The Ncm-1 gene for resistance to Cucumber mosaic virus in yellow lupin (Lupinus luteus): molecular studies and marker development.

This thesis is presented for the degree of Doctor of Philosophy

2012

Dora Agnes Li

BSc (Hons)

Supported by
Grains Research and Development Corportation

State Agricultural Biotechnology Centre
Murdoch University, Western Australia

Under the supervision of
Professor M.G.K Jones1 and Professor R. Appels2

1 State Agricultural Biotechnology Centre, Murdoch University, Western Australia
2 Centre for Comparative Genomics, Murdoch University, Western Australia
I declare that this thesis is my own account of my research and contains as its main content work which has not previously been submitted for a degree at any tertiary education institution.

Dora Agnes Li
Abstract

Cucumber mosaic virus (CMV) is an important virus pathogen of lupins in Australia which causes serious yield losses of up to 60% in epidemic years. In commercially grown lupin (Lupinus angustifolius and L. luteus) crops CMV is spread non-persistently by aphid vectors, but it can also be seed borne and this extends virus infection into successive generations. Resistance to CMV has been identified in L. luteus cv. Wodjil and is the conferred by the Ncm-1 gene. The aims of this research were to study the Ncm-1 gene in order to gain a better understanding of resistance in yellow lupin, and to develop a molecular marker linked to Ncm-1 for use in marker assisted selection.

Previously published data by Jones et al (1996) identified Ncm-1 as being a single dominant resistance gene, however, phenotypic analysis of CMV infection in a segregating L. luteus mapping population in this thesis was consistent with the Ncm-1 gene being a dominant gene modified by at least one other minor gene. The polygenic nature of CMV resistance in this genetic background was further supported by AFLP analysis which identified one major and one minor QTL associated with resistance.

A PCR based approach, using degenerate primers designed on conserved disease resistance protein motifs, was used to identify resistance gene analogues (RGA) in L. luteus. Comparative analysis revealed that RGAs isolated from L. luteus were members of the TIR-NBS-LRR class of R proteins and were similar to the TMV resistance gene N identified in tobacco and the RT4-4 CMV resistance gene from pepper. Extensive comparative analysis using the genomes of model species (including Medicago truncatula, Glycine max, Arabidopsis thaliana and Lotus japonicus) was explored and validated the assignment from L. luteus RGAs to the category of candidate gene for CMV resistance. The RGAs identified in L. luteus were found to be highly
conserved in both the CMV resistant and susceptible varieties tested. SNPs which resulted in non-synonymous mutations were identified using cDNA based 5’ RACE and used to develop a single nucleotide primer extension (SNUPE) assays for MALDI-ToF mass spectrophotometric analysis. As SNUPE is based on the allele specific extension of a single nucleotide, genotyping is highly accurate and provides co-dominant information. Two SNUPE assays were developed based on the RGAs isolated and validated on bulked samples from two *L.luteus* populations segregating for CMV resistance. One assay, SNUPE A^{267→C} was found to associate with CMV resistance. This co-dominant assay is the first of its kind reported for yellow lupin.
Acknowledgements

First and foremost, I would like to thank my supervisors, Professors Mike Jones and Rudi Appels for their support, encouragement and guidance throughout this research. I am truly grateful for the wealth of knowledge I have gained from them and for their faith in me as a student. I would especially like to thank Rudi, who always managed to shine a light into the dark recesses and find wheat among the chaff. Without your guidance I would still be looking for a torch.

To the many friends and colleagues that have helped me on this long and winding road, thank you for all your help and support. Thank you to the staff in the lupin breeding program and in the virology labs at DAFWA who provided me with not only the populations and virus stocks to get started, but also their knowledge and assistance to make some sense of it all. I would also like to thank everyone at the SABC, both in the plant lab and in the DAFWA lab who have supported me throughout this research. Thanks especially to Meredith, Marie and Steve, who were always there to give me encouragement and a nudge when required. Your friendship and support smoothed the bumps on the journey.

To my family, I am eternally grateful for your unwavering love and support. To my parents, thank you for everything, but especially for giving me the desire to learn and the opportunity to try and fulfil it. To my extended family, who always had faith and were there with ready encouragement and baby sitting, no one could ever wish for or get better in-laws. And to my husband Noel, daughter Tara and son Liam, who have made this whole journey worthwhile, this is for you!
Table of Contents

Abstract i
Acknowledgements iii
Table of Contents iv
Abbreviations xii
List of Tables xv
List of Figures xvii

Chapter 1. Literature review 1

1.1 Introduction 1

1.2 Plants viruses 1
   1.2.1 Cucumber Mosaic Virus 3
   1.2.2 Economic impact of CMV in lupins 5

1.3 Disease management 6
   1.3.1 Molecular breeding 7
   1.3.2 Marker assisted selection 7

1.4 Molecular markers : a brief overview 8
   1.4.1 Early marker development 8
   1.4.2 Genetic profiling 8
   1.4.3 PCR-based molecular markers 9
   1.4.4 Multiple Arbitrary Amplicon Profiling (MAAP) 10
   1.4.5 Amplified Fragment Length Polymorphisms (AFLP) 10
   1.4.6 Simple Sequence Repeats (SSR) 11
   1.4.7 Single Nucleotide Polymorphisms (SNP) 12
1.5 SNP genotyping

1.5.1 Cleaved Amplified Polymorphic Sequence (CAPS) markers

1.5.2 Single Nucleotide Primer Extension (SNuPE) Assay

1.5.3 Next generation technologies for high density association mapping of SNPs

1.6 Disease Resistance

1.6.1 The Hypersensitive Response (HR)

1.6.2 Gene-for-gene interaction

1.6.3 R gene structure

1.7 Conserved R protein structures and their functions

1.7.1 Leucine Rich Repeats (LRR)

1.7.2 Nucleotide binding site (NBS)

1.7.3 TIR domain

1.7.4 Non-TIR subfamily

1.7.5 Kinase domain

1.8 R gene – Avr recognition

1.8.1 R gene recognition complexes

1.8.2 Guard model for R gene interaction

1.8.3 Decoy model for R gene interaction

1.8.4 Defence signalling

1.8.5 Diversity and evolution of R genes

1.9 Plant virus resistance genes

1.9.1 N gene mediated resistance to Tobacco mosaic virus (TMV)

1.9.2 RCY1 mediated resistance to CMV in Arabidopsis

1.10 Aims of thesis
# Chapter 2. General materials and methods

2.1 Plant materials and phenotypic scoring

2.1.1 Segregating populations

2.1.2 Surface sterilisation and vernalisation of seeds

2.1.3 Plant material for virus stocks

2.1.4 Sap inoculation of host and segregating plants

2.2 Enzyme-linked immunosorbent assay (ELISA)

2.3 DNA extraction from yellow lupins

2.4 Quantification of DNA

2.5 Production of recombinants and transformation of bacterial cells

2.5.1 Ligation of PCR products into the pGEM-T Easy vector system

2.5.2 Preparation of chemically competent E. coli cells

2.5.3 Transformation of competent E. coli cells

2.5.4 Inoculation of LB broth cultures

2.6 PCR Amplification

2.6.1 General amplification conditions

2.6.2 DNA sequencing and analysis

2.7 Amplified Fragment Length Polymorphism (AFLP)

2.7.1 Restriction/ligation of target templates

2.7.2 Pre-selective PCR

2.7.3 Selective amplification

2.7.4 Post PCR multiplexing of AFLP reactions

2.7.5 Visualisation of fluorescent AFLPs using polyacrylamide gel electrophoresis
Chapter 3. Characterisation of CMV resistance and five domestication traits in two segregating crosses of *Lupinus luteus*  

3.1 Introduction  
3.2 Materials and Methods  
3.2.1 Plants  
3.2.2 Plant Inoculation  
3.2.3 Phenotypic evaluation for CMV resistance  
3.2.4 Alkaloid level  
3.2.5 Growth habit, pod shattering, seed colour and flower colour  
3.2.6 ELISA determination of virus infection  
3.3 Results  
3.3.1 Disease assessment of F2 populations  
3.3.2 Evaluation of F3 populations for CMV resistance
Chapter 4. Identification of molecular markers linked to CMV resistance. 78

4.1 Introduction 78

4.2 Materials and Methods 80
  4.2.1 DNA extraction 80
  4.2.2 Random Amplified Polymorphic DNA (RAPD) 80
  4.2.3 Amplified Fragment Length Polymorphisms 81
    4.2.3.1 AFLP reactions 81
  4.3.2.3 MAPMAKER 3.0 analysis 82
    4.3.2.4 MapManager QTX analysis 83
  4.2.4 Radio-isotope labelling of reactions 83
  4.2.5 Non denaturing polyacrylamide gel electrophoresis 84
  4.2.6 AFLP detection and fragment recovery 84

4.3 Results 85
  4.3.1 RAPD Analysis 85
  4.3.2 AFLP Analysis 86
  4.3.3 Linkage analysis 86
  4.3.4 Isolation of linked fragments 89
  4.3.5 Sequence analysis of AFLP fragments 92
    4.3.5.1 Fragment M6E1T117 92
    4.3.5.2 Fragment M7E4T211 92
    4.3.5.3 Fragment M8E4T169 93
    4.3.5.4 Fragment M4E8P368 94
Chapter 5. Isolation and characterisation of NBS-LRR resistance gene analogues in L. luteus

5.1 Introduction

5.2 Materials and Methods
  5.2.1 Primer design
  5.2.2 RGA amplification and visualisation
  5.2.3 Cloning and sequencing of RGA fragments
  5.2.4 Restriction enzyme digestion of RGA clones
  5.2.5 Sequence analysis of RGA fragments

5.3 Results
  5.3.1 RGA amplification
  5.3.2 Identification of unique RGA fragments
  5.3.3 RGA sequencing

5.4 Discussion

Chapter 6. Analysis of transcribed resistance gene analogues as candidate markers for Ncm-1

6.1 Introduction

6.2 Materials and Methods
  6.2.1 Southern hybridisation
  6.2.2 mRNA extraction
  6.2.3 Reverse transcription PCR for first strand cDNA synthesis
  6.2.4 5’ Rapid amplification of cDNA ends (5’ RACE)
6.2.5 Matrix-Assisted Laser Desorption-Ionisation Time-of-Flight (MALDI-ToF) Mass Spectrometry of Single Nucleotide Primer Extension (SNUPE) products

6.2.5.1 Product amplification and dephosphorylation
6.2.5.2 Single Nucleotid Extension
6.2.5.3 Desalting
6.2.5.4 MALDI-ToF Mass Spectrophotometric Analysis

6.3 Results

6.3.1 Sequence analysis of RGAs within parental genotypes
6.3.2 Southern hybridisation of RGA-1
6.3.3 5’ Rapid amplification of cDNA ends (5’ RACE)
6.3.4 SNUPE assays for marker development
6.3.5 SNUPE for mutation A⁴³⁰→T
6.3.6 Validation of the SNUPE assay for mutation A⁴³⁰→T
6.3.7 SNUPE for mutation A²⁶⁷→C
6.3.8 Validation of the SNUPE assay for mutation A²⁶⁷→C

6.4 Discussion

7. General discussion

7.1 Disease resistance pathway
7.2 Single gene hypothesis for CMV resistance
7.3 AFLP fragments linked to CMV resistance
7.4 Resistance gene analogues homologous to disease resistance genes
7.5 SNUPE assay development for MAS
# Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFLP</td>
<td>Amplified fragment length polymorphisms</td>
</tr>
<tr>
<td>ANGIS</td>
<td>Australian National Genomic Information Service</td>
</tr>
<tr>
<td>AP-PCR</td>
<td>Arbitrarily primed polymerase chain reaction</td>
</tr>
<tr>
<td>Avr</td>
<td>Avirulence</td>
</tr>
<tr>
<td>BAC</td>
<td>Bacterial artificial chromosome</td>
</tr>
<tr>
<td>BSA</td>
<td>Bulked segregant analysis</td>
</tr>
<tr>
<td>CAPS</td>
<td>Cleaved amplified polymorphic sequence</td>
</tr>
<tr>
<td>CC</td>
<td>Coil coiled</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complementary DNA</td>
</tr>
<tr>
<td>CMV</td>
<td>Cucumber mosaic virus</td>
</tr>
<tr>
<td>CP</td>
<td>Coat protein</td>
</tr>
<tr>
<td>DAF</td>
<td>DNA amplification fingerprinting</td>
</tr>
<tr>
<td>ddNTP</td>
<td>Dideoxynucleotide</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>EDS1</td>
<td>Enhanced disease susceptibility locus 1</td>
</tr>
<tr>
<td>eLRR</td>
<td>Extracellular leucine rich repeat</td>
</tr>
<tr>
<td>ELISA</td>
<td>Enzyme-linked immunosorbent assay</td>
</tr>
<tr>
<td>EST</td>
<td>Expressed sequence tag</td>
</tr>
<tr>
<td>ET</td>
<td>Ethylene</td>
</tr>
<tr>
<td>HR</td>
<td>Hypersensitive response</td>
</tr>
<tr>
<td>HSP</td>
<td>High scoring pair</td>
</tr>
<tr>
<td>InDel</td>
<td>Insertion or deletion</td>
</tr>
<tr>
<td>JA</td>
<td>Jasmonic acid</td>
</tr>
<tr>
<td>LB</td>
<td>Luria-Bertani</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>LRR</td>
<td>Leucine rich repeat</td>
</tr>
<tr>
<td>MAAP</td>
<td>Multiple arbitrary amplicon profiling</td>
</tr>
<tr>
<td>MALDI-TOF MS</td>
<td>Matrix-assisted laser desorption/ionisation time-of-flight mass spectrophotometry</td>
</tr>
<tr>
<td>MAMP</td>
<td>Microbe associated molecular patterns</td>
</tr>
<tr>
<td>MAPK</td>
<td>Mitogen activated protein kinases</td>
</tr>
<tr>
<td>MAS</td>
<td>Marker assisted selection</td>
</tr>
<tr>
<td>MP</td>
<td>Movement protein</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger RNA</td>
</tr>
<tr>
<td>NBS</td>
<td>Nucleotide binding site</td>
</tr>
<tr>
<td>NO</td>
<td>Nitric oxide</td>
</tr>
<tr>
<td>ORF</td>
<td>Open reading frame</td>
</tr>
<tr>
<td>PAMP</td>
<td>Pathogen associated molecular patterns</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PRR</td>
<td>Pattern recognition receptors</td>
</tr>
<tr>
<td>PTGS</td>
<td>Post-transcriptional gene silencing</td>
</tr>
<tr>
<td>QTL</td>
<td>Quantitative trait loci</td>
</tr>
<tr>
<td>R</td>
<td>Resistance</td>
</tr>
<tr>
<td>5' RACE</td>
<td>5’ Rapid amplification of cDNA ends</td>
</tr>
<tr>
<td>RAPD</td>
<td>Random amplified polymorphic DNA</td>
</tr>
<tr>
<td>RFLP</td>
<td>Restriction fragment length polymorphisms</td>
</tr>
<tr>
<td>RGA</td>
<td>Resistance gene analogue</td>
</tr>
<tr>
<td>RISC</td>
<td>Ribonucleic acid induced silencing complex</td>
</tr>
<tr>
<td>ROI</td>
<td>Reactive oxygen intermediates</td>
</tr>
<tr>
<td>RP</td>
<td>Replicase protein</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>SA</td>
<td>Salicylic acid</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>SAR</td>
<td>Systemic acquired resistance</td>
</tr>
<tr>
<td>SIPK</td>
<td>Salicylic acid induced protein kinase</td>
</tr>
<tr>
<td>siRNA</td>
<td>Small interfering ribonucleic acid</td>
</tr>
<tr>
<td>SNP</td>
<td>Single nucleotide polymorphism</td>
</tr>
<tr>
<td>SNuPE</td>
<td>Single nucleotide primer extension</td>
</tr>
<tr>
<td>SSR</td>
<td>Simple sequence repeats</td>
</tr>
<tr>
<td>TIR</td>
<td>Toll and Interleukin-1 receptor like</td>
</tr>
<tr>
<td>TM</td>
<td>Transmembrane</td>
</tr>
<tr>
<td>TMV</td>
<td>Tobacco mosaic virus</td>
</tr>
<tr>
<td>UTR</td>
<td>Untranslated region</td>
</tr>
<tr>
<td>WIPK</td>
<td>Wound inducible protein kinase</td>
</tr>
</tbody>
</table>
List of Tables

Chapter 1

Table 1.1 Virus Classification 3
Table 1.2 Classification of plant disease resistance genes 20

Chapter 3

Table 3.1 Phenotypic evaluation of F2 segregating populations for background effects on the expression of the parental phenotype 67
Table 3.2 Distribution F2 individuals of crosses P28212 X Wodjil and P28213 X Wodjil for necrotic lesions following assessment of F3 families by CMV challenge 68
Table 3.3 Phenotypic classification for cross P28213 X Wodjil 70
Table 3.4 Correlation of the % of F3 individuals, that displayed CMV lesions in the P28213 X Wodjil population, to averaged ELISA values 73

Chapter 4

Table 4.1 Primers used for selective AFLP amplification 81
Table 4.2 Scoring system used for the assessment of AFLP fragments generated for linkage analysis on MAPMAKER 82
Table 4.3 Primer combinations used for AFLP analysis. 86

Chapter 5

Table 5.1 DNA sequence of the degenerate oligonucleotide primers tested to amplify RGA in L. luteus cv. Wodjil 103
Table 5.2 RGA primer combinations, and the expected sizes of the amplified product in base pairs (bp). 105
Table 5.3 Sequences producing significant alignments to RGAs identified in L. luteus 113
Table 5.4 Origins of R genes and RGAs used for Kinase 2 tree analysis 123
Chapter 6

Table 6.1  Primers designed for SNuPE analysis  134
Table 6.2  Predicted sizes of SNuPE extension products  142
Table 6.3  *L. luteus* lines used to validate the A$^{430}$$\rightarrow$$T$ SNuPE assay and their genotype  144
Table 6.4  *L. luteus* lines used to validate the A$^{267}$$\rightarrow$$C$ SNuPE assay  148

Chapter 8

Table 8.1  ELISA values for F3 individuals from the Wodil X P28213 population.  168
Table 8.2  Distribution of F2 individuals for the Wojil X P28213 Population based on the scoring of F3 progeny following CMV challenge  173
Table 8.3  ELISA values for F3 individuals from the Wodil X P28212 population  174
Table 8.4  Distribution of F2 individuals for the Wojil X P28212 population based on the scoring of F3 progeny following CMV challenge  177
List of Figures

Chapter 1

Figure 1.1  Organisation of the CMV genome  5
Figure 1.2  Gene for gene interaction  17
Figure 1.3  A schematic representation of the predicted domains of the most prevalent classes of R genes and their location within the cell.  19
Figure 1.4  Typical structure of the NBS-LRR class of R genes  21
Figure 1.5  Comparisons of the popular models for R gene / effector interactions  27
Figure 1.6  Overview of the local signalling networks controlling activation of local defence responses  30
Figure 1.7  Immune pathways for plant pathogen interaction  35
Figure 1.8  Pathways for plant defence during infection with a typical RNA virus  39
Figure 1.9  Proposed model for N mediated resistance  43
Figure 1.10 Components known to be involved in RCY1 mediated resistance  44

Chapter 3

Figure 3.1  CMV infected *L. luteus* cv Wodjil  63
Figure 3.2  Parental phenotypes associated with infection with CMV *L. luteus*  66
Figure 3.3  Distribution of F2 individuals for the P28213 X Wodjil population based on the presence of necrotic lesions following inoculation with CMV  69
Figure 3.4  Distribution of F2 individuals for the P28212 X Wodjil population based on the presence of necrotic lesions following inoculation with CMV  69
Figure 3.5  Distribution of F2 individuals for the P28213 X Wodjil population based on the detection of viral titre following analysis using ELISA  71
Figure 3.6  Distribution of F2 individuals for the P28212 X Wodjil population based on the detection of viral titre following analysis using ELISA  71
Chapter 4

Figure 4.1  A typical RAPD PCR of *L. luteus* parental varieties 85

Figure 4.2  Polymorphic fragments segregating with CMV resistance, low alkaloid levels, the orange flower colour and the non-shattering pod phenotype 88

Figure 4.3  QTL regression analysis scan for QTL (MapManager QTX 20b) for CMV resistance as determined using the averaged ELISA values associated with P28213 X Wodjil population 89

Figure 4.4  Radiolabelled AFLP gels showing fragments isolated for sequence analysis 90

Figure 4.5  Sequence from AFLP fragment M6E1T117 originating from *L. luteus*, Wodjil 92

Figure 4.6  Sequences from AFLP fragment M7E4T211 originating from *L. luteus*, Wodjil 93

Figure 4.7  Sequences from AFLP fragment M8E4T169, originating from *L. luteus*, Wodjil 93

Figure 4.8  Sequences from AFLP fragment M4E8P368, originating from *L. luteus*, P28213 94

Chapter 5

Figure 5.1  RGA fragment amplification using degenerate primers at 47°C annealing 107

Figure 5.2  Colony screening by PCR using SP6 and T7 primers to identify RGA fragment carrying recombinants 108

Figure 5.3a  Restriction endonuclease profile, using *Dpn I*, of putative RGA containing plasmids as seen Figure 5.2 108

Figure 5.3b  Restriction endonuclease profile, using *Taq I*, of putative RGA containing plasmids as seen Figure 5.2 108

Figure 5.4  Sequence alignment of the six RGAs identified from *L. luteus* 110

Figure 5.5  Alignment of the deduced amino acid sequences of the RGAs from *L. luteus* and the kinase domain of other TIR NBS-LRR resistance gene proteins 112
Figure 5.6 Pairwise similarities calculated for the comparison of *L. luteus* Rga kinase sequences and the corresponding kinase domain in *RPS2, L6, RPP5* and *N* resistance proteins

Figure 5.7 Phylogentic tree of the deduced amino acid sequences of Lupin RGAs based on the Neighbour joining method using Mega4.0

Figure 5.8 Distribution of Rga 1 homologous sequences on the eight chromosomes of *M. truncatula*

Figure 5.9 Distribution of Rga 2 homologous sequences on the eight chromosomes of *M. truncatula*

Figure 5.10 Comparative analysis of regions homologous to Rga 1 and Rga 2 from *L. luteus*

Figure 5.11 Dendrogram of the Kinase domain proteins of known R genes, RGAs and lupin RGAs based on the Neighbour joining method using Mega4.0

Chapter 6

Figure 6.1 Comparison of the nucleotide sequences of RGA 1 in *L. luteus* variety Wodjil and accession P28212 using CLUSTALW

Figure 6.2 Comparison of the nucleotide sequences of RGA 2 in *L. luteus* variety Wodjil and accession P28212 using CLUSTALW

Figure 6.3 Southern blot of alleles of RGA1 on parental varieties Wodjil and P28212

Figure 6.4 Sequence comparison of the 5’RACE products from Wodjil and P28212 with RGA1

Figure 6.5 Alignment of the deduced amino acid sequence of the 5’ RACE products from *L. luteus* cv. Wodjil and P28212

Figure 6.7 Spectra of the SNuPE assay for genotyping the A<sup>430</sup>→T mutation observed for *L. luteus* cv Wodjil.

Figure 6.8 Mass spectra of the SNuPE assay for genotyping the A<sup>430</sup>→T mutation observed for *L. luteus* cv P28212

Figure 6.9 Mass spectra of the SNuPE assay for genotyping the A<sup>267</sup>→C mutation observed between *L. luteus* cv Wodjil and P28212

Chapter 7

Figure 7.1 Host mechanisms which may be activated following infection by CMV
| Figure 7.2 | Possible model for *Ncm-1* mediated resistance in *L. luteus* | 158 |
| Figure 7.3 | Overview of the resistance network of *Ncm-1* in *L. luteus* as compared to other closely related viral R genes | 165 |