Qian X-M. & Botha A-M. 1998a. *Differential induction of apoplastic peroxidase and chitinase activities in susceptible and resistant wheat cultivars by Russian wheat aphid induction.* Plant Cell Reports 8: 132-137.

Van der Westhuizen A.J., Qian X-M. & Botha A-M. 1998b. *β-1,3-glucanases in wheat and resistance to the Russian wheat aphid.* Physiologia Plantarum 103: 125-131.

Van der Westhuizen A.J., Qian X-M., Wilding M. & Botha A-M. 2002. *Purification and immunocytochemical localization of a wheat β-1,3-glucanase induced by Russian wheat aphid infestation.* South African Journal of Science 98: 197-202.

Wang Y.H. & Speed T. 2002. *Design of microarray expression experiments. In: DNA microarrays: A molecular cloning manual.* Bowtell, D. & Sambrook J. (Eds). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA. Part 1: 513-525

Whitney L.W., Becker K.G., Tresser N.J., Caballero-Ramos C.I., Munson P.J., Prabhu V.V., Trent J.M., McFarland H.F. & Biddison W.E. 1999. *Analysis of gene expression in multiple sclerosis lesions using cDNA microarrays.* Annals of Neurology 46: 425-428.

Whitham S., Dinesh-Kumar S.P., Choi D., Hehl R., Corr C. & Baker B. 1994. *The product of the tobacco mosaic virus resistance gene N: Similarity to Toll and the interleukin-1 receptor.* Cell 78: 1101-1115.

Wolfinger R.D., Gibson G., Wolfinger E.D., Bennett L., Hamadeh H., Bushel P., Afshari C. & Paules, R.S. 2001. *Assessing gene significance from cDNA microarray expression data via mixed models.* Journal of Computational Biology 8: 625-637.

CHAPTER 5

SUMMARY AND PERSPECTIVE

Chapter 5:

Summary and Perspective

The most challenging aspect of the research done for this PhD was surely the chosen organism, namely wheat. As discussed in the literature review, it is noted that wheat has a very big and complex genome. Thus, to isolate and analyze specific genes, especially those not expressed continuously remains a big challenge. Since wheat is such an important crop in South Africa and all over the world, the study of wheat resistance responses is very important. The amount of evidence pointing to the role of NBS-LRRs in plant resistance is overwhelming as can be seen from the literature review.

The main question behind this PhD was if NBS-LRRs are involved in the resistance response of wheat due to RWA infestation. To answer this, it was considered significant to isolate ESTs from the resistant wheat line 'Tugela DN' after RWA infestation and to ascertain their characteristics as well as their involvement in the resistance response. These ESTs might indicate the involvement or not of NBS-LRRs.

The first technical aspect accomplished in this project was the construction of a cDNA library from the resistant wheat line 'Tugela DN' containing NBS-LRRs. This was done using degenerate NBS primers. This was a very useful technique, being very easy to use and highly repeatable. The results obtained allowed the identification and characterization of numerous ESTs, allocating to each their function and distribution in the genome. Many NBS-LRRs were identified, however, this approach alone did not lend itself to positive identification of these domains in the RWA resistance response. Identification of these domains could not rule out their involvement in other resistance responses in wheat or even other protein-protein interactions.

A new discovery was the use of SSH to identify novel sequences in the RWA resistance response. To our knowledge, SSH has not previously been employed on

wheat to identify any resistance genes. This technique also allowed the compilation of a database containing sequences involved in RWA resistance response. SSH was, generally, a relatively quick method and resulted in the isolation of numerous unknown sequences, of which seven were studied. Three proved to be up-regulated only in the resistant wheat line during RWA infestation and thus, putatively involved in the plant resistance response. This study is one of the first to generate characterization data regarding RWA response sequences. However, none of these sequences showed any homology to NBS-LRRs or any other published sequences. It can also not be conclusively said that these sequences are genes. Future studies should include the attainment of the complete sequences, to determine if they are genes. In this case, the isolation of novel sequences supposedly involved in the plants' response to RWA infestation was successful. The technique does remain a useful way to screen big or complex genomes for novel sequences, however the abundance of these sequences in the genome have to be determined. The normalization step in SSH makes it a worthwhile approach because this leads to the enrichment of differentially expressed genes and high complexity of the subtracted cDNA pool. Future studies could include the use of cDNA-RDA. It would be interesting to compare the results obtained from SSH and that of cDNA-RDA.

The third and most conclusive new result came from the use of micrarray technology to determine the regulation of ESTs during RWA infestation. The first known wheat microarray cDNA chip was made which is in itself a big accomplishment, due to the fact that it can be used in future or other, different studies. The most conclusive results in this part of the study came from the up-regulation of chloroplast ATP-synthase and RGA-2 (NBS-LRR homologue) genes. We know that chloroplast ATP-synthase is encoded in the cell nucleus and is involved in plant metabolism, supplying the cells with energy in the form of ATP. Previous research has also shown that NBS-LRRs have the ability to bind ATP and GTP, thus potentially being an energy carrier. This evidence suggests that both RGA-2 (an NBS-LRR homologue) and the chloroplast ATP-synthase genes are involved in the RWA resistance response. This is done by their involvement in cell maintenance, increasing the plants' tolerance to stress by keeping the photosynthetic machinery intact, and not by direct defence. A future study could involve the transformation of a susceptible wheat line with these two genes, causing over-expression. Would the plant be rendered resistant to RWA

infestation? An improvement in this study would have been to have more clones on the slide. A bigger variety of controls would also have improved the study.

A fourth new result was the implementation of Real-Time PCR to relatively quantify the isolated genes' up- or down-regulation during RWA infestation. This is a relatively new field of study and no published data exists, to my knowledge, using wheat as subject to determine gene regulation. This method proved to be quite expensive and it took a while to optimize. The results, however, were very sensitive, when compared to Northern blots and the repeatability of the experiments were very high. The optimization of this method has led to numerous other studies being done in our lab, especially the confirmation of Northern blot results.

A few technical recommendations can be made regarding future studies. Susceptible plant ESTs could be spotted onto the microarray slide. Then a comparison can be made between the expression profiles of the RWA infested resistant and susceptible plant genes. Genes from a known defence pathway could also be spotted onto a slide to determine each one's specific regulation during RWA infestation. The specific regulation mechanisms of NBS-LRRs, whether as energy carrier or rather as a direct defence response gene, should also be studied further.

The strength of this PhD lies in the implementation of various techniques new to our institute, using wheat as study material. New technology was used for the detection and studying of genes involved in response to RWA infestation. It has paved the way for other comparison studies using microarray technology, as well as Real-Time PCR and SSH.