APPENDIX 1

SUMMARY

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Analysis of gene expression in *Triticum aestivum* L. cv. 'Tugela DN' after Russian wheat aphid (*Diuraphis noxia* Mordvilko) infestation.

Chantal van Niekerk
Promotor: Prof. A-M Botha-Oberholster
Co-promotor: Prof. K.J. Kunert
Department of Genetics
Philosophiae Doctor (Ph.D)

The Russian wheat aphid (*Diuraphis noxia,* Mordvilko; RWA) causes serious damage to wheat crops annually. Breeding efforts have led to the development of wheat lines (e.g. 'Tugela DN'(*Dn1,* SA1684 / Tugela*5) that are resistant to RWA infestation.

During RWA infestation of resistant wheat, differential expression of genes occurs due to the feeding of the aphid. Nucleotide binding sites (NBS) have been found to be conserved regions in many resistance genes from various plant species. In this study a PCR based approach was followed, using degenerate primers to target and amplify NBS sites from cDNA synthesised from RNA isolated from leaf tissue after infestation with RWA. All amplified fragments were isolated, cloned 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.

By making use of 'Tugela', 'Tugela DN' and suppression subtractive hybridisation (SSH), fragments were isolated that may be involved in the wheat plant's resistance to RWA infestation. The SSH fragments were cloned and all clones sequenced. Thereafter, the sequences were subjected to a GenBank database search using the BLASTX algorithm. Interestingly, all obtained sequences showed no significant homology (e-value $< 10^{-5}$) with any known protein. However, Real-Time PCR and Northern blot analysis confirmed involvement in the RWA resistance response through up-regulation from 5-fold to 5.4-fold of selected SSH fragments.

Previous studies show that 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, occur. 256 wheat NBS-LRR sequences were obtained using degenerate primers sets designed from the consensus NBS motif from other genome studies. Purified mRNA from infested material, containing the RWA resistance genes *Dn1* was isolated 0, 2, 5 and 8 days after infestation, post-labeled with Cy3- or Cy5- fluorescent dyes and hybridized to the arrays. Statistical analysis of the expression data revealed the up- and down-regulation of 12% 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 and WIR pathogen R genes were also induced in the RWA resistance response. The expression levels of a subset of clones were verified by Real-Time PCR and Northern blot analysis.

APPENDIX 2

Published in African Journal of Biotechnology 2(4): 75-81

African Journal of Biotechnology Vol. 2 (4), pp. 75-81, April 2003
Available online at www.academicjournals.org/AJB
ISSN 1684-5315 © 2003 Academic Journals

Functional and comparative analysis of expressed sequences from *Diuraphis noxia* infested wheat obtained utilizing the conserved Nucleotide Binding Site

Lynelle Lacock, Chantal van Niekerk, Shilo Loots, Franco du Preez and Anna-Maria Botha*

Department of Genetics, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, ZA0002 South Africa

Accepted 10 March 2003

Russian wheat aphid (*Diuraphis noxia*, Morvilko; RWA) is a major pest on wheat, barley and other triticale in South Africa. Infestation by the RWA results in altered protein expression patterns, which is manifested as differential expression of gene sequences. In the present study, Russian wheat aphid resistant (Tugela DN, Tugela*5/SA2199, Tugela*5/SA463, PI 137739, PI 262660, and PI 294994) and susceptible triticale (Tugela) were infested and cDNA synthesized. A PCR based approach was utilized to amplify the nucleotide binding site conserved region to obtain expressed sequence tags (ESTs) with homology to resistance gene analogs (RGAs). The approach proved highly feasible when the isolation of RGAs is the main objective, since 18% of all obtained ESTs showed significant hits with known RGAs, when translated into their corresponding amino acid sequences and searched against the nonredundant GenBank protein database using the BLASTX algorithm.

Key words: Resistance gene analogs, degenerate PCR, nucleotide-binding site-leucine rich repeat resistance genes, *Aegilops tauschii*.

INTRODUCTION

Russian wheat aphid (*Diuraphis noxia*, Morvilko; RWA) is one of the most adaptable insects that is recognized as a pest of wheat, barley and other triticale (Bryce, 1994; Walters et al., 1980). Infestation can occur shortly after the emergence of the wheat plants and the aphids are found on the newest growth and the axils of the leaves, but damage is greatest when the crops start to ripen. This is due to the twisting and distortion of the heads and the resulting failure to emerge properly (Unger and Quisenbury, 1997). Further symptoms of RWA feeding on susceptible cultivars include longitudinal streaking and leaf rolling, which under severe infestation leads to a drastic reduction in effective leaf area (Walters et al., 1980). Infestation by the RWA also results in altered protein expression patterns, which is manifested as differential expression of total proteins, and specific pathogenesis-related proteins like chitinases, ß-1,3-glucanases and peroxidases (Bahlmann, 2002; Botha et al., 1998; Van der Westhuizen et al., 1998a,b, 2002; Van der Westhuizen and Botha, 1993; Van der Westhuizen and Pretorius, 1996). The use of RWA-resistant cultivars, however, may reduce the impact of this pest on wheat production and in the same time reduce environmental risks and control costs due to chemical spraying (Tolmay et al., 1999). The need for more RWA tolerant plants places emphasis on obtaining resistance candidate genes, as well as on the understanding of the underlying mechanisms of defense against the RWA.

Disease resistance genes have been isolated and characterized at the molecular level in several plant species such as *Arabidopsis*, tobacco, tomato and wheat (Jones and Jones, 1997; Cannon et al., 2002). Resistance gene products specifically recognize and provide resistance towards a large number of pests and pathogens (Seah et al., 1998; Pan et al., 2000). These genes can be divided into four broad, structurally distinct classes. The first class of resistance genes belongs to the serine-threonine kinases (Martin et al., 1993; Ritter and Dangl, 1996). The protein kinases phosphorylate serine/threonine residues and thus control certain signaling networks during the resistance response. The second class of resistance genes encodes putative transmembrane receptors with extracellular leucine rich repeat (LRR) domains (Jones et al., 1994; Dixon et al., 1998). The third class encodes for a receptor-like kinase and combines qualities of both the previous classes. Both the LRR domain and the protein kinase regions are encoded in the same protein. The fourth class, which

*Corresponding Author; E-mail: ambothao@postino.up.ac.za, tel: +27 12 420 3945, fax: +27 12 420 3947

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

represents the majority of plant disease resistance genes cloned so far, is the nucleotide-binding site-leucine rich repeat (NBS-LRR) resistance genes. The NBS-LRR class of genes is abundant in plant species. In *Arabidopsis*, it has been estimated that at least 200 different NBS-LRR genes exist making up to 1% of the genome (Ellis et al., 2000; Sandhu and Gill, 2002).

The NBS-LRR genes contain three distinct domains: a variable N-terminus, a nucleotide-binding site and leucine rich repeats. Two types of N-termini are present in NBS-LRR. One kind contains a leucine zipper or coiled-coil sequence that is thought to facilitate protein-protein interactions. The coiled-coil motif has been found in the N terminus of both dicotyledons and cereals (Pan et al., 2000; Cannon et al., 2002). The second kind of N-terminus has been described only in dicotyledons and is similar to the cytoplasmic signaling domains on the *Drosophila* Toll- or the mammalian interleukin receptor-like (TIR) regions (Whitham et al., 1994; Cannon et al., 2002). These NBS regions are found in many ATP and GTP-binding proteins that act as molecular switches (Jackson and Taylor, 1996). These genes regulate the activity of proteases that can initiate apoptotic cell death. Since defense mechanisms in plants include the hypersensitive response, which is very similar to apoptosis, the common occurrence of NBS domains in both plants and animals could be an indication of similar functioning (Cannon et al., 2002).

NBS-LRR homologues encode proteins that are structurally closely related. This suggests that they have a common function in signal transduction pathways, even though they confer resistance to a wide variety of pathogen types. The conservation between different NBS-LRR resistance genes enables the use of polymerase chain reaction (PCR)-based strategies in isolating and cloning other R gene family members or analogs using degenerate primers for these conserved regions. Strategies using degenerate primers have been successfully utilized in the cloning of other putative NBS-LRR resistance gene analogs (RGA) from potato (*Solanum tuberosum* L.)(Leister et al., 1996), soybean (*Glycine max* L. Merr.)(Yu et al., 1996) and citrus (Deng et al., 2000).

The identification and analysis of expressed sequence tags (ESTs) provide an effective tool to study thousands of genes expressed during plant development and their response to varying environmental conditions (Gyorgyey et al., 2000; White et al., 2000; Yamamoto and Sasaki, 1997) in complex genomes like wheat. The development of EST databases further provides a resource for transcript profiling experiments and studies of gene expression (Mekhedov et al., 2000; Schenk et al., 2000).

The aim of this study was to survey the expressed sequence tags obtained through PCR-based strategies utilizing the conserved nucleotide binding site motifs in an effort to increase the efficacy of isolating resistance gene candidates, from the complex hexaploid wheat genome.

MATERIALS AND METHODS

Plant Material

The plant materials in the study were *Aegilops tauschii*, the near isogenic lines 'Tugela DN' (Tugela*5/SA1684, *Dn1*), Tugela *Dn2* (Tugela*5/SA2199), Tugela *Dn5* (Tugela*5/SA463) and Tugela (RWA susceptible), as well as RWA tolerant lines PI 137739 (SA1684, *Dn1*), PI 262660 (SA2199, *Dn2*) and PI 294994 (SA463, *Dn5*). The plants were 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 24°C, and the plants were watered daily. Half of the wheat seedlings were infested with RWA (10 aphids per plant) at the 3-4-leaf growth stage. The second and third leaves from uninfested and infested plants were removed after one week for analysis. The aphids were removed from the infested leaves under running water to prevent aphid derived nucleic acid contamination during the RNA isolation. The leaves were dried and used immediately for total RNA isolation.

Treatment of glassware, plastic ware and solutions

All glassware, plastic ware and solutions used, up to the second strand cDNA synthesis, were treated and then kept free of RNases. The glassware was treated overnight in 0.1% (v/v) diethyl pyrocarbonate (DEPC), autoclaved for 20 min at 121°C and baked at 200°C for 3-4 hours (Sambrook et al., 1989). The mortars and pestles were washed in 0.25M HCl for 30 min, prior to DEPC treatment, autoclaving and baking. All plastic ware and solutions, except those containing Tris (2-Amino-2-(hydroxymethyl)-1,3-propandiol), were DEPC treated and autoclaved.

Total RNA Isolation and cDNA synthesis

Total cellular RNA was extracted using an acid guanidium thiocyanate-phenol-chloroform extraction method described by Chomczynski and Sacchi (1987). The RNA samples were stored at -80°C for further use. The RNA concentration was determined on a Beckman DU®-64 spectrophotometer, by reading the absorbance at 260 nm. The 260/280 ratio was determined to indicate the level of protein contamination (Sambrook et al., 1989). The integrity of the RNA was confirmed by analyzing both the infested and uninfested total RNA on a 2 % (w/v) agarose gel (Sambrook et al., 1989). The molecular mass standard used was λ DNA digested with *Eco*RI and *Hin*dIII (Sambrook et al., 1989). Isolated RNA was electrophoresed at 100 V for 30 min and visualized under UV light with ethidium bromide (EtBr) staining.

mRNA Isolation

The mRNA was purified from the total RNA using Oligo(dT) Cellulose affinity chromatography (GibcoBRL, Life Technologies). The synthesis of cDNA was carried out using either the Roche Molecular Biochemicals cDNA Synthesis System according to manufacturers specifications, or the RLM-RACE system (GeneRacer Kit, Invitrogen). Both the uninfested and the infested wheat mRNA were used as the substrate for the cDNA synthesis reaction. The ds cDNA was purified by the QIAquick Spin Purification Procedure (QIAGEN). The cDNA was eluted with water and the concentration determined spectrophotometrically and stored at -20°C.

When making use of the RLM-RACE system, the mRNA was dephosphorylated with calf intestinal phosphatase to remove the 5' phosphates and decapped with tobacco acid pyrophosphatase

Table 1. Functional annotation of expressed sequence tags (ESTs) that produced BLASTX hits.

BLASTX annotation	E-value	No. of dbESTs hits
Mitochondrial 26S rRNA (*Triticum aestivum*, Z11889; *Zea mays*, K01868; *Beta vulgaris*, AP000397; *Arabidopsis thaliana*, Y08501)[a]	0.0	20
Mitochondrial DNA (*Secale cereale*, Z14059)	0.0	2
Mitochondrial 18S rRNA (*Zea mays*, X00794)	0.0	4
Mitochondrial 23S rRNA (*Beta lupin*, X87283)	4.00E-13	1
Mitochondrial 16S rRNA (*Beta japonicum* , Z35330)	0.0	2
SSU[b] rRNA (*Drimys winteri*, AF197162; *Tetracentron sinense*, AF193998;*Grevillea robusta*, AF193995; *Trochodendron aralioides*, AF161092; *Nelumbo nucifera*, AF193983; *Platanus occidentalis*, AF161090)	2.00E-87	6
S7 ribosomal fragment (*Triticum aestivum*, X67242)	1.00E-03*	1*
S13 ribosomal fragment (*Triticum aestivum*, Y00520)	0.085*	1*
Chloroplast DNA (*Triticum aestivum*, AB042240; *Oryza sativa*, AB042240.3; *Zea mays*, X15901; *Hordeum vulgare*, X86563; *A. crassa*, X86563.2; *A. columnaris*, X00408; *Clinostigma savoyanum*, X00408.1,)	0.0	15
Long terminal repeat (*Hordeum vulgare*, Z84569)	7.00E-03*	2*
Actin gene (*Triticum monococcum*, AF326781)	8.00E-31	8
Resistance gene analogue2(*Triticum monococcum*, AF326781)	7.00E-13	6
Retrotransposon, MITE[c] (*Hordeum vulgare*, AB022688)	3.00E-76	5
Tonoplast DNA (*Hordeum vulgare*, AF254799)	6.00E-50	3
HMW[d] glutenin gene (*Aegilops tauschii*, AF497474)	1.00E-116	2
Noduline-like-protein (*Triticum monococcum*, AF326781)	2.00E-43	2
Chromosome condensation factor (*Triticum monococcum*, AF326781)	2.00E-43	2
mRNA for SSU, Rubisco (*Triticum aestivum*, K01229; *Secale cereale*, M37328; *Hordeum vulgare*, AJ131738; *Triticum timopheevii*, U43493; *Triticum urartu*, AB020955; *Oryza sativa*, AB020954; *Avena strigosa*, AF052305; *Avena maroccana*, AF097360; *Aegilops squarrosa*, AF104250; *Aegilops bicornis*, AB020938; *Aegilops sharonensis*, AB020936; *Aegilops longissima*, AB020935)	6.00E-58	24
ATP synthase β subunit (*Clinostigma savoryanum*, AB020933; *Elaeis oleifera*, AF449171; *Cyphophoenis nucele*, AY012452; *Howea belmoreana*, AY012445; *Phoenix canariensis*, AY012435; *Linospadix longicruris*, AF209652)	1.00E-168	6
Aldehyde dehydrogenase (*Zea mays*, AF449172)	3.00E-14	1
Microsatellite fragment (*Oryza sativa*, AF348415)	3.00E-07*	2*
Receptor-like kinase (*Triticum aestivum*, AY021654)	6.00E-07*	2*
LRR[e] 19 (*Triticum aestivum*, AF325196)	6.00E-65	2
WIR pathogen R[f] gene (*Triticum aestivum*, AF325196)	1.00E-125	3
Leucine-rich-like protein (*Aegilops tauschii*, X87686)	1.00E-55	2
RGA link to resistance loci in rice (*Oryza sativa*, AB022168)	4.00E-16	1
actin (ACT-1) gene, partial cds (*Triticum monococcum*, AF326781)	4.00E-16	1
putative chromosome condensation factor (CCF) (*Triticum monococcum*, AF326781)	4.00E-16	1
putative resistance protein(RGA-2) (*Triticum monococcum*, AF326781)	4.00E-16	1
putative nodulin-like protein (NLL) gene (*Triticum monococcum*, AF326781)	6.00E-26	1
chloroplast matK gene for maturase (*Cycas pectinata*, AB076238.1; *Zamia angustifolia*, AB076567.1)	6.00E-26	2
clone tac 923.8 3' Ac insertion site sequence (*Zea mays*, AY065582.1)	1.00E-29	1
Genomic seq. BAC F27F5 (*Arabidopsis thaliana*, AC007915.3)	0.011*	1*
Germline Ig heavychain var. region (*Macaca mulatta*, U57580)	0.043*	1*
Genomic DNAChr. 1 BAC clone: OJ1294_F06 (*Oryza sativa*, AP004326.3)	0.0	1
Wheat chloroplast ATP synthase CF-1 gene, beta and epsilon subunits, complete cds, and Met-tRNA gene (*Triticum aestivum*, M16843)	7.00E-29	1
Ty1-copia-like retrotransposon partial pol pseudogene, clone Tbn-1 (*Beta nana*, AJ489202)	2.00E-32	1
microsatellite DNA, CA-repeat (AC) (*S. salar*, Y11455)	0.0	1
Predicted membrane protein (*Clostridium acetobutylicum*, AE007615-4)	1.00E-03*	1*
Nucleotide binding site LRR protein-1 (*Oryza sativa*, AY043283)	4.00E-22	1
Nucleotide-binding leucine-rich-repeat protein 1 (*Oryza sativa*, AF271293)	1.00E-12	1
Nucleotide-binding leucine-rich-repeat protein-like (*Oryza sativa*, AP003802)	4.00E-14	2
Putative NBS-LLR type resistance protein (*Oryza sativa*, AC097447)	3.00E-13	1
Putative disease resistance protein (*Oryza sativa*, AC087181)	1.00E-55	
Resistance gene candidate CC-NBS-LLR Class (*Arabidopsis thaliana*, NM_175742.1)	1.00E-33	1
F12M16 (*Arabidopsis thaliana*, AC008007)	4.00E-43	1
Disease resistance complex protein NBS-LRR class (*Arabidopsis thaliana*, NP_188065.1)	1.00E-43	1
Putative disease resistance protein CC-NBS Class (*Arabidopsis thaliana*, NM_104655)	1.00E-43	1
PRM1 homolog (*Arabidopsis thaliana*, AB028231)	1.00E-43	1
Putative RGA PIC23 -(*Lactuca sativa*, AF017751)	3.00E-16	1
Resistance complex protein I2-C-2 (*Lycopersicon peruvianum*, AF004879)	7.00E-13	1
NBS-LRR resistance protein candidate (*Lactuca sativa*, AF113949)	7.00E-13	1
Serine/threonine kinase protein (*Triticum aestivum*, *Pseudoroegneria*)	1.00E-18	2
Conserved hypothetical protein (*E. coli*, NC_002655, NC_003047)	6.00E-31	2
NBS-LRR type protein (r15) gene (*Oryza sativa*, AF032702)	6.00E-13	1

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

BLASTX annotation	E-value	No. of dbESTs hits
Resistance protein candidate RGC2A pseudogene (*Lactuca sativa*, AF072268)	1.00E-14	1
Polymyxin ß-resistance protein (*Saccharomyces*, S569090)	1.00E-04*	1*
Receptor like protein (*Arabidopsis thaliana*, NP_190339)	3.00E-13	1
Thioredoxcin (*Triticum aestivum*, AJ005840)	1.00E-66	1
18S ribosomal RNA (*Triticum aestivum*, AY049040)	0.440*	1*
Intergrase/recombinase (*Brucella melitensis*, AE009541)	1.00E-03*	1*
Intergrase-like protein (Bacteriophage H191, AJ236875)	0.61*	1*
Microsatellite DNA (*Entandrophragma*, AJ420885, *Cocos*, AJ458311)	3.00E-48	9
mRNA sequence (*Zea mays*, AY105736)	2.00E-287	2
Aminotransferase-like protein (*Oryza sativa*, AF105736)	5.00E-52	3
Genomic DNA chromosome 4 (*Oryza sativa*, AL662950)	9.00E-27	2
Giant Cell protein mRNA (*Lycopersicon esculentum*, L24012)	7.00E-03*	1*
RbcL for Rubisco (*Oryza sativa*, X62117; *Triticum aestivum*, X62119; *Aegilops tauschii* , X04789; *Avena*, L15300)	9.00E-44	9
Mitochondrial gene for tRNA-His (*Oryza sativa*, D13101)	2.00E-35	1

a. NCBI accession number of homologous sequence
b. Small subunit
c. Miniature inverted terminal repeat element
d. High molecular weight
e. Leucine rich-repeat
f. Resistance
* Protein with discernable function

(TAP) to remove the 5' cap. The dephosphorylated, decapped mRNA was ligated to a GeneRacer™ RNA oligo using the GeneRacer Kit (Invitrogen). The ligated mRNA was reverse-transcribed using SUPERSCRIPT™ II RT (Invitrogen) and the GeneRacer™ Oligo dT Primer to create RACE-ready cDNA with known priming sites at the 5' and 3' ends. The 5' ends were amplified using a reverse degenerate nucleotide-binding site primer and the GeneRacer™ 5' Primer. The degenerate oligonucleotide primers were based on the amino acid sequences of two highly conserved motifs of the NBS in the tobacco *N* and *Arabidopsis RPS2* genes (Yu et al., 1996). The 3' ends were amplified using a forward degenerate nucleotide-binding site primer and the GeneRacer™ 3' primer (GCTGTCAACGATACGCTACGTAACGGC ATGA CAGTG$(T)_{18}$). The cycling parameters used for the GeneRacer™ reactions were five cycles consisting of 94°C for 30 sec and 72°C for 1 min, five cycles of 94°C for 30 sec, 70°C for 30 sec and 72°C for 1 min and twenty cycles of 94°C for 30 sec, 68°C for 30 sec and 72°C for 1 min.

Degenerate NBS-PCR

For the amplification of NBS sequences from the synthesized cDNA the following degenerate primers was applied: NBS-F1 (GGAATGG GNGGNGTNGGNAARAC); NBS-R1 (YCTAGTTGTRAYDATDAYY YTRC), where R = A/G, Y = C/T, D = A/G/T, H = A/C/T, N = A/G/C/T. The PCR reaction consisted of 50 µM of each primer, 50 ng of the RT template, 1X reaction buffer (Promega), 2.5 mM $MgCl_2$, 0.2 mM of each dNTP and 2.5U of *Taq* DNA polymerase, and 1.3 M betaine to increase primer annealing. Thirty cycles of PCR, consisting of 95°C for 1 min, 55°C for 1.5 min, and 72°C for 1 min, were performed in a Perkin-Elmer GeneAmp PCR System 9700 DNA thermal cycler (Applied Biosystems).

Cloning and Analysis of NBS-PCR Products

The PCR products were purified from an agarose gel slice using a Geneclean III Kit (Bio101). These fragments were cloned into the pGEM®-T Easy vector system (Promega). Ligation mixtures were used to transform competent *E. coli* (JM109) cells. Plasmid DNA was isolated from candidate clones and purified. Sense and antisense strands of the clones were used in cycle sequencing using the dideoxy-DNA chain-termination method with the BigDye Terminator Cycle Sequencing Reaction kit (Perkin-Elmer) on the ABI-3100 Prism Automated sequencer (Perkin Elmer).

Sequence identity and functional annotation

The sequence identities were obtained after BLAST searching and alignment to other published sequences in GenBank (Altschul et al., 1997). Functions were assigned to ESTs 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. Sequences that produced hits to proteins with E values greater than 10^{-5} were also considered to have an unknown function. Sequences with hits to proteins with no discernable function were placed into the miscellaneous category. Sequences with hits to plant defense (pest and pathogen) were placed into the Secondary metabolism category. The remaining sequences were placed into five broad functional categories: protein synthesis and modification, metabolism, regulatory, structural and genes of unknown function (miscellaneous).

RESULTS

We constructed cDNA libraries from Russian wheat aphid infested wheat leaves at the 3-4-leaf growth stage. The average titer of the cDNA libraries collectively were approximately 2 x 10^6 CFU, and with the average cDNA insert size of approximately 1kB. Following a single-pass, 5'-end sequencing approach, we obtained a total of 207 ESTs with sizes that ranged from 230 to 772 bases, and an average size of 489 bp.

To assign function to the proteins encoded by nonredundant sequences, the DNA sequences were translated into their corresponding amino acid sequences and searched against the nonredundant GenBank protein database using the BLASTX algorithm. A maximum probability threshold for a sequence match was set at 10^{-5}. Following this approach we obtained a total of 194

ESTs with significant E-values already present in GenBank (Table 1).

After the sequence identities were obtained from GenBank, functions were assigned based on the results returned after BLAST searching of the obtained ESTs (Figure 1). The annotated functions comprise of 25% of sequences involved in protein synthesis and modification, such as the translation factors, tRNA ligases, protein kinases and hydrolases; 25% of the sequences were involved in structural functions, such as membrane-bound and cytoskeleton proteins; 22% of the sequences were involve in the general metabolic activities required for energy production. Only 3.5% of the obtained sequences represented hits with regulatory function. Of the obtained sequences, 6.5 % failed to give a significant hit with any known protein function and thus represent the miscellanous portion. Following this approach we obtained 18% sequences with functions assigned to the secondary metabolism, and most of these had significant hits to either specific resistance gene analogs or putative RGAs.

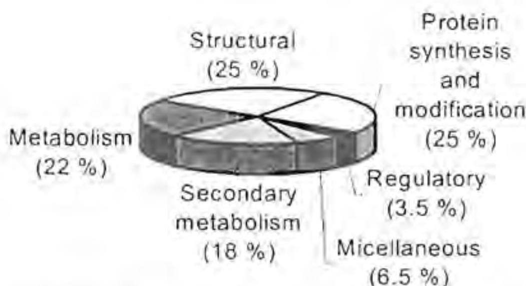

Figure 1. Percentage of nonredundant sequences grouped as genes of unknown function and genes classified into functional groups. Protein synthesis and modification: translation factors, tRNA ligases, protein kinases and hydrolases; Metabolism: proteins with a defined metabolic function like those involved in energy, redox, lipid, or carbohydrate metabolism; Structural: membrane-bound, cytoskeleton, and ribosomal proteins; Regulatory: kinases, transcription factors and proteins involve in cell cycle control; Secondary metabolism: pathogenesis-related proteins; Miscellaneous: proteins with no discernable function. Expressed sequence tags (ESTs) that did not produce a BLASTX hit, or with hits with E-values greater than 10^{-5}, were considered to have an unknown function.

The obtained RGAs were grouped accordingly to the main resistance gene classes (Table 2), and represent the major groups of resistance resistance genes, which include the serien/threonine kinases (2), transmembrane receptors (2); leucine-rich repeats (2); nucleotide binding sites (10) and leucine zippers (2). No hits were obtained that fall within the grouping of toll/interleukin-1. A further 18 sequences gave significant hits with functions either defined as putative resistance proteins or proteins with known linkages to pathogen resistance, but which does not fall within the assigned groupings.

Table 2. Expressed sequence tags (ESTs) that produced BLASTX hits with significance to resistance (R) genes.

Structural domain classes	E-value	No. of dbESTs hits
Serine/threonine kinases		
Serine/threonine kinase protein	1.00E-18	2
Transmembrane receptor		
Receptor-like kinase	6.00E-07*	2
Leucine-Rich Repeats		
LRR19	6.00E-65	2
Nucleotide binding sites		
NBS-LRR resistance protein candidate	7.00E-13	1
NBS-LRR type protein (r15) gene	6.00E-13	1
Putative NBS-LLR type resistance protein	1.00E-55	1
Nucleotide binding site LRR protein-1	4.00E-22	1
Nucleotide-binding leucine-rich-repeat protein 1	1.00E-12	1
Nucleotide-binding leucine-rich-repeat protein-like	4.00E-14	2
Resistance gene candidate CC-NBS-LLR Class	1.00E-33	1
Disease resistance complex protein NBS-LRR class	1.00E-43	1
Putative disease resistance protein CC-NBS Class	1.00E-43	1
Resistance complex protein I2-C-2	7.00E-13	1
Toll/interleukin-1		
Leucine Zipper		
Leucine-rich protein	1.00E-55	2
Other[a]		
Resistance gene analogue2	7.00E-13	6
WIR pathogen R gene	1.00E-125	3
RGA link to resistance loci in rice	4.00E-16	1
Putative resistance protein(RGA-2)	4.00E-16	1
Putative disease resistance protein	1.00E-55	1
Putative RGA PIC23	3.00E-16	1
Resistance protein candidate RGC2A pseudogene	1.00E-14	1
Polymyxin ß-resistance protein	1.00E-04*	1
Thioredoxcin	1.00E-66	1
PRM1 homolog	1.00E-43	1

[a] Structural class is not well defined
* Protein with discernable function

DISCUSSION

The majority of plant disease resistance genes cloned so far contain nucleotide-binding sites (NBS) and leucine-rich repeat (LRR) domains. This class of R genes belongs to a superfamily that is present in both dicotyledons and monocotyledons as suggested from

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

sequence comparisons made between these isolated genes (Bent et al., 1994; Lagudah et al., 1997; Meyers et al., 1998).

The use of PCR based approaches with degenerate oligonucleotide primers designed from the NBS region of cloned disease resistance genes has led to the cloning of resistance gene-like sequences in several plant species (Leister et al., 1998; Seah et al., 1998; Garcia-Mas et al., 2001). Co-segregation of some of these sequences with known disease resistance gene loci has been reported.

In the present study we tested the feasibility of using such a PCR-based approach. The degenerate oligonucleotide primers designed from conserved motifs in the NBS domain, was used to clone several disease resistance gene homologues from wheat lines. Out of the 207 ESTs obtained, 37 gave hits with significant homology to plant defense (E-values < 10^{-5}). In the present study, a clear bias for obtaining resistance gene analogs were found, when compared to other similar but randomized studies (Kruger et al., 2002; White et al., 2000; Yamamoto and Sasaki, 2000). In a similar study, where the expressed genes from *Fusarium graminearum* infected wheat spikes were analyzed, most of the obtained nonredundant ESTs were of miscellaneous nature, followed by sequences related to general metabolism and of importance to cell structure (Kruger et al., 2002).

The NBS and LRR domains are conserved amongst several disease resistance genes and this has led to the hypothesis of cloning additional resistance genes based on the homology to these conserved sequences. The procedure can be complicated by an excess of genes that contain the NBS region, but are not related to resistance genes (Yu et al., 1996). This is also true for this study, as only 8% of the RGAs could be linked to specific resistant genes, and 50% could be assigned to specific groupings, whereas the others contained only the specific conserved motif. Also many homologous resistance genes may be located throughout the genome in a plant species. Thus, the sequence homology among these genetically independent and functionally distinct disease-resistance genes will present a difficulty in isolating individual clones, which correspond to a specific resistance gene by hybridization. However, it proved useful in the present study, as these isolated clones will be utilized in a gene expression study approach in a future study.

ACKNOWLEDGEMENTS

We wish to express our sincere gratitude to the Winter Cereal Trust, National Research Foundation (NRF), Technology and Human Resources for Industry Programme (THRIP) of the NRF and Department of Trade and Industry (DTI) for financial support, and the University of Pretoria for the financial support and provision of infrastructure.

REFERENCES

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucl. Acids Res. 25: 3389-3402.

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

Bent AF, Kunkel BN, Dahlbeck D, Brown KL, Schmidt RL, Giraudat J, Leung JL, Staskawicz BJ (1994). *RPS2* of *Arabidopsis thaliana*: A leucine-rich repeat class of plant disease resistance genes. Science 265: 1856-1860.

Botha AM, Nagel MAC, van der Westhuizen AJ, Botha FC (1998). Chitinase isoenzymes in near-isogenic wheat lines challenged with Russian wheat aphid, exogenous ethylene and mechanical wounding. Bot. Bull. Acad. Sin. 39:99-106.

Bryce B (1994). The Russian wheat aphid: Fact sheet. Centre for Environmental; and Regulatory Information Services, http://www.ceris.pudue.edu/napis/pest/rwa/facts.txt.

Cannon SB, Zhu H, Baumgarten AM, Spangler R, May G, Cook DR, Young ND (2002). Diversity, distribution, and ancient taxonomic relationships within the TIR and non-TIR NBS-LRR resistance gene subfamilies. J. Mol. Evol. 54: 548-562.

Chomczynski P, Sacchi N (1987). Single-step method of RNA isolation by acid guanidium thiocyanate-phenol-chloroform extraction. Anal. Biochem. 162: 156-159.

Deng Z, Huang S, Ling P, Chen C, Yu C, Weber CA, Moore GA, Gmitter Jr. FG (2000). Cloning and characterization of NBS-LRR class resistance-gene candidate sequences in citrus. Theor. Appl. Genet. 101: 814-822.

Dixon MS, Hatzixanthis K, Jones DA, Harrison K, Jones JDG (1998). The tomato *Cf-5* disease resistance gene and six homologs show pronounced allelic variation in leucine-rich repeat copy number. Plant Cell 10: 1915-1925.

Ellis J, Dodds P, Pryor T (2000). Structure, function and evolution of plant disease resistance genes. Curr. Opin. Plant Biol. 3: 278-284.

Garcia-Mas J, Van Leeuwen H, Monfort A, De Vincente MC, Puigdomènech P, Arús P (2001). Cloning and mapping of resistance gene homologues in melon. Plant Sci. 161: 165-172.

Gyorgyey J, Vaubert D, Jimenez-Zurdo JI, Charon C, Troussard L, Kondorosi A, Kondorosi E (2000). Analysis of *Medicago truncatula* nodule expressed sequence tags. Mol. Plant-Microbe Interact. 13: 62-71.

Jackson AO, Taylor CB (1996). Plant-microbe interactions: Life and death at the interface. Plant Cell 8: 1651-1668.

Jones DA, Jones JDG (1997). The role of leucine-rich repeat proteins in plant defenses. Advances in Botanical Research 24: 89-167.

Jones DA, Thomas CM, Hammond-Kosack KE, Balint-Kurti PJ, Jones JDG (1994). Isolation of the tomato *Cf-9* gene for resistance to *Cladosporum fulvum* by transposon tagging. Science 266: 789-793.

Kruger WM, Pritsch C, Chao SM, Muehlbauer GJ (2002). Functional and comparative bioinformatic analysis of expressed genes from wheat spikes infected with *Fusarium graminearum*. Mol. Plant-Microbe Interact. 15: 445-455.

Lagudah ES, 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. Nat. Genet. 14: 421-429.

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

Martin GB, Brommonschenkel SH, Chunwongse J, Frary A, Ganal MW, Spivey R, Wu T, Earle ED, Tanksley SD (1993). Map-based cloning of a protein kinase gene conferring disease resistance in tomato. Science 262: 1432-1436.

Mekhedov S, de Ilarduya OM, Ohlrogge J (2000). Towards a functional catalog of the plant genome. A survey of genes for lipid biosynthesis.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Plant Physiol. 122: 389-401.

Meyers BC, Chin DB, Shen KA, Sivaramakrishnan S, Lavelle OD, Zhang Z, Michelmore RW (1998). The major resistance gene cluster in lettuce is highly duplicated and spans several megabases. Plant Cell 10: 1817-1832.

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.

Ritter C, Dangl JL (1996). Interference between two specific pathogen recognition events mediated by distinct plant disease resistance genes. Plant Cell 8: 251-257.

Sambrook J, Fritsch EF, Maniatis T (1989). Molecular cloning: A laboratory manual, Second Edition. Cold Spring Harbour Laboratory, NY.

Sandhu D, Gill KS (2002). Gene-containing regions of wheat and the other grass genomes. Plant Physiol. 128: 803-811.

Schenk PM, Kazan K, Wilson I, Anderson JP, Richmond T, Somerville SC, Manners J (2000). Coordinated plant defense responses in Arabidopsis revealed by microarray analysis. Proc. Natl. Acad. Sci. USA 97: 11655-11660.

Seah S, Sivasithamparam K, Karakousis A, Lagudah ES (1998). Cloning and characterization of a family of disease resistance gene analogs from wheat and barley. Theor. Appl. Genet. 97: 937-945.

Tolmay VL, Van der Westhuizen MC, Van Deventer CS (1999). A six week screening method for mechanisms of host plant resistance to *Diuraphis noxia* in wheat accessions. Euphytica 107: 79—89.

Unger LM, Quisenbury SS (1997). Effect of antibiotic plant resistance on the reproductive fitness of the Russian wheat aphid (Homoptera: Aphididae). J. Econ. Entomol. 90: 1697-1701.

Van der Westhuizen AJ, Botha FC (1993). Effect of the Russian wheat aphid on the composition and synthesis of water soluble proteins in resistant and susceptible wheat. J. Agron. Crop Sci. 170: 322-326.

Van der Westhuizen AJ, Pretorius Z (1996). Protein composition of wheat apoplastic fluid and resistance to the Russian wheat aphid. Austrl. J. Plant Physiol. 23: 645-648.

Van der Westhuizen AJ, 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 Rep. 8: 132-137.

Van der Westhuizen AJ, Qian X-M, Botha A-M (1998b). ß-1,3-glucanases in wheat and resistance to the Russian wheat aphid. Physiol. Plant. 103: 125-131.

Van der Westhuizen AJ, Qian X-M, Wilding M, Botha A-M (2002). Purification and immunocytochemical localization of a wheat ß-1,3-glucanase induced by Russian wheat aphid infestation. SA J. Science 98: 197-202.

Walters MC, Penn F, Du Toit F, Botha TC, Aalbersberg K, Hewitt PH, Broodryk SW (1980). The Russian wheat aphid. *Farming in South Africa*. Wheat: Winter Rainfall G.3/1980.

Whitham S, Dinesh-Kumar SP, 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.

White JA, Todd J, Newman T, Focks N, Girke T, de Ilarduya OM, Jaworski JD, Ohlrogge JB, Benning C (2000). A new set of *Arabidopsis* expressed sequence tags from developing seeds. The metabolic pathway from carbohydrates to seed oil. Plant Physiol. 124: 1582-1594.

Yamamoto K, Sasaki T (2000). Large-scale EST sequencing in rice. Plant Mol. Biol. 35: 1350-144.

Yu YG, Buss GR, Maroof MAS (1996). Isolation of a superfamily of candidate disease-resistance genes in soybean based on a conserved nucleotide-binding site. Proc. Natl. Acad. Sci. USA 93: 11751-11756.

APPENDIX 3

Published in Proceedings of the 10th international wheat genetics symposium, Paestum, Italy. 1-6 September 2003. Vol. 1 p. 1281-1283

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

USING SUPPRESSION SUBTRACTIVE HYBRIDISATION (SSH) TO SCREEN FOR NOVEL SEQUENCES EXPRESSED IN RESPONSE TO RUSSIAN WHEAT APHID FEEDING.

Chantal van Niekerk & Anna-Maria Botha
Department of Genetics, Forestry and Agriculture Biotechnology Institute, University of Pretoria, Hillcrest, Pretoria, ZA0002, South Africa. (Phone: +27 420 3945; Fax: +27 420 3947; E-mail: ambothao@postino.up.ac.za).

ABSTRACT

The Russian wheat aphid (*Diuraphis noxia,* Mordvilko; RWA) causes serious damage to wheat crops worldwide. Breeding efforts have led to the development of wheat lines (e.g. 'Tugela DN' (*Dn1,* SA1684 / Tugela*5) that are resistant to RWA infestation. By making use of 'Tugela DN' and suppression subtractive hybridisation (SSH), fragments were isolated that may be involved in the wheat plant's resistance to RWA infestation. The SSH fragments were cloned into pGEM-T Easy Vector and all clones sequenced. Thereafter, the sequences were subjected to a GenBank database search using the BLASTX algorithm. Interestingly, all obtained sequences showed no significant homology (e-value $< 10^{-5}$) with any known protein. However, Real-Time PCR and Northern blot analysis confirmed involvement in the RWA resistance response through up-regulation from 5-fold to 5.4-fold of selected SSH fragments.

INTRODUCTION

Suppression subtractive hybridization (SSH) is the ideal method for the detection of low copy number, differentially expressed genes (Birch *et al.*, 2000). For SSH, two cDNA populations are needed, one in which the specific expressed 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. 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).

In this study we applied SSH to isolate unique expressed sequence tags (ESTs) that were being expressed in 'Tugela DN' plants after Russian wheat aphid (*Diuraphis noxia,* Mordvilko; RWA) infestation. By using SSH, we could compare two cDNA populations with one another and isolate any uniquely expressed sequences. 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, 1989) 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.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

MATERIALS AND METHODS

Two wheat lines were used, 'Tugela DN' (resistant line) and 'Tugela' (susceptible line), to synthesize cDNA for the SSH. Two separate subtractive hybridizations were done using the Clontech (USA) PCR-Select cDNA subtraction kit. Firstly, 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). A second, independent hybridization (SSHb) was done using 'Tugela DN' for both the tester and driver cDNA. In SSHb, RWA infested material was used for the tester cDNA and uninfested material was used for driver cDNA. All subsequent sequences were cloned (pGem-T Easy vector kit), sequenced and submitted to GenBank using the BLASTX algorithm for functional annotation (Altschul *et al.*, 1997). Resulting sequences of interest were subjected to Southern and Northern hybridization as well as Real-Time PCR to determine origin and expression patterns. Genomic 'Tugela DN' DNA for Southern blot analysis was extracted (Edwards *et al.*, 1991) using leaf tissue.

RESULTS

Seven bands were obtained from SSHa and SSHb. These bands were cloned, sequenced, and annotated in GenBank using the BLASTX algorithm. All the sequences showed no significant homology to any known proteins. Five sequences, ABO 00010, ABO 00011, ABO 00013, ABO 00014 & ABO 00027 were selected and labeled as probes for Northern and Southern blots (Fig. 1).

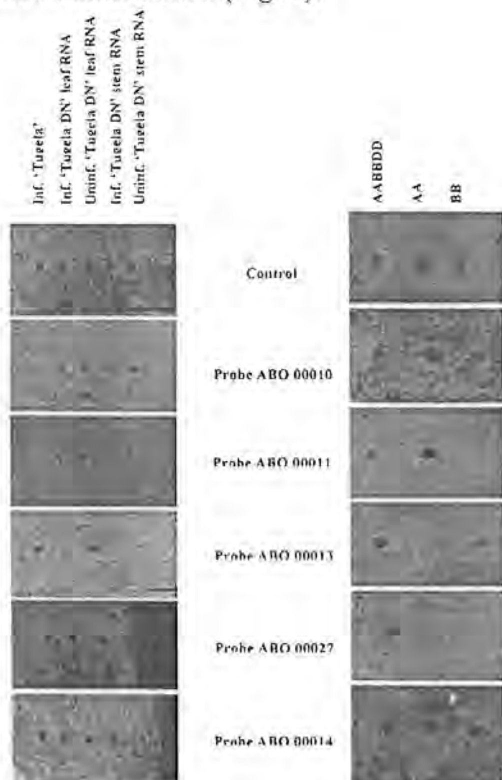

Figure 1. RNA (a) and DNA (b) dot blot analysis utilising cross hybridisation of the probes to different wheat tissue types. Control = ubiquitin probe, probe ABO 00010 & ABO 00011 = sequences identified during the SSHa subtraction, probe ABO 00013, ABO 00027 & ABO 00014 = sequences identified during the SSHb subtraction. RNA 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. DNA samples were obtained from the 'Tugela DN', *Triticum urartu* and *Aegilops speltoides*, representing theAABBDD genome, the AA genome, and the BB genome, respectively.

(a) (b)

Primers were designed from the sequences of ABO 00013, ABO 00027 and ABO 00014 and Real-Time PCR was done to determine up- or down-regulation (Fig. 2).

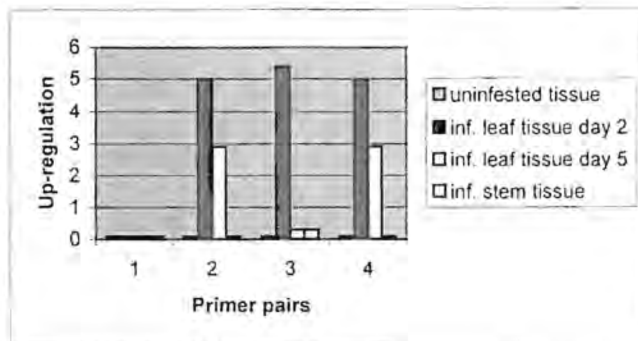

Fig. 2. Histogram illustrating up-regulation of SSH ESTs obtained after Real-Time PCR. Primer pair 1 = ubiquitin control; Primer pair 2 = EST ABO 00013 (SSHb); Primer pair 3 = EST ABO 00027 (SSHb) ; Primer pair 4 = EST ABO 00014 (SSHb). Leaf and stem tissue from RWA infested material was used. Leaf tissue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

DISCUSSION

Southern hybridisations confirmed that the ESTs obtained during SSHa and SSHb originate from the wheat genome and is not due to contamination (Janke *et al.*, 2001)(Fig.1b). Probes ABO 00013 and ABO 00027 obtained from SSHb had higher cross-hybridized mainly to the AABBDD genome ('Tugela DN'), thus implying hybridisation potentially to the DD genome, that carries the *Dn1* resistance gene (Du Toit, 1989). The northern hybridisations were done to determine up- or down-regulation of the identified ESTs (Fig.1a), whilst Real time PCR was used to quantify the level of regulation (Fig. 2). Probes ABO 00010 and ABO 00011 (SSHa) hybridised to leaf RNA and stem RNA. No preference to infested or uninfested RNA could be determined. Probe ABO 00013 (SSHb) showed a higher hybridisation to uninfested leaf RNA and almost no hybridisation to stem RNA. Probes ABO 00027 and ABO 00014 (SSHb) hybridised equally well to all samples. Real-Time PCR showed relative quantification of ESTs ABO 00013, ABO 00027 and ABO 00014. All three ESTs were up-regulated up to 5.4 times in RWA infested tissue on day 2 and showed less up-regulation on day 5 (Fig. 2). The results obtained demonstrated the efficacy of the SSH method to screen a big genome such as that of wheat for single or low copy-number genes.

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.

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.

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.

Edwards K, Johnstone C and Thompson C. 1991. A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. Nucleic Acids Research 19: 1349.

Janke B., Dobrindt U., Hacker J. & Blum-Oehler G. 2001. A subtractive hybridisation 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 Microbiology Letter 199: 61-66.

Van Niekerk C. & Botha A-M. 2002. Isolation and characterization of cDNA sequences from Russian wheat aphid induced 'Tugela DN' (*Dn1*) libraries (GenBank accession numbers BU808657-60 & CA407984. www.ncbi.nlm.nih.gov/BLAST).

Acknowledgements The authors express their sincere gratitude to the UP, National Research Foundation (NRF), Winter Cereal Trust, and Technology and Human Resources for Industry Programme (THRIP) of South Africa for financial support.

APPENDIX 4

Published in Proceedings of the 10th international wheat genetics symposium, Paestum, Italy. 1-6 September 2003. Vol. 1 p. 334-338

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

GENE EXPRESSION PROFILING DURING *DIURAPHIS NOXIA* INFESTATION OF *TRITICUM AESTIVUM* CV. 'TUGELA DN' USING MICROARRAYS.

Botha AM[1,3]., L. Lacock[1,3], C. van Niekerk[1,3], M.T. Matsioloko[1,3], F.B. du Preez[1,3], A.A. Myburg[1,3], K. Kunert[2,3] and C.A. Cullis[4]
1Department of Genetics, 2Department of Botany, 3Forestry and Agriculture Biotechnology Institute, University of Pretoria, Hillcrest, Pretoria, ZA0002, South Africa. 4 Department of Biology, Case Western Reserve University, Cleveland, Ohio, USA.*Email: ambothao@postino.up.ac.za; Phone: +27 12 420 2935; Fax: +27 12 420 3947

Introduction

Diuraphis noxia (Mordvilko, Russian wheat aphid, RWA) are probing and sucking insects, which feed on wheat, barley and other triticale (Bryce, 1994). In a study on feeding behaviour of RWA it was confirmed that the RWA probed more and fed less on resistant cultivars, resulting in the formation of more lesions on the resistant cultivars (Bahlmann *et al*., 2003). Infestation by the RWA results in altered protein expression patterns, which is manifested as differential expression of total proteins, and specific pathogenesis-related proteins (PR) like chitinases, ß-1,3-glucanases and peroxidases (Botha *et al*., 1998; Van der Westhuizen *et al*., 1998a,b). Feeding by RWA also wounds the tissue. Results on chitinase profiling (IWF and Western-blot analysis) indicated that there are distinct differences between the obtained isoenzymes and chitinase subunits after RWA infestation, wounding and exogenous ethylene treatments (Botha *et al*., 1998). Studies conducted on the intercellular washing fluid of wheat (*Triticum aestivum* L.) resistant to the Russian wheat aphid ('Tugela DN'), showed that proteins were induced within six days of infestation. The response is induced within the first 12h after infestation by the RWA in resistant cultivar 'Tugela DN', but not in the susceptible near isogenic line 'Tugela' (Botha *et al*., 1998; Van der Westhuizen *et al*., 1998a,b).

Recently, cloning of multiple R genes from various plant species has revealed conserved domains at the amino acid level. The most notable being the presence of nucleotide binding sites (NBS) and leucine rich repeat regions (LRR). The presence of a NBS and a LRR is consistent with the protein products playing a significant role in signal transduction and having a putative role in protein-protein interactions (Whitham *et al*., 1994; Jackson and Taylor, 1996; Pan *et al*., 2000; Cannon *et al*., 2002). In the comparative analysis using infested material containing the RWA resistance genes *Dn1*, *Dn2* and *Dn5*, and utilizing degenerate primer sets designed from the consensus NBS motif from other genome studies (e.g. *Arabidopsis* and rice), suppression subtractive hybridization (SSH), RACE-PCR and cDNA libraries, several NBS-LRR sequences were obtained (Lacock *et al*., 2003). The feasibility of using the degenerate PCR-based approach was tested, and it was found that 18% of all the obtained ESTs showed significant homology to resistance genes from other plants on amino acid level (E value $< 10^{-5}$) rendering the approach highly feasible if resistance gene analogs are the target of interest.

In this study, we have analyzed the expression profiles of selected gene sequences obtained from our previous study (Lacock *et al*., 2003), as well as from other studies. To follow the expression profiles of these gene sequences, we hybridized the microarray against cDNA synthesized from leaf tissue of RWA

resistant cultivar 'Tugela DN' pre- (day 0) and post-infestation (days 2, 5 and 8), in an effort to identify gene sequences with significance to the RWA defense response.

Material and Methods

Russian wheat aphid resistant cultivar 'Tugela DN' (Tugela*5/SA1684, *Dn1*) was infested with RWA (10 aphids per plant) at the 3-4-leaf growth stage (Botha *et al*., 1998). Total RNA isolation, purification of mRNA, cDNA synthesis, cDNA library construction and sequencing was performed as previously described (Lacock *et al*., 2003). After sequencing of clones, sequence identities were annotated through BLAST searching and alignment to other published sequences in GenBank (Altschul *et al*., 1997). Functions were assigned to dbESTs based on the results (E value $< 10^{-5}$) returned from searches using the BLASTX algorithm (Lacock *et al*., 2003).

Target cDNA for spotting on the microarray were 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). PCR products were quantified by electrophoresis on 0.8% agarose gels (w/v) and visualized by ethidium bromide staining. Microarrays were printed on a BioRobotics Generation II Arrayer according to the manufacturer's instructions. Arrays were printed on aminosaline slides and each target DNA were spotted 8 times. 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.

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

Microarrays were pre-hybridized 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 ddH2O 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 eppendorf and 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 1xSSC/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. Hybridization was repeated over biological material, made use of direct comparisons, and a reference, 2x2 factorial design (Yang and Speed, 2002).

An Axon GenePix 4000 A 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

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

for further analysis. Background fluorescence values were automatically calculated by the GenePix program and subtracted from all feature intensities, before further calculations were performed. Genes of interest were identified by computational analysis using 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.*, 2000).

Results and Discussion

The custom-designed microarrays used in this analysis were screened with cDNA synthesized from mRNA samples derived from wheat leaves at the three to four leaf stage, pre- (day 0) and post-infestation (days 2, 5 and 8). Previous studies on RWA feeding induced responses, indicated the induction of PR-proteins and other defense related proteins, e.g. chitinases, peroxidases, ß-1,3-glucanases (Botha *et al.*, 1998; Van der Westhuizen *et al.*, 1998a,b), lipoxigenase, ROS (Mohase and van der Westhuizen, 2002) 3 to 12 days post-infestation. Two hundred and fifty-six wheat NBS-LRR and other sequences were obtained using degenerate primer sets designed from the consensus NBS motif from other genome studies (e.g. *Arabidopsis* and rice), suppression subtractive hybridization (SSH), RACE-PCR and cDNA libraries (Lacock *et al.*, 2003). 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-labeled with Cy3- or Cy5-fluorescent dyes and hybridized to the arrays.

Statistical analysis of the expression data revealed the up-regulation and down-regulation of 5% of all the spotted gene fragments (Fig. 2) at a threshold log-2 expression ratio of 1.5 and $P \leq 0.05$ (Fig. 1).

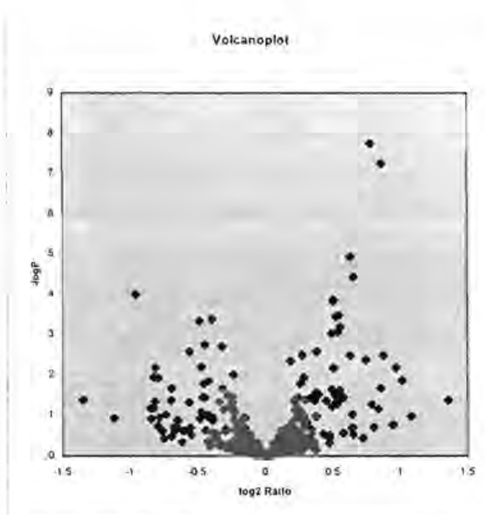

Figure 1. Significance plot for data set.

Figure 2. Hierarchical clustering of data to demonstrate regulation of genes upon RWA feeding (Days 0, 2, 5 and 8 post-infestation; Green = down-regulation; Red = up-regulation; Black = zero regulation).

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Real time quantitative PCR of selected cDNA clones to quantify the observed regulation of data.

Clone ID[a]	Day 2*	Day 5*
Unknown protein #23	+ 2.6	+ 4.4
Triticum monococcum putative resistance protein (RGA-2)	+ 1.2	- 8.0
Wheat chloroplast – ATP synthase	- 2.1	- 11.3
Unknown protein #57	- 2.3	- 3.8
BARE-1 long terminal repeat	+ 3.5	+ 5 200.0
Unknown protein #310	+ 2200.0	- 109.0
Ty1-copia-like retrotransposon for putative reverse transcriptase	+ 32.0	+ 20 000.0
Unknown protein #314	+ 77.7	+ 1.4

[a]EST (array spot #) with no significant homology (E value $<10^{-5}$) to dbEST in GenBank
*times up (+) or down(-) regulation

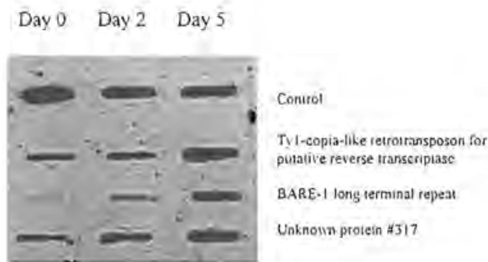

Figure 3. Northern blot analysis of cDNA clones spotted on the Microarray. Control demonstrating equal loading of cDNA.

Several unknown ESTs obtained from a previous study (sequences with no significant homology to published dbESTs in GenBank at a E value $< 10^{-5}$, Lacock *et al.*, 2003) showed significant regulation in response to RWA feeding (Fig. 2, 3, Table 1). Wheat homologs to *RGA*-2 and chloroplast-ATP synthase genes were down-regulated in response to RWA feeding on days 2 and 5 post-infestation (Table 1), whereas wheat homologs to BARE-1 long terminal repeat and Ty1-copia-like retrotransposon for putative reverse transcriptase were significantly up-regulated on days 2 and 5 post-infestation (Figure 3, Table 1).

References

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

Bahlmann, L., Govender, P. and Botha, A.-M. (2003) The leaf epicuticular wax ultrastructure and trichome presence on Russian wheat aphid (*Diuraphis noxia*) resistant and susceptible leaves. *African Entomology* **11**, 59-64.

Botha, A-M., Nagel, M.A.C., van der Westhuizen, A.J. and Botha, F.C. (1998) Chitinase isoenzymes in near-isogenic wheat lines challenged with Russian wheat aphid, exogenous ethylene and mechanical wounding. *Bot. Bul. Acad. Sinica* 39, 99-106.

Bryce, B. (1994) The Russian wheat aphid : Fact sheet. Centre for Environmental, and Regulatory Information Services, http://www.ceris.pudue.edu/napis/pest/rwa/facts.txt.

Chu, T.-M., Weir, B. and 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. and Young, N.D. (2002) Diversity, distribution, and ancient taxonomic relationships within the TIR and non-TIR NBS-LRR resistance gene subfamilies. *J. Mol. Evol.* 54, 548-562.

Eisen, M.B., Spellman, P.T., Brown, P.O. and Botstein, D. (1998) Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA* 95, 14863-8.

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

Lacock, L., van Niekerk, C., Loots, S., du Preez, F.B. and Botha, A.-M. (2003) Functional and comparative analysis of expressed sequences from *Diuraphis noxia* infested wheat obtained utilizing the Nucleotide Binding Site conserved motif. *Afr. J. Biotech.* 2, 75-81.

Mohase, L. and van der Westhuizen, A.J. (2002) Salisylic acid is involved in resistance responses in the Russian wheat aphid-wheat interaction. *J. Plant Physiol.* 159: 585-590.

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

SAS Institute Inc. SAS/STAT Software version 8 (SAS Institute, Cary, NC, 1999) http://brooks.statgen.ncsu.edu/ggibson/Pubs.htm).

Van der Westhuizen, A.J., Qian, X.-M. and 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. and Botha, A.-M. (1998b) ß-1,3-glucanases in wheat and resistance to the Russian wheat aphid. *Physiol. Plant.* **103**, 125-131.

Whitham, S., Dinesh-Kumar, S.P., Choi, D., Hehl, R., Corr, C. and 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. and Paules, R.S. (2001) Assessing gene significance from cDNA microarray expression data via mixed models. *J. Computat. Biol.* **8**, 625-637.

Wang, Y.H. and 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.

Acknowledgements The authors wish to express their sincere gratitude to Mr. Danie Theron (ACTG Microarrayer Facility, UP) for printing of microarrays and technical advice. We greatly acknowledge the Winter Cereal Trust, National Research Foundation (NRF), Technology and Human Resources for Industry Programme (THRIP) of the NRF and Department of Trade and Industry (DTI) for financial support, and the University of Pretoria for the financial support and provision of infrastructure.