Identification and Characterization of Common Bacterial Blight Resistance Genes in the Resistant Common Bean (*Phaseolus vulgaris*) Variety OAC Rex

By
Denise M. Cooper

A Thesis
Presented to
The University of Guelph

In partial fulfilment of requirements
for the degree of
Master of Science
in
Plant Agriculture

Guelph, Ontario, Canada
© Denise Cooper, July, 2015
Identification and Characterization of Common Bacterial Blight Resistance Genes in the Resistant Common Bean (Phaseolus vulgaris) Variety OAC Rex

Denise M. Cooper
University of Guelph, 2015

Advisor: Dr. K. Peter Pauls

Common bacterial blight (CBB), caused by the bacterium, Xanthomonas axonopodis pv. phaseoli (Xap), is a major disease of common bean (Phaseolus vulgaris L.). CBB resistance candidate genes (CGs) were identified in the genomic library sequence of resistant variety OAC Rex based on pathogen resistance gene homology. Four CGs (CBB1-4) and other genes in the surrounding sequence were characterized. The CGs had homology to a putative leucine-rich-repeat disease resistance gene in Oryza sativa and retrotransposons. CG expression was not observed in either OAC Rex or OAC Seaforth (susceptible variety) after Xap or water treatment, indicating they are unlikely to be involved in CBB resistance. The other genes encoded ascorbate oxidase, callose synthase and hypothetical proteins. To test Xap susceptibility of Arabidopsis thaliana, detached leaves were inoculated with bacteria. Chlorosis and bacterial growth were observed, suggesting it may be a useful system for testing CGs.
ACKNOWLEDGEMENTS

Over the long course of this project, I have been blessed with the help of many individuals. I would like to acknowledge the help of my advisor Dr. Peter Pauls, whose guidance lead me to get the most of my work and results. Dr. Gregory Perry was a constant and invaluable source of help in the lab, especially during the bioinformatics analyses, and my “technical difficulties”. I would also like to thank Drs. Alireza Navabi and Paul Goodwin for their assistance and feedback on my thesis.

I would like to thank all of the Pauls lab members, who helped me learn new procedures and perspectives. Our discussions improved my knowledge on many things inside and out of the lab, and made my time here truly a world class education. I must mention Jerlene Nessia, Mahbuba Siddiqua, Loo-Sar Chia, Weilong Xie and Jan Brazolot, who were my troubleshooting pals throughout my lab experiments.

I would like to thank my family and friends for their encouragement and support throughout my research and personal struggles. My husband Peter was a constant source of support in many ways during my studies, and I thank you for all your love and encouragement.

I am grateful to the various forms and sources of financial support provided by the University of Guelph, making it possible for me to complete this phase of my education.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>ACKNOWLEDGEMENTS</th>
<th>iii</th>
</tr>
</thead>
<tbody>
<tr>
<td>TABLE OF CONTENTS</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>x</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xii</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS</td>
<td>xiv</td>
</tr>
</tbody>
</table>

1.0 Literature Review ........................................... 1

1.1 Common Bean .................................................. 1

1.1.1 Nutritional Importance of Dry Bean .................. 1

1.1.2 Centres of Origin ........................................ 2

1.1.2.1 Middle American Landraces .......................... 3

1.1.2.2 Andean Cultivated Landraces ......................... 4

1.1.2.3 Common Bean Diversity and Domestication ........... 5

1.1.3 Common Bean Genetics ..................................... 6

1.2 Common Bacterial Blight of *P. vulgaris* ................ 6

1.2.1 Pathogen Spread, Infection and Symptoms in *P. vulgaris* .. 7

1.2.2 CBB Mode of Pathogenicity ............................... 8

1.2.3 Disease Prevention ....................................... 9

1.3 CBB Resistance in *P. vulgaris* ............................ 9

1.3.1 CBB Resistance Quantitative Trait Loci ............... 10

1.3.1.1 Frequently Used CBB Resistance QTL Markers ....... 11

1.3.1.2 Other Significant CBB Resistance Markers .......... 13

1.4 OAC Rex ....................................................... 15

1.4.1 Genetics of CBB Resistance in OAC Rex ................. 15

1.4.2 Major CBB Resistance QTL in OAC Rex .................. 16

1.5 Plant-Pathogen Interactions ................................ 17

1.5.1 Mechanisms of Plant Pathogen Resistance ............... 17

1.5.1.1 Pathogen Associated Molecular Pattern Triggered Immunity ................. 18
# 2.0 Identification and Characterization of Candidate Common Bacterial Blight Resistance Genes

## 2.1 Abstract

## 2.2 Introduction

- 2.2.1 CBB in Common Bean
- 2.2.2 OAC Rex Genomic Library
- 2.2.3 R Gene Classifications
- 2.2.4 R Gene Identification in OAC Rex

## 2.3 Materials and Methods

- 2.3.1 R Gene Homology Identification in OAC Rex Library Clones
- 2.3.2 Confirmation of Contig Sequences
- 2.3.3 Gene Identification and Characterization
- 2.3.4 CG and Surrounding Sequence Genome Location
  - 2.3.4.1 OAC Rex Genome Assembly Location
  - 2.3.4.2 G19833 Genome Assembly Location
    - 2.3.4.2.1 Repetitive DNA Analysis
- 2.3.5 Characterization of Predicted Proteins
  - 2.3.5.1 Domain and Secondary Structure Analysis
  - 2.3.5.2 CG-Specific Analysis
- 2.3.6 Isolation of CGs
  - 2.3.6.1 Amplification of CGs
  - 2.3.6.2 Cloning of CGs

## 2.4 Results

- 2.4.1 Gene Identification in Selected OAC Rex Contig Sequences
  - 2.4.1.1 R Gene Homology Identification
  - 2.4.1.2 Contig Sequence Assembly Confirmation
    - 2.4.1.2.1 Contig1701 Assembly Confirmation
    - 2.4.1.2.2 Contig1455 Assembly Confirmation
  - 2.4.1.3 Contig1701 Gene Annotation
  - 2.4.1.4 Contig1455 Gene Annotation
2.4.2 OAC Rex Gene Location in OAC Rex Genome Assembly .... 59
2.4.2.1 G19833 Location for Contig1701 Identified Genes 59
2.4.2.2 G19833 Location for Contig1455 Identified Genes 62
2.4.3 Contig1701 Gene and Protein Characterization ............... 67
2.4.3.1 Characterization of CGs .................................. 67
2.4.3.1.1 CG Similarities ..................................... 67
2.4.3.1.2 CG Domain Analysis ................................ 67
2.4.3.1.3 Contig Sequence Retrotransposon Analysis 68
2.4.3.1.4 CG Secondary Structure Prediction ............ 78
2.4.3.2 Characterization of Genes with Putative Functions 78
2.4.3.2.1 Ascorbate Oxidase Gene Homologs ............ 78
2.4.3.2.2 β-Glucan Synthase-Like Homologs .......... 80
2.4.3.2.3 Hypothetical Gene Annotations .......... 90
2.4.4 Isolation and Amplification of CGs ................................. 91
2.5 Discussion ....................................................................... 94
2.5.1 CG Characteristics .................................................. 94
2.5.1.1 R Gene Characteristics of CGs ............................. 94
2.5.1.2 Unique/Unusual Characteristics of CGs .............. 96
2.5.1.3 Retrotransposon Characteristics of CGs .......... 97
2.5.2 Additional CGs on Contig1701 .............................. 99
2.5.2.1 Ascorbate Oxidases and Disease Resistance ...... 99
2.5.3 CG Location in the Bean Genome ............................. 100
2.5.3.1 CBB Resistance Markers and CG G19833
Genome Location .................................................. 103
2.6 Conclusion .................................................................. 105

3.0 Detached Leaf Xanthomonas axonopodis pv. phaseoli Susceptibility

Assay in Arabidopsis thaliana ................................................. 106
3.1 Abstract ......................................................................... 106
3.2 Introduction ..................................................................... 107
3.2.1 Common Bean Improvement .................................. 107
3.2.2 CBB in A. thaliana .................................................. 107
3.3 Materials and Methods ................................................................. 109
  3.3.1 Bacterial and Plant Growth ......................................................... 109
  3.3.2 Plant Wounding and Inoculation ............................................... 109
  3.3.3 Assessment of Disease Symptoms .............................................. 111
    3.3.3.1 Image Collection and Analysis ............................................ 111
    3.3.3.2 CFU Analysis ........................................................................ 111
  3.3.4 Statistical Analyses ................................................................. 112
3.4 Results ....................................................................................... 113
  3.4.1 Analysis of Variance ................................................................. 113
  3.4.2 Visual Observations and IA ......................................................... 113
    3.4.2.1 Common Bean Visible Leaf Symptoms .................................... 113
    3.4.2.2 Arabidopsis Visible Leaf Symptoms ....................................... 115
  3.4.3 Leaf Sample CFUs ..................................................................... 118
    3.4.3.1 Common Bean CFUs ............................................................... 118
    3.4.3.2 Arabidopsis CFUs ................................................................. 120
3.5 Discussion ................................................................................... 127
  3.5.1 Symptom Development ............................................................. 127
    3.5.1.1 Common Bean Symptom Evaluation by IA .............................. 127
    3.5.1.2 Arabidopsis Symptom Evaluation by IA ................................. 129
    3.5.1.3 Symptom Evaluation by CFUs ............................................... 130
  3.5.2 Suitability of Detached Arabidopsis Leaf CBB Assay .................. 131
    3.5.2.1 Variability in CFU Observations ............................................ 131
    3.5.2.2 Detached Leaf vs. Attached Leaf Symptom Progression .......... 133
3.6 Conclusion .................................................................................. 135

4.0 Candidate Common Bacterial Blight Resistance Gene Expression in a Resistant (OAC Rex) and Susceptible (Seafort) Variety of Phaseolus vulgaris ........................................................................................................ 136
  4.1 Abstract ....................................................................................... 136
  4.2 Introduction .................................................................................. 137
    4.2.1 Significance of Common Bean ............................................... 137
4.2.2 CBB in Common Bean ................................................. 137
4.2.3 CG Testing in Common Bean ........................................ 138
4.3 Materials and Methods .................................................. 139
  4.3.1 Plant and Inoculum Preparation ................................. 139
    4.3.1.1 Plant Growth ................................................. 139
    4.3.1.2 Inoculum Preparation ...................................... 139
    4.3.1.3 Plant Inoculation .......................................... 139
  4.3.2 Sample Collection and Analysis ................................ 140
    4.3.2.1 Sample Collection and Preparation ....................... 140
    4.3.2.2 RNA Isolation .............................................. 140
    4.3.2.3 cDNA Transcription ....................................... 142
    4.3.2.4 Primer Design and PCR Amplification .................... 142
4.4 Results ........................................................................ 146
  4.4.1 CG Expression ....................................................... 146
    4.4.1.1 CBB Susceptible Genotype (Seaforth) .................... 146
    4.4.1.2 CBB Resistant Genotype (OAC Rex) ....................... 150
4.5 Discussion .................................................................... 153
  4.5.1 Observed CG Expression .......................... ........................ 153
  4.5.2 Expected CG Expression Patterns ......................... 154
4.6 Conclusion .................................................................... 156
5.0 Future Directions and Recommendations ............................ 157
  5.1 OAC Rex Genome Sequence Assembly ............................ 157
  5.2 Candidate CBB R Gene Analysis ................................. 157
    5.2.1 Identification and Characterization of New CGs .......... 157
    5.2.2 CG Expression Analysis ...................................... 159
  5.3 Model Plant CBB Inoculation Procedure ....................... 159
6.0 Summary ..................................................................... 161
7.0 References ................................................................... 164
8.0 Appendices .................................................................. 193
| Table 2.1 | DNA sequences and annealing temperatures for primers used to identify of OAC Rex genomic library clones associated with the PV-ctt001 QTL region of Pv04 | 40 |
| Table 2.2 | PV-ctt001 associated OAC Rex genomic library clones | 40 |
| Table 2.3 | NCBI accession numbers for the NBS-LRR genes used to search for candidate common bacterial blight resistance genes | 45 |
| Table 2.4 | Locations within the contig1701 sequence assembly and genetic features of contig1701 gene annotations | 58 |
| Table 2.5 | Locations within the contig1455 sequence assembly and genetic features of contig 1455 gene annotations | 60 |
| Table 2.6 | Summary of candidate common bacterial blight resistance gene characteristics | 61 |
| Table 2.7 | Predicted locations in G19833 and expression of annotated genes in OAC Rex contig1701 | 64 |
| Table 2.8 | Predicted gene locations on OAC Rex contig 1455 in bp with associated G19833 chromosome number, location (in bp), gene or element and possible function | 66 |
| Table 2.9 | Similarity between candidate common bacterial blight resistance genes | 69 |
| Table 2.10 | Phyre2 secondary structure modelling for candidate gene proteins CBB1, CBB2, CBB3 and CBB4 | 77 |
| Table 2.11 | Similarity between contig1701 genes of interest and identified homologues | 79 |
| Table 2.12 | Phyre2 secondary structure modelling for gene 5, Phvul.006G011700, Glyma.20G051900 and 1ASP_A | 87 |
| Table 2.13 | Phyre2 secondary structure modelling for gene 7, G2, and Glyma.08G361500 | 89 |
Table 2.14  Primers used to amplify candidate genes CBB2 and CBB3 from OAC Rex DNA  ......................................................... 93

Table 3.1  LSmeans for common bacterial blight symptoms on OAC Rex, Nautica and Arabidopsis leaves at 144 hours post inoculation (HPI).  117

Table 3.2  LSmeans for common bacterial blight symptoms on OAC Rex, Nautica and Arabidopsis leaves at 192 hours post inoculation (HPI).  117

Table 3.3  LSmeans for colony forming units for *X. axonopodis pv. phaseoli* (*Xap*) bacteria at 48 hours post inoculation (HPI)  122

Table 3.4  LSmeans for colony forming units for *X. axonopodis pv. phaseoli* (*Xap*) bacteria at 96 hours post inoculation (HPI)  122

Table 3.5  LSmeans for colony forming units for *X. axonopodis pv. phaseoli* (*Xap*) bacteria at 144 hours post inoculation (HPI)  123

Table 3.6  LSmeans for colony forming units for *X. axonopodis pv. phaseoli* (*Xap*) bacteria at 192 hours post inoculation (HPI)  123

Table 3.7  Inoculated *A. thaliana* leaf symptom progression  ................. 124

Table 3.8  Inoculated common bacterial blight susceptible *P. vulgaris* variety leaf symptom progression  ........................................ 125

Table 3.9  Inoculated common bacterial blight resistant *P. vulgaris* variety leaf symptom progression  ........................................ 126

Table 4.1  Candidate common bacterial blight resistance gene primer sequences for real time quantitative RT-PCR  ....................... 145
### LIST OF FIGURES

<p>| Figure 2.1 | Alignment of selected OAC Rex genomic library clones around CBB resistance marker PV-ctt001 | 43 |
| Figure 2.2 | OAC Rex contig1701 sequence confirmation, with alignments between OAC Rex sequence elements and the contig1701 and contig1455 sequence assemblies | 55 |
| Figure 2.3 | Locations, coding sequences and potential functions of predicted genes on contig1701 | 57 |
| Figure 2.4 | Location, coding sequence and potential function of predicted genes on contig1455 | 60 |
| Figure 2.5 | Candidate common bacterial blight (CBB) resistance genes | 61 |
| Figure 2.6 | Sequence comparisons between OAC Rex contig1701 and syntenic regions in G19833 | 63 |
| Figure 2.7 | Alignments of contig1455 and associated match locations within G19833 | 65 |
| Figure 2.8 | Alignment of candidate common bacterial blight resistance gene coding sequences | 72 |
| Figure 2.9 | Alignment of candidate gene protein sequences | 73 |
| Figure 2.10 | LRR sequence of candidate genes | 74 |
| Figure 2.11 | Secondary structure models of predicted protein for candidate CBB resistance genes | 76 |
| Figure 2.12 | Alignment and annotation of ascorbate oxidase coding sequence | 83 |
| Figure 2.13 | Alignment and annotation of L-ascorbate oxidase amino acid sequences | 85 |
| Figure 2.14 | Secondary structure model of OAC Rex and G19833 L-ascorbate oxidase protreins | 86 |</p>
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.15</td>
<td>Secondary structure model for predicted β glucan synthase-like genes</td>
<td>89</td>
</tr>
<tr>
<td>2.16</td>
<td>Candidate gene PCR amplification from OAC Rex genomic DNA and BiBAC2 library DNA sources</td>
<td>93</td>
</tr>
<tr>
<td>2.17</td>
<td>Common bacterial blight resistance candidate gene clones</td>
<td>94</td>
</tr>
<tr>
<td>2.18</td>
<td>Location of clone sequence match within contig1455</td>
<td>94</td>
</tr>
<tr>
<td>3.1</td>
<td>Symptom progression in inoculated <em>P. vulgaris</em> leaves</td>
<td>114</td>
</tr>
<tr>
<td>3.2</td>
<td>Symptom progression in inoculated <em>A. thaliana</em> rosette leaves</td>
<td>116</td>
</tr>
<tr>
<td>3.3</td>
<td>Pathogen growth in inoculated <em>P. vulgaris</em> leaf samples</td>
<td>119</td>
</tr>
<tr>
<td>3.4</td>
<td>Pathogen growth in inoculated <em>A. thaliana</em> leaf samples</td>
<td>121</td>
</tr>
<tr>
<td>4.1</td>
<td>Locations of RNA expression primers within candidate genes</td>
<td>144</td>
</tr>
<tr>
<td>4.2</td>
<td>Initial candidate gene expression in inoculated <em>Phaseolus vulgaris</em> leaves</td>
<td>147</td>
</tr>
<tr>
<td>4.3</td>
<td>Subsequent candidate gene expression in second experiment replication of <em>Phaseolus vulgaris</em> leaf inoculation</td>
<td>148</td>
</tr>
</tbody>
</table>
LIST OF ABBREVIATIONS

AA – Ascorbic Acid
AO – Ascorbate Oxidase
Avr – Avirulence
BAC – Bacterial Artificial Chromosome
BiBAC – Binary Bacterial Artificial Chromosome
BLAST – Basic Local Alignment Search Tool
CBB – Common Bacterial Blight
CBB1-4 – Common Bacterial Blight Resistance Candidate Genes 1-4
CC – Coiled Coil
CFU – Colony Forming Unit
CG – Candidate Gene
CGM – Candidate Gene Markers
DIRS – Dictyostelium Intermediate Repeat Sequence
EST – Expressed Sequence Transcript
ETI – Effector Triggered Immunity
eLRR – Extracellular Leucine Rich Repeat
HPI – Hours Post Innoculation
HR – Hypersensitive Response
hsp – Hypersensitive Response and Pathogenicity
IA – Image Analysis
Line – Long Interspersed Nuclear Element
LRR – Leucine-Rich Repeat
LTR – Long Terminal Repeat
MAMP – Molecular Associated Molecular Pattern
MAPK – Mitogen-Activated Protein Kinase
MAS – Marker Assisted Selection
NBS – Nucleotide Binding Site
PAMP – Pathogen Associated Molecular Pattern
PCD – Programmed Cell Death
PCR – Polymerase Chain Reaction
PG – Polygalacturonase
PGIP – Polygalacturonase Inhibiting Protein
PLE – Penelope-Like Element
PRR – Pattern Recognition Receptor
PTI – Pathogen Associated Molecular Pattern Triggered Immunity
QTL – Quantitative Trait Locus
R – Resistance
RAPD – Random Amplified Polymorphic DNA
RLCK – Receptor-Like Cytoplasmic Kinase
RLK – Receptor-Like Kinase
RLP – Receptor-Like Protein
ROS – Reactive Oxygen Species
RT – Reverse Transcriptase
SCAR – Sequence Characterized Amplified Region
SINE – Short Interspersed Nuclear Element
SNP – Single Nucleotide Polymorphism
SSR – Simple Sequence Repeat
T3SS – Type Three Secretion System
TIR – Toll-Interleukin Receptor
1.1 Common Bean

Common bean (*Phaseolus vulgaris* L.) is a member of the *Fabaceae* family, and is an important grain legume (pulse) crop species, grown in Canada and globally. It is a highly produced food legume crop in the world, with the majority of production occurring in developing countries (Gepts *et al.* 2008). There are many coloured and white varieties of edible dry bean in a range of different seed sizes. The most important bean market classes grown in Canada are navy, black, pinto, dark- and light-red kidney, great northern and cranberry beans (Bewley *et al.* 2006; Pulse Canada 2007). Dry beans produced in Canada are mainly exported to the US and Europe (Agriculture and Agri-Food Canada 2011), accounting for four times the amount consumed domestically (Goodwin 2005). The main production areas in Canada are Manitoba, Ontario and Alberta (Goodwin 2005; Pulse Canada 2015).

1.1.1 Nutritional Importance of Dry Bean

Pulses are recommended as a meat alternative in the human diet by Health Canada (Health Canada, 2014). Dry beans contain high levels of high-quality proteins (Yañez *et al.* 1995), no cholesterol, and are an excellent source of complex carbohydrates, such as dietary fibre, starch and oligosaccharides, as well as important vitamins and minerals such as folate (Tosh and Yada, 2010). A large portion of the protein component of common bean is made up of the phaseolin protein (Ma and Bliss, 1978), which has been extensively studied (Bollini and Chrispeels, 1978; Bollini *et al.*
and has been shown to be rich in lysine but poor in methionine (Bressani, 1983; Gepts and Bliss, 1984).

Despite the many beneficial nutritional aspects of incorporating dry beans into the human diet, there are several factors that contribute to reduced consumption in developed countries. The first is the presence of anti-nutritional factors such as protease and amylase inhibitors, lectins, polyphenols and oligosaccharides, whose effects may be decreased by different cooking methods, but still contribute to low digestibility (Gupta, 1987). The second is the time required to cook beans, especially when packaged dry. The long preparation time is a reason they are passed over for faster prepared foods in developed countries. However, beans are still a major protein source in some developing countries, especially in South America, and provide up to 40% of the daily protein intake in less developed nations in Africa (Gepts et al. 2008).

1.1.2 Centres of Origin

Common bean originates from Middle- and South-America, with wild common bean growing from the north of Mexico to the north of Argentina (Gepts and Debouck, 1991). Two major gene pools exist, termed Middle American and Andean, with an intermediate gene pool from the region between Peru and Ecuador, (Gepts and Bliss, 1986; Gepts, 1988; Debouck et al. 1993). Variation in the sequences of five genes from Middle American, Andean and Peru-Ecuador wild common bean accessions indicated that the centre of origin is Mexico, and that subsequent spread and domestication established the Andean and Peru-Ecuador gene pools (Bitocchi et al. 2012). This was further supported by Schmutz et al. (2014), using pooled resequencing of 30 individuals from Middle American and Andean wild genepools, using ~663,000 polymorphic sites at
least 5 kb away from genes and not in repeat sequences. This analysis, which was facilitated by the genome sequence for common bean variety G19833 (reported in the same paper), confirmed that the wild Andean gene pool originated from the wild Middle American population, indicating that it was the centre of origin for common bean. Due to further domestication and geographic separation, the two groups are distinguished not only by the location, from which they originate, but also by their agronomic and molecular traits, which are used to sub-classify them (Singh et al. 1991).

1.1.2.1 Middle American Landraces

Within the Middle American group, which is found in wild populations located from Mexico to Venezuela, there are three races; Mesoamerica, Durango and Jalisco as was described by Singh et al. (1991). The race Mesoamerica is from the lowlands and encompasses the black, navy and small red market classes. Varieties of this race grow either type II (indeterminate growth with continued vegetative growth after flowering, vegetative terminal buds and strong upright architecture) or type III (indeterminate growth with weak branches and twining and pods located primarily at the base of the plant) growth habits (Singh, 1982). Each pod produces 6 to 8 small, cylindrical, kidney or oval shaped seeds (<25g/100 seed), which can be all possible seed colours in pods that are slender and 8-15 cm long. The leaves and internode length can be small, intermediate or large, with cordate, hastate or ovate terminal leaf shape. Flowers are striped with broad cordate or lanceolate bracteoles, petals can be white, white with pink stripes or purple, and inflorescences are multi-noded. The Durango race is from the semi-arid highlands of Mexico and encompasses pinto, great northern, medium red and pink market classes. These varieties predominantly have indeterminate, prostrate, type
III growth habits. They have small to medium ovate or cordate leaflets and thin stems and branches with short internodes. Fruiting mainly occurs from the basal nodes, and the flowers produce 5 to 8 cm long flattened pods containing 4 to 5 medium sized seeds (25-40g/100seeds). Seed colour ranges from tan to yellow, cream, gray, black, white, red or pink and may have stripes or spots. Finally, the Jalisco race is located in the southern part of the Durango range and encompasses black, small red and yellow market classes. These varieties have indeterminate type IV growth habit and can reach total plant height of over 3 m in native environments. The leaves are hastate, ovate or rhombohedric, stems and branches are weak, and internodes are medium to long. Fruiting occurs along the entire length of the plant, but is concentrated in the upper part. The flowers produce 5-8 small to medium sized round or oval seeds in pods that are 8-15 cm long.

1.1.2.2 Andean Cultivated Landraces

Singh et al. (1991) also described three races that fall within the Andean centre of origin, and which are found in Peru, Chile, Bolivia and Argentina; these are called Nueva Granada, Chile and Peru. The Nueva Granada race, which encompasses the majority of large seeded and horticultural snap bean market classes, is found in a variety of altitudes and climates as well as continents. These varieties have type I (determinate growth with reproductive terminal buds and limited or no continued vegetative growth after flowering) (Singh, 1982), II and III growth habits, with medium to long internodes and large hastate, ovate or rhombohedric leaves with dense hairs. The flowers have small to medium bracteoles that are ovate, lanceolate or triangular, and produce large pods (10-20 cm) with 4-6 seeds. The seeds are medium (25-40g/100
seeds) to large (>40g/100 seeds) and come in a variety of seed colours. Varieties within the Chile race are found in the southern Andes and have indeterminate type III growth habits, with small or medium sized hastate, rhombohedric or ovate leaves and small stem internodes. The flowers are light pink or white with small to medium sized triangular, spatulate or ovate bracteoles, and produce medium sized pods (5-8 cm) with 3-5 round-to-oval shaped seeds, which are predominantly harvested as green, immature seeds. The final Andean race, Peru, is grown in the Andean highlands, and encompasses varieties which have either indeterminate or determinate type IV climbing growth habits (Debouck et al. 1988). These varieties have large hastate or lanceolate leaves, and long, weak internodes, which corresponds with the fact that they are generally grown in association with maize and other crops in their natural habitat (Singh et al. 1991). Fruiting occurs along the length of the plant or concentrated at the top, and produces pods that are 10-20 cm long, with large round or oval seeds.

1.1.2.3 Common Bean Diversity and Domestication

As can be seen in the summary above of the more detailed observations presented by Singh et al. (1991), *P. vulgaris* is a very diverse species, both between and within centres of origin and races. This is not only reserved to phenotype, but also extends to nutritional contents within the seeds and pods (Singh et al. 1991). Much of the diversity of observed traits can be attributed to the domestication process which occurred between ~8500--~8200 years ago for the Mesoamerica genepool and ~7000-~6300 years ago for the Andean genepool (Mamidi et al. 2011). Domestication genes that have been identified previously are involved with plant inflorescence structure, seed colour number and size, and harvestability (reviewed by Glémin and Bataillon, 2009).
Genome sequence analysis by Schmutz *et al.* (2014) identified 1,835 Mesoamerican and 748 Andean candidate domestication genes, with the number and type being variable across the subpopulation groups discussed above. Examples of the types of candidate genes identified include flowering, vernalization, photoperiod regulation, increased plant size, nitrogen metabolism and increased seed size and weight.

### 1.1.3 Common Bean Genetics

Common bean is a diploid species with $2n = 2x = 22$ chromosomes. The genome is estimated to be 600 Mb (Gepts *et al.* 2008). *P. vulgaris* is closely related to *Glycine max* (soybean), and both are *Phaseoleae* legumes. There is evidence that the soybean genome is the product of a duplication event and further fractionation and reassembly of the common ancestor for it and dry bean (McClean *et al.* 2010; Schmutz *et al.* 2014). This relationship makes the soybean genome an excellent resource for genetic comparisons and gene discovery. Even though comparisons between the two genomes have shown that common bean has evolved more rapidly than *G. max* (Schmutz *et al.* 2014), it can still be used as a reference for gene number, function and organization, and has proven to be useful for ongoing *P. vulgaris* sequencing efforts as described below.

### 1.2 Common Bacterial Blight of *P. vulgaris*

Common bacterial blight (CBB) is a significant seed borne disease of common bean, caused by the Gram-negative bacterial pathogen *Xanthomonas axonopodis* pv. *phaseoli* (*Xap*) and its fuscans variant *Xanthomonas fuscans* subsp. *fuscans* (*Xff*) (Gabriel *et al.* 1989; Vauterin *et al.* 1995; Schaad *et al.* 2006). Common bean is the primary host of CBB, although alternative hosts can be found among other members of
the Fabaceae family (Hayward, 1993), and Arabidopsis ecotype Columbia has been shown to be susceptible to manual inoculation (Perry, 2010). The disease is seen wherever beans are produced, and can reduce yield from 10-45%, depending on the environmental conditions and genotype (Saettler 1989; Gillard et al. 2009). The effect of the disease is most severe on non-resistant varieties grown in warm, humid growing conditions.

1.2.1 Pathogen Spread, Infection and Symptoms in *P. vulgaris*

The primary source of *Xap* inoculum is infected seeds, but the pathogen may also be introduced into established fields by infected crop debris or secondary hosts growing alongside fields (Vidaver 1993). Infection by *Xap* occurs through wounds, stomata, or from infected cotyledons (Zaumeyer and Thomas, 1957). The initial bacterial introduction is followed by cell proliferation and accumulation of extracellular polysaccharide in the intercellular spaces (Rudolph, 1993). Leaf symptoms begin as water-soaked areas appearing anywhere on the leaf between 4 to 10 days after inoculation, corresponding to the regions of bacterial cell proliferation. The bacteria will also enter degrading plant cells adjacent to colonies and begin to multiply, resulting in the expansion of water-soaked lesions on the leaves (Rudolph, 1993). As the spots grow, they may merge, becoming necrotic and sometimes developing a surrounding chlorotic zone (Vidaver, 1993).

With the progression of the disease, bacterial cells may enter the vascular tissue and spread throughout the plant (Rudolph, 1993). Although stem symptoms are less frequently seen, they have a similar look and progression to leaf symptoms (Vidaver, 1993). Pod symptoms develop differently however, beginning as water-soaked lesions
and becoming slightly sunken and brown or red-brown over time. Seed symptoms are only visible on light coloured seed coats, and appear as yellow to brown irregular spots. Seeds may also rot or shrivel, which greatly decreases the economic value of diseased yield (Vidaver, 1993).

1.2.2 CBB Mode of Pathogenicity

The ability of bacteria to infect a given plant, pathogenicity, involves large numbers of genes. An excellent example of this are the hypersensitive response and pathogenicity (hrp) genes, which are a conserved group of bacterial genes involved in the interaction between gram negative bacteria, such as Xanthomonas, and their hosts (Kamoun and Kado 1990; Bonas et al. 1990). These genes tend to be associated with the Type III secretion system (T3SS) in the bacterial cell wall (Van Gijsefem et al. 1995). However, they have also been implicated in the regulation of expression and secretion of proteins involved in pathogenicity directly from the bacterial cell to the plant host cell, as well as bacterial cell motility (Van Gijsefem et al. 1995; Agrios, 2005).

These secreted proteins, called effectors, have a variety of functions related to pathogenicity. Examples include the formation of motility structures such as the flagella in Pseudomonas solanacearum (Van Gijsefem et al. 1995); proteases such as prt1 and prt2, and pectin degrading polygalacturonases (PGs) which are involved in Xanthomonas campestris pv. campestris (Xcc) infection of crucifers, causing black rot (Dow et al. 1990; Wang et al. 2008a). The complement of proteins that are employed in pathogenicity are specific to the bacteria injecting them. In certain cases, host species have evolved mechanisms to recognize these effectors and use their presence to initiate resistance reactions (Agrios, 2005), as will be discussed below.
1.2.3 Disease Prevention

One of the most common practices for prevention of $Xap$ infection in affected areas is the use of disease free seed. This seed is produced in arid areas where a low frequency of disease is found and where preventative management to limit introduction of the pathogen is performed. However, in a study performed by Darrasse et al. (2007), it was found that inoculated common bean plants grown in environments that were suboptimal for $Xap$ proliferation were still colonized by the bacteria, although symptoms were not observed. This is a major concern to bean growers, since the use of disease free seed is so important for disease prevention in a production setting. Given that the presence of common bacterial blight in a common bean crop can reduce yield significantly (Saettler, 1989; Tar’an et al. 2001; Gillard et al. 2009), and that chemical application control measures provide no significant reduction in infection severity and effect (Weller and Saettler, 1976), the development and use of CBB resistant varieties of common bean has become a more practical and attractive alternative in the efforts to reduce yield loss to the disease.

1.3 CBB Resistance in $P. vulgaris$

$P. vulgaris$ has little natural resistance to CBB, however, there are currently a few registered varieties of common bean that are resistant to infection by $Xap$. Resistance has been introduced to common bean via inter-specific crosses to close relatives of $P. vulgaris$ such as tepary bean ($Phaseolus acutifolius$) (Parker, 1985) and scarlet runner bean ($Phaseolus coccineus$) (Miklas et al. 1994). $P. acutifolius$ shows differential levels of resistance to different $Xap$ isolates, which Opio et al. (1996) suggested indicates a gene-for-gene interaction between host and pathogen, with 5 gene pair
interactions providing the simplest explanation. However, inoculation of the progeny of four resistant by susceptible crosses with a single isolate (484a) indicated that resistance was controlled by one or two dominant genes for three resistant parents, and quantitative resistance for the fourth resistant parent (Urrea et al. 1999).

This wide range in the resistance responses depending on the Xap isolate used has also been observed in different resistant common bean genotypes (Opio et al. 1996; Mutlu et al. 2008), and multiple genomic regions have been shown to control the reaction (Shi et al. 2011). Resistance to CBB in common bean is characterized by a high to moderate slowing of disease symptom progression, and a lack of internal seed infection (Aggour et al. 1989). This partial resistance suggests that not all of the multiple CBB resistance-related genes were transferred to P. vulgaris through the interspecific crosses, or that the transferred genes don’t function at full capacity in the common bean background.

Current advances in developing CBB resistant common bean varieties include combining multiple sources of resistance (pyramiding). This approach may lead to higher levels of resistance and may reduce the chances that the resistance may be overcome by the pathogen. This may ultimately result in the production of varieties containing a resistance system which is more robust and would reduce the potential for development of resistance in the pathogen to plant defences.

1.3.1 CBB Resistance Quantitative Trait Loci

Resistance to common bacterial blight in P. vulgaris is a quantitative trait, meaning it segregates in a non-Mendelian fashion and involves multiple genes (St. Clair, 2010). Genes involved in the trait are affected by other genes as well as the
environment. Quantitative trait loci (QTLs) are regions in the genome containing
gene(s) for the trait of interest. To map QTLs, individuals in a population are analyzed
for genetic markers to determine the relationship to a particular phenotype, for example,
resistance to a disease. The phenotypic and genotypic data obtained from the
population as a whole is then used to determine the location of the QTL (St Clair, 2010),
and markers shown to be associated with particular QTL are then used to determine its
presence or absence in a plant for the purpose of trait selection.

Since the introduction of CBB resistance through interspecific crosses with *P. acutifolius*, many linkage maps have been developed to determine the locations of
QTLs for CBB resistance (i.e. Miklas *et al.* 2000; Tar’an *et al.* 2001; Blair *et al.* 2003;
Cordoba *et al.* 2010; Shi *et al.* 2011), which identified 22 minor and major effect CBB
resistance QTLs across the 11 linkage groups. Although these maps do not identify
specific genes associated with resistance to CBB, markers linked to QTL can be used
for selection, as they identify the regions in the genome which likely contain genes for
resistance to the disease. Markers also provide the opportunity to identify genes that
may be related to a particular QTL when the linkage map is compared to a physical map
(i.e. Perry *et al.* 2013). Thus, QTL, linkage maps and markers can, and have been
excellent tools, not only for the selection of traits related to CBB resistance in common
bean, but to study the genetic mechanisms of resistance.

**1.3.1.1 Frequently Used CBB Resistance QTL Markers**

Markers which have been frequently used for CBB resistance selection, and
account for the most variability in resistance phenotypes across multiple genetic
backgrounds are SU91 and BC420 (Pedraza *et al.* 1997; Yu *et al.* 2000b). The
resistance associated with both sequence characterized amplified region (SCAR) markers originated from the dry bean line XAN159, which resulted from an interspecific cross with *P. acutifolius* accession PI31944.

Marker SU91 has been reported to be on Pv08 (Miklas *et al.* 2006; Shi *et al.* 2011; Perry *et al.* 2013) and BC420 has been reported to be on Pv06 (Yu *et al.* 2000b; Shi *et al.* 2011). In the study by Vandemark *et al.* (2008), it was shown that both of these dominant markers are part of a recessive epistatic interaction. In this interaction, the recessive su91//su91 genotype suppresses the expression of heterozygous or dominant homozygous BC420. The most resistant phenotypes were seen when at least one dominant allele was present for both BC420 and SU91. However, since both markers are dominant, it is difficult to determine when a given genotype is homozygous for the dominant allele of either or both markers. This makes selection for genotypes that are homozygous dominant for both BC420 and SU91 difficult.

BC420 and SU91 are the most often used markers for determining the presence of resistance QTL in common bean accounting for 64% and 17% of phenotypic variation respectively (Miklas *et al.* 2006). BC420 is estimated to be 7.1 cM away from the associated resistance QTL (Yu *et al.* 2004), and the distance between SU91 and its associated resistance gene is estimated to be 1-2 cM (Xie *et al.*, personal communication). These two markers are not perfect indicators of resistance, and may result in the identification of false positives or negatives, but they are still useful tools when breeding for CBB resistance.

To improve prediction capabilities for the QTL which SU91 and BC420 are associated with, Shi *et al.* (2012) developed new markers through the use of a genomic
library available for resistant common bean variety HR45 (Liu et al. 2010). Candidate
gene markers (CGMs) were developed for genes present in bacterial artificial
chromosome (BAC) library clones associated with both SU91 and BC420 (Shi et al.
2012).

Although common bean is recalcitrant to transformation, Shi et al. (2012) were
able to determine that several of the CGMs accounted for more of the variability in
resistance phenotypes than the original markers did. This was shown by testing for the
presence of the CGM expression in resistant (Heterozygous dominant BC420 and
SU91) and susceptible (heterozygous recessive BC420 and SU91) near isogenic lines
(NILs) under disease pressure. Since the CGMs were developed by identifying genes
shown to be involved in disease resistance within the SU91 and BC420 QTL regions, it
was expected that they would provide a more specific method of selection for the trait
during breeding. This assumption was validated for SU91 in a recombinant inbred line
(Shi et al. 2012), which showed that CGMs accounted for greater phenotypic variation
both in the growth room ($R^2$ value up to 0.36) and in the field ($R^2$ value up to 0.42). One
of the CGMs was also found to be co-dominant, making identification of the
homozygous or heterozygous states easier.

1.3.1.2 Other Significant CBB Resistance Markers

Two other significant markers used for identification of CBB resistance QTL
include simple sequence repeat (SSR) marker PV-ctt001 and SCAR marker SAP6. The
PV-ctt001 marker was identified in the resistant common bean variety OAC Rex,
accounts for 42.2% of variation (Tar'an et al. 2001), and is located at the end of the
Pv04 (Yu et al. 2000a; Perry, 2010). However, subsequent studies using OAC Rex as a
resistant parent have not shown this marker to be involved in CBB resistance (Larsen, 2005; Durham et al. 2013). This suggests that PV-ctt001 may have been a false positive detection in the original study, and that other abiotic (ie temperature and humidity) and biotic (ie. common bean disease *Colletetrichum lindemuthianum*) factors may have been responsible for the increased variability accounted for by PV-ctt001 (personal communication; Alireza Navabi). The effect of the common bean disease anthracnose (*C. lindemuthianum*) is a likely possibility, since the same region of the Pv04 has been the focus of studies on a cluster of multiple resistance (R) genes and major effect QTLs for resistance to anthracnose (Geffroy et al 1999; Geffroy et al 2000; Ferrier-Cana et al. 2003; Campa et al. 2011). However, multiple resistance type genes are predicted to be present in a large cluster in the same region of Pv04 (Schmutz et al. 2014), suggesting other potential resistance interactions

SAP6 is a dominant marker for CBB resistance that was identified in great northern cultivar Montana No. 5 on Pv10 (Miklas et al. 2000) and accounts for 35% of resistance variation in an F2 population (Miklas et al. 2003). Although this marker was shown to be associated with a major effect QTL for CBB resistance, when Vandemark et al. (2009) attempted to combine SAP6 and SU91 in a single population, it was found that of the two, SU91 was primarily responsible for resistance. They discussed the possibility that the linkage between SAP6 and its associated QTL might not be strong, and that recombination was the cause of the lack of resistance conferred by the presence of the marker, as was evidenced by the presence of the marker in susceptible bean cultivar Matterhorn (Miklas et al. 2003). Although markers provide a good indication of the desired trait, the SAP6 marker highlights the need for identifying the
gene(s) that are responsible for CBB resistance. With gene-based markers there is an improved ability to select for the desired phenotype.

1.4 OAC Rex

OAC Rex was the first CBB-resistant variety of common bean registered in Canada, and is the result of a cross between the HR20-728 and MBE 7 bean lines (Michaels et al. 2006). Resistance to CBB originated from an inter-specific cross with naturally resistant tepary bean (P1440795) (Parker, 1985). OAC Rex produces white navy bean seeds and displays resistance to CBB in leaves and pods (Michaels et al. 2006). Although resistance to infection by Xap in P. acutifolius is characterized as a hypersensitive response (HR), and likely involves gene-for-gene resistance (Opio et al. 1996), OAC Rex lacks HR when infected. Instead, a decrease in disease symptom severity is observed, allowing for increased yield in resistant varieties under disease pressure when compared to non-resistant varieties under the same conditions (Tar’an et al. 2001; Gillard et al. 2009).

1.4.1 Genetics of CBB Resistance in OAC Rex

In the case of OAC-Rex and other varieties that are resistant to Xap but do not exhibit HR in the presence of the pathogen, it is likely that some, but not all of the genes for resistance were transferred from the resistant donor. Assuming this is the case, key signalling or recognition genes may be missing, which slows the resistance reaction. This is supported by the observation that disease progression and symptom development is often delayed and less severe in these resistant varieties as opposed to the small contained necrotic region associated with HR. Although complete resistance isn't present, the slowed disease progression is sufficient to protect yield compared to
susceptible varieties under disease pressure (Tar’an et al. 2001, Gillard et al. 2009). It
must also be noted that a strategy to use multiple partial resistance genes may be a
more robust system of resistance, since it may be more difficult for the pathogen to
overcome that kind of resistance.

1.4.2 Major CBB Resistance QTL in OAC Rex

QTL analysis of CBB resistance in OAC Rex performed by Tar’an et al. (2001)
indicated the involvement of at least one major R gene and two minor QTLs. The major
R gene, associated with the PV-ctt001 marker, was originally placed on the Pv05
linkage group (Tar’an et al. 2001) but the original SSR mapping study by Yu et al.
(2000a) had reported the position of Pvctt001 on Pv04. Also, subsequent studies
localized the marker on Pv04 (Perry, 2010; Perry et al. 2013). The location of this
marker at the end of the linkage group in both cases is most likely the reason for
differing reports. Since the most recent report (Perry, 2010) suggested the marker was
present on the Pv04 linkage group, the region associated with the PV-ctt001 marker on
this linkage group has been used for a preliminary candidate gene search as described
below.

Another disease resistance marker, SU91, was not initially identified in OAC Rex
as being associated with a resistance QTL (Tar’an et al. 2001). However, the SU91
sequence is present in the OAC Rex genome as confirmed by Perry et al. (2013), and
further testing has shown the marker to be associated with a major effect QTL in
progeny of OAC Rex (Durham et al. 2013). Association mapping performed by Shi et
al. (2011), which used 469 bean cultivars and breeding lines and 132 bean single
nucleotide polymorphisms (SNPs) has indicated marker SU91 is located on the Pv08
linkage group, and comparisons between the genome sequence of Andean common bean line G19833 and OAC Rex by Perry et al. (2013) provide support for this location in OAC Rex.

1.5 Plant-Pathogen Interactions

The majority of microbes are non-pathogenic towards most plant species, but in some cases, a parasitic, or disease interaction may evolve. Due to their stationary nature and lack of a circulatory immune system such as those found in mammals, plants have developed complex systems for pathogen recognition and signalling defense responses. This involves a diverse complement of genes related to pathogen resistance, which encode proteins specific to the role they are needed for. These genes may be major effect resistance genes, which have a high level of specificity and efficacy on their own, or they may be minor effect resistance genes, which act in conjunction with other minor and major effect genes to provide resistance to a particular disease, race or strain. In either case, the expression of these genes can be induced in response to an invading pathogen or constitutively expressed in the plant as a whole, or in specific locations.

1.5.1 Mechanisms of Plant Pathogen Resistance

One of the most important steps in resisting a pathogen for any organism is the recognition that there is an attacking entity present. The absence or retardation of pathogen recognition can mean the difference between disease resistance and susceptibility, as it can mean responses may not be started or may be too slow to be effective. Plants are no exception, and just as there are a variety of pathogen types with diverse strategies for plant infection, there are a variety of mechanisms for
recognizing them. These mechanisms fall into two broad categories called pathogen associated molecular pattern (PAMP) or microbe associated molecular pattern (MAMP) triggered immunity (PTI), or effector triggered immunity (ETI), and which may overlap, or complement each other (reviewed by Jones and Dangl 2006; Block et al. 2008; Padmanabhan et al. 2009; Rafiqi et al. 2009).

1.5.1.1 Pathogen Associated Molecular Pattern Triggered Immunity

The initial, more general form of recognition and defense is called PTI (Jones and Dangl 2006). This line of defense involves the recognition of molecules which are conserved among many classes of microbes (PAMPs/MAMPs), but are not necessarily involved in pathogenicity. Examples of PAMPs/MAMPs include the flagellin in bacterial flagella, or chitin which is a cell wall component specific to fungal cells (Zipfel and Felix, 2005). These molecules are recognized via pattern recognition receptors (PRRs), which are generally membrane bound proteins with recognition domains that are located extra-cellularly, and which may have an intracellular kinase domain for further signalling (Zipfel, 2008). Their role is to monitor for the presence of microbes, by recognizing the conserved PAMPs/MAMPs, whether or not it is generally pathogenic toward that species.

Once a pathogen is recognized, different intracellular responses ensue, such as signalling via MAP kinase phosphorylation cascades (Asai et al. 2002). Ultimately, a defense response occurs via the activation of defense genes. The effect of PTI on the plant may range from increased cell wall thickness, closing of stomata to prevent bacterial access, increases in hormone levels, production of reactive oxygen species production and defense gene expression (Taguchi et al. 2003; Desaki et al. 2006;
Melotto et al. 2006). PTI is an effective defense strategy for the majority of microbes that plants come into contact with, on a daily basis. However, the initial recognition of a microbe can be overcome by a virulent species through mutation of the conserved molecule, which allows it to avoid PTI and become pathogenic (Zipfel, 2008). Resistance is then dependent on the recognition of other molecules produced by the pathogen, whether it is another PAMP or a protein directly involved in pathogenicity, as will be discussed below.

1.5.1.1.1 Recognition of Flagellin: An Example of PTI

A well-studied example of PTI is the recognition and response to flagellin. This molecule is the building block molecule of the bacterial flagella, the membrane protrusion responsible for cellular motility, and is an important factor for many bacteria, whether plant pathogenic or otherwise. Although the internal parts of flagellin can be very different between bacterial species, there are well conserved regions at the ends (Murthy et al. 2003), which are the targets of recognition of PRRs (Gómez-Gómez et al. 1999; Asai et al. 2002). The N-terminal domain is the most highly conserved part of flagellin, and a synthesized peptide (flg22) from this region has been shown to induce typical defense responses, such as increased ethylene and reactive oxygen species production and induction of defence related genes (Felix et al. 1999; Gómez-Gómez et al. 1999; Asai et al. 2002).

Recognition of the flagellin molecule occurs via the PRR flagellin sensing 2 (FLS2), which is a LRR-receptor like kinase (LRR-RLK) that is expressed in flowers, stems, leaves and roots, even in the absence of flg22 (Gómez-Gómez and Boller, 2000). The FLS2 protein has an LRR domain for PAMP/MAMP recognition, a
membrane spanning domain, and a serine/threonine protein kinase domain (Gómez-Gómez and Boller, 2002). It acts in the presence of flg22 with another LRR-RLK called BRI1-associated receptor kinase 1 (BAK1) due to its association with the brassinosteroid receptor BRI1 (Li et al. 2002; Chinchilla et al. 2007). Defense responses are induced by the FLS2-BRI1 association by initiating signalling cascades via mitogen-activated protein kinases (MAPKs), resulting in the activation of defense-related genes via WRKY transcription factors (Asai et al. 2002). In Arabidopsis, downstream responses such as oxidative burst, callose deposition, ethylene production and the expression of defense-related genes FRK1, GST1, PR1 and PR5 were observed at different time points post-treatment (Asai et al. 2002)

1.5.1.2 Effector Triggered Immunity

Although PTI may be effective in responding to the presence of PAMPs/MAMPs, it can also be overcome by a successful pathogen, by either altering the conserved molecule, or by producing molecules that restrict host resistance responses such as kinase signalling and protein transport (Reviewed by Jones and Dangl, 2006; Block et al. 2008; Padmanabhan et al. 2009). These proteins that are involved with initiating and maintaining pathogenesis are called effectors. They are secreted by the pathogen into the host cell or surrounding extracellular region, and may trigger the form of resistance called ETI. Recognition of effectors generally occurs within the cell via the protein products of R genes. R genes are generally major effect genes, and are most commonly of the NBS-LRR class. ETI can involve direct or indirect recognition of the molecule. Direct recognition of the pathogen involves the interaction of the LRR domain of the R protein with the pathogen effector. Indirect recognition, which is the more
common form of pathogen recognition, is mediated by NBS-LRR proteins, and involves monitoring a host molecule for changes made by a pathogen effector protein (Jones and Dangl, 2006; Padmanabhan et al. 2009).

As with PTI, recognition of effectors leads to a series of responses resulting in ETI. This includes an oxidative burst, changes in hormone ratios, and signalling cascades via kinase activation and transcription factors, and eventually leads to a resistant response (Ryals et al. 1996; Dorey et al. 1997; Torres et al. 2005). Alternatively to PTI, ETI may lead to the more severe hypersensitive response (HR), in which infected and surrounding cells undergo apoptosis to restrict expansion of the pathogen. This method of resistance is only effective on obligate biotrophs, such as bacteria, because they require a living host cell to obtain nutrients and survive. By sacrificing the infected cell, the plant stops the progression of the infection into the surrounding healthy tissue. Conversely, if the pathogen were necrotrophic, killing cells to obtain nutrients from them, the HR would not decrease pathogen expansion since it would still be able to obtain nutrients from the dead cells (reviewed by Glazebrook, 2005).

1.5.1.2.1 Downy Mildew Resistance in Arabidopsis: An Example of ETI

An example of ETI are the Recognition of *Peronospora parasitica* genes (RPP), which confer resistance to downy mildew (*Hyaloperonospora arabidopsidis*) (*Hpa*) in *Arabidopsis* (Parker et al. 1997; Botella et al. 1998; van der Biezen et al. 2002). The RPP genes are present in two clusters, the RPP5 and RPP1 clusters, which are located on chromosomes 4 and 3 respectively (Parker et al. 1997; Botella et al. 1998; Noël et al. 1999). The functional members of these large gene clusters encode Toll interleukin
receptor (TIR) nucleotide binding site NBS-LRR class of resistance proteins (discussed below), and recognize the presence of the *Hpa* effector proteins *Arabidopsis Thaliana* Recognized (ATR) (Parker *et al.* 1997; Botella *et al.* 1998; van der Biezen *et al.* 2002; Gunn *et al.* 2002; Krasileva *et al.* 2011; Goritschnig *et al.* 2012). There is variability seen in the LRR domain sequence, however, the NBS region for RPP gene family members are highly conserved (Noël *et al.* 1999). The variability which is seen in the different RPP gene family members corresponds to the range in specificities required to recognize the different *Hpa* effectors in a gene-for-gene fashion (Botella *et al.* 1998; Rehmany *et al.* 2005; Fabro *et al.* 2011; Krasileva *et al.* 2011). In the case of RPP1, recognition of ATR1 occurs through the association of the LRR domain of RPP1 with the effector protein, and the TIR domain facilitates signalling and induction of defense genes (ie PAL, PBS2/3 and SID1/2), and hormone (ie. salicylic acid) production, which ultimately results in HR (van der Biezen *et al.* 2002; Krasileva *et al.* 2010).

### 1.5.2. Diversity in Plant-Pathogen Interactions

Throughout the interactions between the different plant species and the pathogens that affect them, similarities and trends can be seen, and classifications of the related genes and proteins can be made. However, the mechanisms for individual plant-pathogen interactions are complex, and may also be affected by several interchangeable factors. These factors include the genetic background of the plant and/or the pathogen, the environment in which the interaction is occurring, and the developmental stages of the plant and pathogen.
1.5.2.1 Genetic Variability and Pathogen Resistance

The genetic background of the plant is an important consideration, especially when dealing with highly cultivated crops. Some pathogen resistance mechanisms require the presence of multiple genes for recognition of a pathogen to produce a complete resistance reaction. Examples include the FLS2/BAK1 complex required for complete downstream resistance reaction to flagellin in Arabidopsis (Chinchilla et al. 2007), the complex formed between Pi5-1 and Pi5-2 that confer resistance to rice blast (Magnaporthe oryzae) in rice (Lee et al. 2009), and the complexes formed between Prf, Pto and Pto-homologues which confer resistance to bacterial speck disease (Pseudomonas syringae pv. tomato) in tomato (Chang et al. 2002; Mucyn et al. 2006; Gutierrez et al. 2010). In these cases, the absence of regulatory genes or genes that would normally work together to form a complex translates into a reduction in, or non-reaction to the pathogen.

There is also a complex interplay between the genetics of the plant host and the pathogen. The interaction between resistance genes in the host plant and avirulence (Avr) genes in the pathogen, combine to produce the resistance reaction on the host that we see. In general terms, when the pathogen has superior Avr expression, disease symptoms are seen, and when the plant expresses superior resistance genes, disease symptoms are either not present, or significantly reduced. Pathogens with different genetic backgrounds that are of the same species are called races, to make a distinction between the differences in pathogenic abilities. This interplay between host and pathogen genetic variability is called race-specific resistance. An excellent example of race-specific resistance is the interaction between races of anthracnose...
(Coletotrichum lindemuthianum) and common bean. There are multiple races of bean anthracnose, and multiple resistance genes, called Co, have been identified, and they provide resistance to different anthracnose races when present in different combinations (reviewed by Kelly and Vallejo, 2004; Ferreira et al. 2013).

1.5.2.2 Transposable Elements and Pathogen Resistance

In most cases, the genetic background of a plant species refers to which genes and alleles are present, and how their presence and combination affects the reaction to the presence of a pathogen. However, sometimes other genetic elements, such as transposable elements, can have an effect on gene function.

1.5.2.2.1 Classifications of Transposable Elements

Transposable elements are repetitive sequences within a genome that are able to move within the genome. There are two classifications of transposable element, which is based on the intermediate used for their movement (transposition) within the genome: Type I retrotransposons and Type II DNA transposons, which use RNA and DNA as their transposition intermediate, respectively (Finnegan, 1989). These classifications have been further divided into orders and super-families, to reflect the differences in insertion mechanisms and enzyme characteristics (Wicker et al. 2007).

Type I retrotransposons are divided into the orders long terminal repeat (LTR), Dictyostelium intermediate repeat sequence (DIRS), Penelope-like elements (PLE), long interspersed nuclear element (LINE) and short interspersed nuclear element (SINE), which are found in plants, mammals, fungi and other eukaryotes. Despite differences in their size, organization, and the number and type of encoded proteins, the main distinguishing features of retrotransposons are the presence of a reverse
transcriptase, and the production of a new copy with each replication cycle (for complete review see Wicker et al. 2007).

Type II DNA transposons are classified into two subclasses based on the number of DNA strands cut during transposition: subclass I encompasses orders TIR and crypton, which mainly use transposase to “cut and paste” from and into double stranded sites; subclass II is comprised of the helitron and maverick orders, which use a range of proteins to “copy and paste” from and into single stranded sites. Unlike the Type I retrotransposons, the main unifying feature of Type II DNA transposon subclasses are the use of the DNA intermediate for transposition (for complete review of mechanisms and proteins see Wicker et al. 2007).

1.5.2.2.2 Retrotransposons and R Gene Function

In addition to their role as a major source of highly repetitive regions, retrotransposons also have been shown to affect the function of disease resistance genes in plants. In many species, retrotransposons are found in areas where gene clusters are rapidly evolving, such as disease resistance related genes in rice (Miyao et al. 2003), lettuce (Meyers et al. 1998), barley (Marcel et al. 2007), common bean (Vallejos et al. 2006; David et al. 2009), poplar (Lescot et al. 2004), sugar beet (Kuykendall et al. 2009) and Arabidopsis (Yi and Richards, 2007). Retrotransposons have also been shown to cause transcription of neighboring disease resistance genes (Hernández-Pinzón et al. 2009), cause chimeric proteins with resistance genes (Kashkush et al. 2003; Wang and Warren, 2010), and re-functionalize inactive resistance genes (Hayashi and Yoshida, 2009).
1.5.2.3 Genetic by Environmental Interaction and Pathogen Resistance

In addition, the genetic background is affected by the environment, which can make the evaluation of resistance gene effectiveness difficult. This genetic (G) by environmental (E) interaction (G x E) is not unique to pathogen resistance, but the effect is two-fold in that the conditions present during plant growth and development affect host plant physiology and host gene expression as well as the ability of the pathogen to infect the host. Plant resistance genes have been shown to be affected by temperature (Fu et al. 2009; Jorgensen, 2012), soil nutrient composition (Jorgensen, 2012) humidity (Xiao et al. 2003; Zhou et al. 2004a) and light levels (Xiao et al. 2003), and may be affected at either high or low extremes. The typical response measured in these experiments is the lack of HR, which indicates that resistance genes with high resource and energy requirement are the primary target to conserve energy in increased abiotic stress. However, there are likely other less visibly affected resistance genes that have yet to be studied.

1.5.2.4 R Gene Expression at Different Developmental Stages

Some resistance genes are differentially expressed depending on the developmental state of the plant (Reviewed by Develey-Rivière and Galiana, 2007). Resistance has been shown to be affected by transitions between developmental stages such as *Zea mays* resistance to *Puccinia sorghi* which is dependent on the expression of adult characteristics (Abedon and Tracy, 1996). It may also be affected by the maturity of tissues or organs, such as the increased resistance to *Venturia inaequalis* with increased leaf maturity in apple (Li and Xu, 2002), or tissue rank, such as the increased resistance in rice from juvenile leaf ranks to adult ones (Koch and
Resistance has also been shown to increase over several stages once it begins to have an effect, such as resistance to *Xanthomonas oryzae pv. oryzae* in rice, which starts at the transition from the juvenile to adult stage, and increases to maximize at the full adult state (Mazzola *et al.* 1994; Century *et al.* 1999).

### 1.6 R Gene Classes

Just as resistance genes have been shown to be involved in a variety of pathogen recognition roles, there are several different classes of proteins which these genes encode. This includes the nucleotide binding site – leucine-rich repeat (NBS-LRR), extracellular leucine rich repeat (eLRR) containing proteins such as LRR-receptor like kinases (RLKs), receptor like proteins (RLPs), polygalacturonase inhibiting proteins (PGIPs), and receptor like cytoplasmic kinases (RLCKs). Many genes are also involved in peripheral roles, such as downstream signalling (i.e. MAPKs) or physical prevention of pathogenesis (i.e. callose production). Although there are multiple classes of genes and resultant proteins involved in pathogen resistance pathways in plants, the most commonly found, and most well studied group of resistance genes encode NBS-LRR proteins (Mindrinos *et al.* 1994; Ori *et al.* 1997; Wang *et al.* 1999; Meyers *et al.* 2003).

#### 1.6.1 Nucleotide Binding Site-Leucine Rich Repeat Proteins

The basic NBS-LRR protein contains two main domains: the nucleotide binding site domain, and the leucine rich repeat domain. These domains have different motifs, and play different roles in the recognition and defense pathway. The NBS domain of R proteins is involved in binding nucleotides for induction of downstream resistance processes (McHale *et al.* 2006). The LRR region is involved in recognition of a pathogen via protein-protein interaction (Jones and Jones, 1997).
1.6.1.1 NBS Domain

The NBS region of R proteins is made up of several conserved motifs, including the P-loop, kinase-2, resistance nucleotide binding site (RNBS)-B, Gly-Leu-Pro-Leu (GLPL) and MHDV subdomains (Meyers et al. 2002; Liu et al. 2007). These motifs are important for anchoring and hydrolyzing ADP molecules, which is the beginning of the signal transduction cascade for the resistance response (Tameling et al. 2002; McHale et al. 2006; Mestre and Baulcombe, 2006). The NBS site is kept inactive by other domains of the protein, or with host proteins, and is activated by the presence of a specific pathogen effector molecule either by direct interaction or by conformational changes in the host protein it is monitoring (Reviewed by Muthamilarasan and Prasad, 2013).

1.6.1.2 Role of Amino-Terminal Domain

In addition to the nucleotide binding region, an NBS-LRR protein may have other domains on the N terminal end. These extra domains are either the coiled coil (CC) or toll-interleukin receptor (TIR) motifs, and are the basis for two subfamilies of NBS-LRR proteins having different sequences and downstream pathways they affect. Both the CC and TIR domains are comprised of ~ 200 amino acids (Meyers et al. 2003; Bernoux et al. 2011). Crystallography has shown the CC domain structure to be a rod-shaped dimer of α-helices (Maekawa et al. 2011), and the TIR structure is comprised of a β-sheet connected by α-helices (Chan et al. 2010; Bernoux et al. 2011). It has been shown for both N terminal domains that self-association occurs, and is an important component for initiation of downstream signalling and ultimately the induction of HR (Moffett et al. 2002; Rairdan et al. 2008; Bernoux et al. 2011; Maekawa et al. 2011).
Bernoux et al. (2011) hypothesized that dimerization of the TIR region after effector binding, might provide a target for binding proteins that would cause downstream signalling, however, it is unclear whether this is a unique or more common action.

1.6.1.3 LRR Domain

The LRR region is involved in protein-protein interactions (Kobe and Kajava, 2001). The amino acid sequence is hyper-variable, yet maintains a basic consensus sequence of LxxLxxLxx/cnx, where L represents the location of leucine or other allopathic amino acids, and x represents the location of any amino acid (Kobe and Kajava, 2001; Kajava and Kobe, 2002; McHale et al. 2006). The leucine rich consensus sequence forms a β-strand followed by a β-turn structure, and makes up the backbone of the LRR domain (Kobe and Deisenhofer, 1993; Kobe and Kajava, 2001). The other variable amino acids comprise the exposed region, and changes in these amino acids may alter the specificity of the protein (Parniske et al. 1997; Botella et al. 1998; Hwang et al. 2000; Krasileva et al. 2010) as well as the structure (for review see Kobe and Kajava, 2001).

1.6.2 Extracellular LRR Proteins

Another class of genes that have been shown to be involved in pathogen resistance is the extracellular leucine rich repeat protein encoding genes. There are several classifications of these proteins based on their domains, whether they are membrane spanning, and if they contain a cytoplasmic domain. Their unifying feature is mainly the presence of an eLRR, as well as their localization to the cell membrane.
1.6.2.1 eLRR Domain

The extracellular leucine rich repeat domain is comprised of 20-29 repeats of the leucine rich repeat motif found in cytoplasmic genes with a variation in the leucine rich repeat motif consensus sequence (Zhang et al. 2013). The eLRR consensus sequence is xxLxxLxxLxxLxxNxLt/sGxIP where x may be any amino acid and L represents either leucine or other allopatic residues (Kobe and Kajava, 2001; Zhang et al. 2013). Like the LRR motif of cytoplasmic proteins, the eLRR has been shown to be involved in pathogen ligand binding, such as recognition of flg22 by the eLRR domain of the receptor like kinase FLS2 (Chinchilla et al. 2006).

1.6.2.2 LRR Receptor-Like Kinases

RLKs are comprised of an extracellular leucine-rich repeat domain, a membrane spanning domain, an intracellular serine/threonine kinase domain and a signal peptide (Sun et al. 2004; Zipfel et al. 2006). Two variations of RLK domain content have also been identified in various plants. The variants RLCK and RLP have also been shown to be involved in disease resistance reactions (discussed below). While the eLRR domain of RLKs have been shown to bind pathogen ligands as was discussed above, the kinase domain is involved in the direct and indirect initiation of signalling for downstream resistance responses by phosphorylation (Ellis et al. 2000; Lu et al. 2010; Zhang et al. 2013).

RLKs have been shown to be involved in resistance to fungal and bacterial pathogens both by positive and negative regulation (Reviewed by Yang et al. 2012). Some examples of isolated and characterized RLK pathogen resistance genes are
Xa26 and OsBISERK1 in rice (Sun et al. 2004; Song et al. 2008), EFR, FLS2 and BAK1 in Arabidopsis (Zipfel et al. 2006; Yang et al. 2012).

1.6.2.3 Receptor-Like Cytoplasmic Kinases

Receptor-like cytoplasmic kinases have similar domain content to RLKs, but are lacking the eLRR domain. Although there is no LRR ligand-binding domain, RLCKs have been implicated in signalling roles in disease resistance. For example, resistance to bacterial pathogens was demonstrated by the BIK1 gene which encodes an RLCK that mediates signalling post-recognition of flagellin by FLS2/BAK1 complex in Arabidopsis (Lu et al. 2010). Interestingly, the BIK1 protein was first reported to be involved in Arabidopsis resistance to fungal pathogens Botrytis cinerea and Alternaria brassicicola, and decrease resistance to the bacterial pathogen Pseudomonas syringae pv. tomato (Veronese et al. 2006). This indicates the existence of more diverse roles for proteins involved in disease resistance signalling than those involved in pathogen recognition.

1.6.2.4 Receptor-Like Proteins

RLPs have a similar structure to RLKs as they contain an eLRR and transmembrane domain and lack only the cytoplasmic kinase domain (Wang et al. 2008b). Within these two major domains exist several smaller conserved domains including a putative signal peptide, a cysteine-rich domain, the eLRR domain which consists of two LRR sequences interrupted by a non-LRR region, a spacer, an acidic domain, a single transmembrane domain and a short cytoplasmic region (Reviewed by Kruijt et al. 2005). RLPs have been implicated in resistance to fungal pathogens such as the tomato Cf-9 gene that confers resistance to Cladosporum fulvum (Jones et al.
1994) and LepR3 gene involved in resistance to *Leptosphaeria maculans* in *Brassica napus* (Larkan *et al*. 2013), as well as bacterial pathogens such as AtRLP30, which has been shown to be involved in non-host resistance in Arabidopsis to *Pseudomonas syringae pv. phaseolicola* (Ellendorff *et al*. 2008; Wang *et al*. 2008b).

### 1.6.2.5 Polygalacturonase Inhibiting Proteins

Polygalacturonase inhibiting proteins (PGIPs) are extracellular proteins that are involved in both the general PTI and specific ETI forms of pathogen resistance in a plant (reviewed by Dangl and Jones, 2001; Nüremberger *et al*. 2004; Federici *et al*. 2006). These proteins have been well studied for their role in inhibiting fungal pathogen infection in many dicot and monocot plants (reviewed by De Lorenzo *et al*. 2001; Ferrari *et al*. 2002; Federici *et al*. 2006), but have also been reported to be involved in resistance to bacterial pathogens (Huang and Allen, 2000; Hwang *et al*. 2010). PGIPs are also implicated in stress-induced and developmental roles, due to their inhibition of the pectinase polygalacturonase (PG) (for review see Protsenko *et al*. 2008).

Pectinases are used by pathogenic microbes in the initial stages of infection to degrade the cell wall either for nutrients, or to gain access to the cytoplasm (Collmer and Keen, 1986; Alghisi and Favaron, 1995). PG in particular has been shown to be the first protein secreted by pathogenic microorganisms during the initial stages of penetration (Jones *et al*. 1972). PGIPs inhibit PG activity in two ways. The first method of inhibition is by physically occupying the PG cleavage site on the homogalacturonan molecules (Protsenko *et al*. 2008). The second method is by binding the PG enzyme and regulating its action (Leckie *et al*. 1999; Manfredini *et al*. 2005). This leads to the release of degraded plant cell wall oligogalacturide fragments in the apoplast (Cook *et
al. 1999), which elicit further resistance reactions in surrounding unaffected cells (De Lorenzo and Ferrari, 2002).

There are multiple plants from which PGIPs have been isolated, such as common bean, soybean, Arabidopsis, tomato and pear (reviewed by De Lorenzo et al. 2001). These proteins have different triggers that induce their expression as well as different specificities for both the pathogen and the range of PGs they employ (Desiderio et al. 1997). For example, in *P. vulgaris*, four PGIPs have been identified: PvPGIP1-4. PvPGIP1 is involved with resistance to *Fusarium moniliforme* via binding of its PG, whereas PvPGIP2 is able to inhibit the of PGs secreted by both *F. moniliforme* and *Aspergillus niger*, but has also been shown to inhibit *Botrytis cinerea* PG1 (Sicilia et al. 2005).

1.7 R Gene Specificity Evolution

Resistance gene specificity for a particular pathogen is an important component for the resistance reaction since loss of specificity in a given gene may cause a resistant plant to become susceptible to the associated disease. New specificities may be achieved through different means which have been observed in multiple plants for both NBS-LRR and kinase genes (Meyers et al. 1998; Meyers et al. 2003; Kim et al. 2009; Schmutz et al. 2010; Huard-Chauveau et al. 2013). Events that lead to new specificities include inter-allelic recombination, unequal pairing and mispairing, gene duplication of small and large DNA fragments and single genes or gene clusters, diversifying selection, intergenic sequence diversion and rearrangements due to transposable element activity (reviewed by Michelmore and Meyers, 1998; Meyers et al. 2005; Liu et al. 2007; Joshi and Nayak, 2013).
These events cause shuffling and duplication of large and small chunks of DNA, which cause changes in amino acid sequences of the resistance proteins, which are seen primarily as high sequence variability and changes in the specificity area of the LRR region (Michelmore and Meyers, 1998). This not only promotes the development of increased, or new pathogen specificities, but also the occurrence of clusters of R genes and pseudogenes. The genes in these clusters are typically related, but have highly divergent sequences indicating duplication events and followed by divergence caused by chromosomal rearrangements (Hulbert et al. 2001; David et al. 2009). Selection of these changes can occur either through positive or balancing selection, depending on the plant and the genes observed (Bakker et al. 2006; Chen et al. 2010; Gos et al. 2012).

1.8 Disease Resistance Research Genomic Resources

1.8.1 Common Bean Genomics

Recently, the genomic sequence of an Andean line G18933, primarily used as a parent of the principal CIAT (Centro Internacional de Agricultura Tropical, Cali Columbia) mapping population was released by a group in the US (available at www.phytozome.net, Schmutz et al. 2014). The sequencing of the G19833 BAC genetic library was performed using the Roche 454 platform with additional information from BAC and fosmid end sequences. Assembly was performed using the Arachne2 system (Jaffe et al. 2003) in combination with genetic markers and the G. max genomic sequence (Schmutz et al. 2010) as a reference. The released assembly is approximately 521.1 Mb, with 27,197 genes which encode 31,638 transcripts, and
transposons accounting for 41% of the genome (www.phytozome.net, Schmutz et al. 2014).

Although the sequencing was performed on an Andean variety of common bean and some rearrangements can be expected between it and Mesoamerican varieties, it is still an excellent tool for characterization and comparison of gene organization, structure and function between the gene pools. It also serves as a useful reference for other *P. vulgaris* sequencing efforts. Current projects include the South American-Spanish “PhaseIbeAm” consortium focused on sequencing the Mesoamerican variety BAT93, as well as the Canadian “Applied Bean Genomics & Bioproducts” effort, focused on sequencing CBB resistant varieties HR67 and OAC Rex, which will be discussed below.

### 1.8.2 OAC Rex Genomic Resources

A binary bacterial artificial chromosome 2 (BiBAC2) genomic library has been created for the OAC-Rex genome (Perry, 2010), providing a resource for genetic analysis of CBB resistance in OAC-Rex. With the use of this library, the OAC Rex has been sequenced using several sequencing methods. Initial sequencing was performed on 16 BiBAC2 clones selected by markers surrounding the PV-ctt001 CBB disease resistance marker using Roche 454 sequencing. The obtained sequence was then assembled into 1,309 contigs based on homology to the *G. max* chromosome 19. Further sequencing of the entire OAC Rex genome was performed using the Illumina (Illumina Inc, California) and PacBio (Pacific Biosciences Inc., California) platforms, which provided a more complete assembly based on homology to the *P. vulgaris* variety G19833 genome (Schmutz et al. 2014).
1.8.3 Other Plant Genomic Resources

Multiple other plant genomes have been sequenced, including *A. thaliana*, *G. max*, *Oryza sativa*, and annotations of genes involved in plant-pathogen interactions are available through the National Centre for Biotechnology Information (NCBI - www.http://www.ncbi.nlm.nih.gov). These available gene sequences provide an excellent resource for homologue comparison to locate genes of interest within the OAC Rex genome. The collection of gene and protein sequences in the NCBI database, as well as their related studies, also make it possible to determine putative gene function, characteristics and protein structure, which can be further tested in the plant.

1.9 Hypothesis and Objectives

It was hypothesized that candidate CBB resistance genes could be identified in the genomic sequence of OAC Rex genomic library clones selected from the PV-ctt001 region of Pv04 using homology to previously identified disease resistance genes with an LRR domain. The first of four objectives relating to the exploration of this hypothesis was to identify and characterize candidate common bacterial blight resistance genes associated with disease resistance marker Pv-ctt001 on Pv04. The second objective was to isolate the predicted candidate genes for future disease resistance testing in the model plant Arabidopsis. The third objective was to develop a detached Arabidopsis leaf tissue culture assay to be used in the testing of candidate genes in the model plant. The last objective was to measure candidate gene expression in susceptible and resistant varieties of common bean, to observe differences in gene expression.
Chapter 2: Identification and Characterization of Candidate Common Bacterial Blight Resistance Genes

2.1 Abstract

Common bacterial blight (Xanthomonas axonopodis pv. phaseoli) is a serious bacterial disease of common bean (Phaseolus vulgaris L.) which causes lesions on the leaves, stems, pods and seeds of the plant. Since the disease reduces the quantity and quality of seed and fresh pod harvests, efforts to breed for resistance have been ongoing for several decades. To improve the understanding of the genetic mechanisms underlying CBB resistant variety OAC Rex, resistance gene homologs were identified within selected BiBAC2 clones from the OAC Rex genomic library. Characterization of the gene homologs identified four candidate genes, as well as six genes in the surrounding sequence that have properties associated with genes involved in disease resistance. Identified genomic features include retrotransposons, a predicted ascorbate oxidase, a callose/glucan synthase and a hypothetical gene. Genetic and protein analyses of these genes are described.
2.2 Introduction

2.2.1 CBB in Common Bean

CBB is a major disease of common bean, caused by the bacterial pathogen \textit{Xap} and its fuscans variant \textit{Xff}. Symptoms produced by the disease include the development of lesions on the leaves, stems, pods and seeds of the plant (Vidaver 1993). The disease affects seed quality and can reduce yield by up to 45% (Saettler 1989; Gillard \textit{et al}. 2009). There is little natural resistance found in wild or domesticated common bean varieties. However, sources of resistance within the same genus have been used to introduce improved CBB resistance into breeding populations; including \textit{Phaseolus acutifolius} (tepary bean) (Parker 1985) and \textit{Phaseolus coccineus} (scarlet runner bean) (Miklas \textit{et al}. 1994).

The first CBB resistant variety of common bean, OAC Rex, was registered in Canada in 2002. It was the result of a cross between HR20-728 and MBE7 (Michaels \textit{et al}. 2006), and resistance originated from an inter-specific cross with \textit{P. acutifolius} accession P1440795 (Parker 1985). Two major resistance QTLs were found in OAC Rex, associated with PV-ctt001 and SU91, as well as two minor affect QTL, associated with BNG71\textsubscript{Dral} and BNG21\textsubscript{EcoRV} (Tar'an \textit{et al}. 2001; Perry \textit{et al}. 2013). PV-ctt001 was first reported in OAC Rex on linkage group Pv05 by Tar'an \textit{et al}. (2001) and was mapped 21.6 cM away from a major CBB gene locus, and accounted for 42.2% of the phenotypic variation. Subsequently, a physical location for PV-ctt001 was identified on chromosome 4, through the creation of a genomic library for OAC Rex (Perry 2010; Perry \textit{et al}. 2013), which is in agreement with a previous study performed by Yu \textit{et al}. (2000a) in a different mapping population. The difference in the location of the marker
between reports was attributed to the location of the marker at the end of the linkage group, which is difficult to map accurately in the absence of a physical map (Tar’an et al. 2001; Perry 2010).

### 2.2.2 OAC Rex Genomic Library

Resistance to CBB in *P. vulgaris* has been a main focus of bean improvement for decades. However, the genes involved have yet to be discovered. To assist in CBB *R* gene identification, the OAC Rex genomic library created by Perry (2010) was used to select clones associated with the PV-ctt001 region of Pv04. Selection was performed by Southern hybridization of library membranes, using primers for PV-ctt001, as well as *P. vulgaris* variety G19833 library clone end sequences PV_GBa00015K06 5’, PV_GBa000015K06 3’ and PV_GBa00100K10 3’ (Schuleter et al. 2008) as probes (Table 2.1). Sixteen positive clones were confirmed to have either 1 or 2 of the probe sequences by PCR (Table 2.2), and the locations of the probes, as well as clone end sequences that were used to orient the sequences into three contigs (Figure 2.1). These 16 clones were sequenced using 454 sequencing, yielding 433,233 reads which were pooled and assembled into 1,550 contigs. The current work was designed to identify candidate CBB resistance genes associated with the PV-ctt001 marker in the contig sequences.

### 2.2.3 R Gene Classifications

One of the main classifications of pathogen R genes are the NBS-LRR protein encoding genes (Jones and Dangl 2006; Padmanabhan et al. 2009). Two basic components of NBS-LRR proteins are NBS domains, which bind and hydrolyze ADP for downstream signalling (Tameling et al. 2002), and LRR domains, which are involved in
Table 2.1. DNA sequences and annealing temperatures for primers used to identify OAC Rex genomic library clones associated with the PV-ctt001 QTL region of Pv04 (Adapted from Perry 2010)

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5’-3’)</th>
<th>Annealing Temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td>PV-ctt001 Forward</td>
<td>GAGGGTGTTTCACATTGTCACCTGC</td>
<td>48°C</td>
</tr>
<tr>
<td>PV-ctt001 Reverse</td>
<td>TTCATGGATGGTGAGGAGGACAG</td>
<td>48°C</td>
</tr>
<tr>
<td>PV_GBa00015K06 5' Forward</td>
<td>CACCTAGACTTCTGCATAAAC</td>
<td>50°C</td>
</tr>
<tr>
<td>PV_GBa00015K06 5' Reverse</td>
<td>TCTTGGTAATGGCCTATTAGG</td>
<td>50°C</td>
</tr>
<tr>
<td>PV_GBa00015K06 3' Forward</td>
<td>GGATATGACCTAATTGCC</td>
<td>50°C</td>
</tr>
<tr>
<td>PV_GBa00015K06 3' Reverse</td>
<td>TCCCTCCCTTCTCAG</td>
<td>50°C</td>
</tr>
<tr>
<td>PV_GBa00100K10 3' Forward</td>
<td>GCCGAGTATAATTCAACCAGC</td>
<td>48°C</td>
</tr>
<tr>
<td>PV_GBa00100K10 3' Reverse</td>
<td>ATTCAGACTTCTGATGATG</td>
<td>48°C</td>
</tr>
</tbody>
</table>

Table 2.2 PV-ctt001 associated OAC Rex genomic library clones. Presence (+) or absence (-) of markers is indicated for each Binary Bacterial Artificial Chromosome 2 (BiBAC2) library clone. (Adapted from Perry 2010)

<table>
<thead>
<tr>
<th>Clone</th>
<th>Marker</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PV_GBa00015K06 5'</td>
</tr>
<tr>
<td>Rex001A06</td>
<td>+</td>
</tr>
<tr>
<td>Rex012B03</td>
<td>-</td>
</tr>
<tr>
<td>Rex020F10</td>
<td>-</td>
</tr>
<tr>
<td>Rex021F10</td>
<td>+</td>
</tr>
<tr>
<td>Rex028A07</td>
<td>-</td>
</tr>
<tr>
<td>Rex045C01</td>
<td>-</td>
</tr>
<tr>
<td>Rex056D11</td>
<td>-</td>
</tr>
<tr>
<td>Rex057D11</td>
<td>-</td>
</tr>
<tr>
<td>Rex063G02</td>
<td>-</td>
</tr>
<tr>
<td>Rex075F07</td>
<td>-</td>
</tr>
<tr>
<td>Rex076C04</td>
<td>-</td>
</tr>
<tr>
<td>Rex090F06</td>
<td>-</td>
</tr>
<tr>
<td>Rex091D05</td>
<td>-</td>
</tr>
<tr>
<td>Rex092G02</td>
<td>+</td>
</tr>
<tr>
<td>Rex213E05</td>
<td>+</td>
</tr>
<tr>
<td>Rex281F03</td>
<td>+</td>
</tr>
</tbody>
</table>
protein-protein interactions and pathogen specificity (Jones and Jones 1997; Ellis et al. 1999; Dodds et al. 2001). Additional small motifs may also be present at the amino terminal ends, including the CC or TIR motifs, which help mediate signalling and the induction of HR (Meyers et al. 2002; Moffett et al. 2002; Rairdan et al. 2008; Bernoux et al. 2011; Maekawa et al. 2011).

Another type of gene involved in plant pathogen resistance are the eLRR genes. This includes RLKs, which contain an eLRR domain, a transmembrane domain and a kinase domain (Sun et al. 2004; Zipfel et al. 2006), RLPs which contain an eLRR domain, transmembrane domain and small cytoplasmic tail (Wang et al. 2008b), and PGIP which is located in the apoplast and contains only an eLRR (Di Matteo et al. 2003; Sakamoto et al. 2003). The secondary structure has been resolved by X-ray crystallography for PGIP2, which was isolated in *P. vulgaris* (Di Matteo et al. 2003), and may be used as a reference to model the structure of LRR domain containing proteins. RLCKs are similar to RLKs, but do not contain an eLRR and are anchored to the cell membrane and have a kinase domain (Veronese et al. 2006; Lu et al. 2010).

### 2.2.4 R Gene Identification in OAC Rex

Both NBS-LRR and the eLRR genes have been studied in many different pathogen systems. However, they have yet to be linked to resistance to CBB in the common bean. The aim of this research was to identify and characterize CGs in the region of the Pv04 linkage group associated with the PV-ctt001 marker in the OAC Rex genome. It was hypothesized that the gene associated with this CBB resistance QTL encoded an LRR protein, since it is the most commonly reported pathogen R gene domain. It was further hypothesized that the conserved sequence of LRR genes could
be used to identify candidate CBB R genes in OAC Rex that could be tested for their ability to reduce the disease.
Figure 2.1 Alignment of selected OAC Rex genomic library clones around CBB resistance marker PV-ctt001. The ~75 kb region of BiBAC2 library clone 21 (Rex021F10; ~150 kb) that represents the likely location for contig1701 is highlighted in blue to indicate the region of focus for this chapter. (Adapted from Perry 2010; Perry et al. 2013).
2.3 Materials and Methods

2.3.1 R Gene Homology Identification in OAC Rex Library Clones

Contigs assembled from the sequence raw reads obtained from clones 1, 12, 20, 21, 28, 45, 56, 57, 63, 75, 76, 90, 91, 92, 213 and 281 (Figure 2.1) were analyzed for homologous regions to known LRR genes (Table 2.3) in other plants using BLAST. The query was run through the CLC Genomics Workbench, and comparisons were based on a collection of plant LRR genes available at the National Centre for Biotechnology Information (NCBI) website (www.ncbi.nlm.nih.gov). The collection included sequences from *A. thaliana*, *G. max*, *O. sativa*, *Populus trichocarpa* and *Z. mays*, which were deposited in the CLC Genomics Workbench for BLAST against the selected BiBAC clone contig sequences (Table 2.3). When selecting which regions of LRR to explore further, the size of contig and the size and number of homologous regions were taken into account. Preference was given to large contigs containing multiple predicted gene regions with E values (probability of obtaining the hit by chance) less than 1E-20.

2.3.2 Confirmation of Contig Sequences

Two contigs, contig1455 and contig1701, were analyzed for sequence continuity. Raw reads from PacBio (Pacific Biosciences, CA) and Illumina sequencing systems were assembled into a consensus sequence using each contig as a reference to produce a consensus sequence. The consensus sequences were aligned to their respective contig sequence (from Roche 454 sequencing). Scaffolds, contigs and raw sequencing reads were also used to confirm the sequence assembly and associated gene annotations of contig1701. The sequence for contig1701 from 35,528 – 53,077 bp was determined to be an assembly error (results presented in section 2.4.1.2), and
Table 2.3 NCBI accession numbers for the NBS-LRR genes used to search for candidate common bacterial blight resistance genes.

<table>
<thead>
<tr>
<th>Species of Origin</th>
<th>NCBI Accession Number</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Arabidopsis Thaliana</em></td>
<td>NM_001123924.1</td>
</tr>
<tr>
<td></td>
<td>NM_001125847.1</td>
</tr>
<tr>
<td></td>
<td>NM_001161067.1</td>
</tr>
<tr>
<td></td>
<td>NM_102893.2</td>
</tr>
<tr>
<td></td>
<td>NM_112307.2</td>
</tr>
<tr>
<td></td>
<td>NM_114955.4</td>
</tr>
<tr>
<td></td>
<td>NM_118071.5</td>
</tr>
<tr>
<td></td>
<td>NM_123196.1</td>
</tr>
<tr>
<td></td>
<td>NM_124542.2</td>
</tr>
<tr>
<td></td>
<td>NM_128428.4</td>
</tr>
<tr>
<td></td>
<td>NM_180323.3</td>
</tr>
<tr>
<td></td>
<td>NM_180841.2</td>
</tr>
<tr>
<td><em>Glycine Max</em></td>
<td>EU839671.1</td>
</tr>
<tr>
<td></td>
<td>EU888329.1</td>
</tr>
<tr>
<td></td>
<td>EU888330.1</td>
</tr>
<tr>
<td></td>
<td>EU888332.1</td>
</tr>
<tr>
<td></td>
<td>FJ014886.1</td>
</tr>
<tr>
<td><em>Oryza Sativa</em></td>
<td>NM_001055547.2</td>
</tr>
<tr>
<td></td>
<td>NM_197391.1</td>
</tr>
<tr>
<td><em>Populus Trichocarpa</em></td>
<td>XM_002301537.1</td>
</tr>
<tr>
<td><em>Zea Mays</em></td>
<td>NM_001112339.1</td>
</tr>
</tbody>
</table>
analyses performed on the 1-35,527 bp section, only, are reported below. However, the final consensus sequence was produced after the majority of analyses were performed on the entire contig1701 sequence (1-53,077 bp). This includes the CG CBB3, which was identified and characterized using the same procedures as the other CGs, and is included in the results reported below. The sequence for contig1455 was aligned to the ~7 kb sequence surrounding CBB2 (25,469-32,210 bp) to determine the similarity of the two sequences.

2.3.3 Gene Identification and Characterization

Genes were identified in contig1701 and contig1455 using Fgenesh (www.softberry.com) to identify transcription start points, exons and poly A tails. The coding sequences and the corresponding proteins were predicted for each identified gene by Fgenesh. Confirmation of gene predictions were performed by identifying the presence of promoter elements (TAT box or CAAT core) in the putative promoter region, expected intron start (GT) and stop (AG) (Luehrsen et al. 1994), internal content of plant introns (Lorković et al. 2000), and poly A tail sequences. Confirmed genes which also had a corresponding resistance gene homology match hit were designated CGs.

2.3.4 CG and Surrounding Sequence Genome Location

2.3.4.1 OAC Rex Genome Assembly Location

Sequences of the contigs as well as the individual genes were BLASTed against the most current assembly and raw sequencing data from the OAC Rex genome. The most current OAC Rex assembly was comprised of 11 pseudochromosomes, 8,532 scaffolds and 30,023 contigs using Illumina, and 1,550 contigs using Roche 454
sequencing data (Perry 2010). The most current raw sequencing data (March 2015) was produced using the PacBio (Pacific Biosciences, CA) sequencing system, and consisted of 992,295 total raw reads.

### 2.3.4.2 G19833 Genome Assembly Location

The coding sequences and full sequences for genes identified in contig1455 and contig1701, as well as the contig sequences were BLASTed against the sequence assembly reported by Schmutz et al. (2014) for *P. vulgaris* variety G19833, to determine their locations within the common bean genome ([http://phytozome.jgi.doe.gov/pz/portal.html](http://phytozome.jgi.doe.gov/pz/portal.html)). An initial indication of putative gene type and function was obtained when a OAC Rex gene matched to a previous gene annotation or other features (i.e. EST, transcript, protein match, RNA expression). In the case where multiple good (>E63) matches within the G19833 sequence assembly were found, the first match, with the highest E-value, was used for further comparisons and analyses.

#### 2.3.4.2.1 Repetitive DNA Analysis

The presence of repetitive DNA elements (i.e. transposable elements) was identified within the regions of the G19833 genome that matched to genes in selected contigs using the Phytozome RepeatMasker feature. The OAC Rex contigs were analyzed using LTR_Finder (Xu and Wang 2007) and LTRphyler ([http://tools.bat.infspire.org/Ltrphyler/](http://tools.bat.infspire.org/Ltrphyler/)) to identify the presence of long terminal repeats (LTRs) and LTR retrotransposon domains.
2.3.5 Characterization of Predicted Proteins

The protein sequences were predicted for genes which were associated with annotations of the G19833 sequence assembly. The annotation types which were used to select OAC Rex genes for further analysis included previously identified genes (including hypothetical genes/proteins), locations where ESTs or transcripts matched to, or where RNASeq data indicated expression was occurring. OAC Rex genes that matched to locations which contained transposable elements were the only annotations that were not included, except for CGs.

2.3.5.1 Domain and Secondary Structure Analysis

The protein sequences were analyzed for the presence of signal peptide using SignalP 4.1 (Petersen et al. 2011), and transmembrane domain prediction using TMHMM Server v. 2.0 (www.cbs.dtu.dk/services/TMHMM). To identify functional domains, the protein sequences were analyzed using InterPro (http://www.ebi.ac.uk/interpro; Mitchell et al. 2014). The secondary structures for the peptide sequences were modelled using Phyre2 (Kelley and Sternberg 2009). BLAST searches at UniProt (www.uniprot.org), NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi), InterPro (http://www.ebi.ac.uk/interpro; Mitchell et al. 2014) and PvGEA (http://plantgrn.noble.org/PvGEA/blasttranscript.jsp; O'Rourke et al. 2014) were performed for each of the proteins to confirm potential functions. Alignments of the coding and peptide sequences for the genes with putative functions with a previously predicted gene of the same type from G19833 or other species were performed to identify whether key features were conserved.
2.3.5.2 CG-Specific Analysis

CG protein sequences were manually screened for leucines and other allopathic amino acids to determine the locations of leucine rich repeat motifs. Leucine rich repeats were identified using the consensus LRR motif LxxxLxLxLxxL. This motif was designed based on the general consensus motif reported in literature (Tornero et al. 1996; Kajava and Kobe 2002; Di Matteo et al. 2003), but modified to fit the recurring consensus in the predicted proteins encoded by the CGs. Manual analysis of the CG protein sequences was also performed to identify NBS, CC and TIR motifs within the sequence, using the motifs described by Meyers et al. (2003) for Arabidopsis TIR- and CC- NBS-LRR genes. Alignment of CG protein sequences was performed using ClustalW (http://www.genome.jp/tools-bin/clustalw), and the locations of LRR motifs were indicated to show protein homology as well as the percent similarity between the CG proteins. CG protein sequences were also aligned to PvPGiP-2 and Swiss-Model (Arnold et al. 2006) was used to provide an additional secondary structure prediction for CGs using PvPGiP-2 as a reference template.

2.3.6 Isolation of CGs

2.3.6.1 Amplification of CGs

Primers were designed for the identified CG regions from contig1701 (Table 2.3) using CLC Genomics. Polymerase Chain Reaction (PCR) was performed using the OAC Rex BiBAC library clone 21 DNA as the template with the following conditions. Each PCR reaction consisted of 2.5 µl of 10x PCR buffer, 2.5 µl of 15 mM MgCl₂, 1 µl of 10 mM dNTP mix, 1 µl of reverse primer, 1 µl of forward primer, 0.5 µl of JumpStart Taq DNA Polymerase (Sigma-Alderich, St. Louis, MO, USA), 1 µl of DNA template and 14.5
µl of sterile distilled water for a total of 25 µl. The PCR parameters included an initial denaturation cycle at 94°C for 1 minute, followed by 30 cycles of denaturation (94°C for 30 seconds), annealing (30 seconds at the temperature specific to the primers – see Table 2.3), and extension (72°C for 3 minutes), this step was followed by 10 minutes of final extension at 72°C, and held at 4°C. The PCR products were separated by electrophoresis through a 1% agarose gel in 1x TBE buffer (108 g Tris, 55 g Boric Acid, 40 mL EDTA in 10 L of distilled water) stained with 2.5% ethidium bromide, and visualized by exposure to UV light. To isolate DNA fragments of interest from the agarose gel the region containing the band of interest was cut from the gel and extracted using PureLink Quick Gel Extraction Kit (Invitrogen, CA).

2.3.6.2 Cloning of CGs

Isolated fragments were inserted into TOPO cloning vectors using a TOPO TA™ Cloning Kit (Invitrogen), and the ligations were then transformed into TOPO E. coli cells using heat shock, according to manufacturer instructions. After overnight culture at 37°C on solid LB selective medium (1 µl ampicillin/mL LB), white colonies were selected and cultured in liquid ampicillin selective medium (1µLampicillin/mL LB). PCR was performed for cultures that grew in the selective liquid media to determine the size of the ligated fragment, with the following conditions. Each PCR reaction consisted of 2.5 µl of 10x PCR buffer, 2.5 µl of 15 mM MgCl₂, 1 µl of 10 mM dNTP mix, 1 µl of reverse primer, 1 µl of forward primer, 0.5 µl of JumpStart Taq DNA Polymerase (Sigma-Aldrich, St. Louis, MO, USA), 1 µl of each positive culture as the DNA template and 14.5 µl of sterile distilled water for a total of 25 µl. The PCR parameters included an initial denaturation cycle at 94°C for 1 minute, followed by 30 cycles of denaturation
(94°C for 30 seconds), annealing (30 seconds at the temperature specific to the primers – see Table 2.3), and extension (72°C for 3 minutes), this step was followed by 10 minutes of final extension at 72°C, and held at 4°C. The PCR products were separated by electrophoresis through a 1% agarose gel in 1x TBE buffer (108 g Tris, 55 g Boric Acid, 40 mL EDTA in 10 L of distilled water) stained with 2.5% ethidium bromide, and visualized by exposure to UV light.

For cultures containing an insert of the appropriate size, the plasmid was isolated using PureLink Quick Plasmid Mini Prep kit (Invitrogen). For each isolated plasmid, 1 kb of the 5’ and 3’ ends of the ligated fragment was sequenced at the University of Guelph Advanced Analysis Centre Genomics Facility (Guelph, ON, CA). The 1 kb fragments were sequenced using M13 forward and reverse primers which correspond to locations on the 5’ and 3’ ends of the cloning site, to confirm the correct sequence was cloned. The end sequences were then compared to CG and contig sequences to determine the locations the fragments that were isolated.
2.4 Results

2.4.1 Gene Identification in Select OAC Rex Contig Sequences

2.4.1.1 R Gene Homology Identification

Several contigs from OAC Rex BiBAC2 library clone 21 were identified as having significant homology to previously reported NBS-LRRs (Figure 2.1). Contigs1701 and 1455 were selected for analysis. Contig1701 was selected for further analysis because of its large size (~50 kb) and the presence of 3 NBS-LRR homologous regions in its sequence. Contig1455 was added after the initial BLAST and selection of CGs, although it was not originally analyzed. Interest in this contig arose from cloning and sequencing work in later experiments (described below).

2.4.1.2 Contig Sequence Assembly Confirmation

2.4.1.2.1 Contig1701 Assembly Confirmation

When the contig1701 sequence was BLASTed against the current OAC Rex assembly and raw sequence data, 2 scaffolds (scaffold1045 and scaffold296), 4 contigs (contig9955, contig9956, contig4106 and contig4107) and 3 PacBio raw reads (PBR2, PBR4 and PBR11) matched the best (between 87-99% similarity) to 1-35,527 bp, with some overlapping sequence (Figure 2.2). When the contig1701 sequence was BLASTed to the OAC Rex pseudochromosomes there were no locations with high similarity observed. Within the contig1701 sequence, the1-19,586bp section was 98% similar to scaffold1045, and the 19,668-35,527 bp section was 97% similar to scaffold296 (Figure 2.2). Contigs and raw reads matched within these two sections, but not between, and therefore did not support the assembly of a continuous sequence for contig1701 as illustrated in Figure 2.2. A section of scaffold1045 (5,968-26,642 bp)
matched to contig1701 at 1-19,626 bp, and corresponded with contig9955, contig9956, and pbr2. The 1-5,967bp overhang region of scaffold1045 which was missing on the contig1701 sequence corresponded to the same overhang region of contig9955. The section of scaffold296 (113,508-130,206 bp) that matched to contig1701 at 20,885-35,527 bp, corresponded with contig4107, contig4106, pbr4 and pbr11. The same region of scaffold296 (113,508-130,206 bp) also had high similarity to the region between 35,647-49,671 bp in the reverse compliment direction compared to the 20,885-35,527 bp region. Alignment of the two regions (19,668-35,527 bp and 35,647-49,671 bp) within contig1701, which corresponded to scaffold296 (113,508-130,206 bp), showed them to be 87% similar, with large stretches of identical sequence (>1kb) interrupted by smaller stretches of no, or low similarity (<200 bp).

Since a single location could not be identified for the entirety of the contig1701 sequence within the OAC Rex sequence assembly, it was divided into four sections based on homology to OAC Rex contig and scaffold matches (Figure 2.2). The first section extended from 1-19,586 bp (contig1701-1), the second from 19,668-35,527 bp (contig1701-2), the third from 35,647-49,671 bp (contig1701-3) and the fourth and final section from 50,608-53,077 bp (contig1701-4).

The junctions between sections contig1701-1 and contig1701-2, and contig1701-3 and contig1701-4 were confirmed on the basis of the similarity between the contig1701 sequence and the consensus sequence produced when PacBio raw reads were assembled using the contig1701 sequence as a reference. The contig1701 sequence and the assembly consensus sequence were 93% similar for the junction between contig1701-1 and contig1701-2 sections, and were 97% similar for the junction...
between contig1701-3 and contig1701-4 sections. No consensus assembly was obtained for the junction between contig1701-2 and contig1701-3 sections. Therefore, because the sequence from 35,527-53,077 bp was highly similar to the sequence from 1-35,527 bp, and the number of PacBio raw reads that matched the 35,527-53,077 bp sequence was significantly lower (Figure 2.2), further analyses were performed only on the genes predicted in the 1-35,527 bp sequence. However, characterization of CBB3 from the discarded region (39,732-42,310 bp), is reported below, and in Chapter 4, since these analyses were performed on this gene previous to the confirmation of the continuity of the contig1701 sequence assembly.

### 2.4.1.2.2 Contig1455 Assembly Confirmation

Comparison of contig1455 to the most recent assembly of the OAC Rex genome identified high similarities to one scaffold, one contig and one raw sequence read (Table 2.7). The 1-6,365 bp region of contig1455 matched to scaffold296 (117,948-124,325) which also corresponded to contig4106, and the 47-7,720 bp region matched to pbr11. Alignment of contig1455 to the section within contig1701 (25,469-32,210 bp), which also matched to contig4106 and pbr11, indicated 92% similarity, and the predicted genes were present in similar arrangements in all three locations (Figure 2.2). This confirmed the continuity of the contig1455 sequence assembly.

### 2.4.1.3 Contig1701 Gene Annotation

When FGeneSH was used to identify genes on contig1701, 16 genes were identified within the whole sequence (1-53,077 bp) (Figure 2.2). Within the confirmed sequence (1-35,527 bp), 10 genes were identified (Figure 2.2-2.3: Table 2.4). Promoter elements were identified for all ten genes, but poly A tails could only be identified for
Figure 2.2 OAC Rex contig1701 sequence confirmation, with alignments between OAC Rex sequence elements and the contig1701 and contig1455 sequence assemblies. A) Frequency of OAC Rex PacBio raw read match locations within contig1701 when it was used as the reference for a sequence assembly. The number of reads mapping to the reference is indicated by the lefthand scale. B) Alignments between contig1701 and PacBio raw reads, Illumina sequencing assembled contigs and scaffolds, and contig1455. Contig1701 and contig1455 are highlighted in blue. Sections within contig1701 where sequence integrity was in question are highlighted in green. Sequences with high homology to contig1701 (>85%) are indicated by solid lines. Sequences with low homology to contig1701 (<85%) are indicated by dashed lines. Arrows indicate genes with high homology between sequences. C) Summary of OAC Rex genome assembly elements, their percent similarity and location of similarity to contig1701 and contig1455.
genes 4-10 (Table 2.4). The predicted exon/intron splice points were confirmed for each of the ten genes.

### 2.4.1.3.1 Contig1701 CGs

Three resistance gene homologous regions were identified on contig1701, based on similarity to a reported resistance gene from *O. sativa* (NM_197391.1; Table 2.3). The first homologous region was 662 bp long, 57% similar to the matching contig sequence (2E-20), and extended from the promoter region to the second coding region of a predicted gene, which was designated candidate CBB resistance gene CBB1 (123-2,274 bp) (Figure 2.5A; Table 2.4). CBB1 was 2,249 bp in size, with 3 exons consisting of 804 bp of coding sequence (Figure 2.3; Table 2.4 and 6). The second homologous region was 266 bp long, 65% similar to the matching contig sequence (3E-23), and located between the first and second coding regions of a predicted gene, designated CG CBB2 (27,481-31,229 bp) (Figure 2.5B; Table 2.4). CBB2 was 3,922 bp in size, with 4 exons comprising 1,452 bp of coding sequence (Figure 2.3; Table 2.4 and 6). The third significant resistance gene homologous region was 488 bp long, 59% similar (2E-24), and extended from the first coding region into the second coding region of a predicted gene, designated candidate CBB resistance gene CBB3 (39,732-42,310 bp) (Figure 2.5C; Table 2.4 and 2.6). CBB3 is 2,806 bp in size with 3 exons comprising 1,020 bp of coding sequence (Figure 2.3; Table 2.5).

### 2.4.1.4 Contig1455 Gene Annotation

FGeneSH identified three genes on contig1455 (Figure 2.4; Table 2.5). Promoter elements were identified for all three genes, but poly A tails could only be identified for genes 1-2, indicating they were likely not whole genes (Table 2.5).
Figure 2.3 Locations, coding sequences and potential functions of predicted genes on contig1701. Candidate genes CBB1, CBB2 and CBB3 are indicated in blue, yellow and red respectively. Genes with solid fills have promoter and poly A elements, and outlined genes are missing both the promoter elements and a poly A. A BLAST analysis against the G19833 genome sequence associated predicted genes with retroelement (red and candidate genes), L-ascorbate oxidase/multi-copper oxidase (pale green), and glucan-like synthase/callose synthase annotations.
<table>
<thead>
<tr>
<th>Gene</th>
<th>Contig1701 Location</th>
<th>TATA Box</th>
<th>CAAT Core</th>
<th>Coding Sequence</th>
<th>Poly A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gene 1  (CBB1)</td>
<td>123 – 2,274</td>
<td>2,294-2,299</td>
<td>2,367-2,371</td>
<td>CDS1 – 2,119-2,158</td>
<td>-</td>
</tr>
<tr>
<td>Gene 2</td>
<td>2,426 – 5,867</td>
<td>5,891-5,898</td>
<td>6,193-6,197</td>
<td>CDS1 – 3,965-4,145</td>
<td>-</td>
</tr>
<tr>
<td>Gene 3</td>
<td>8,273 – 9,780</td>
<td>-</td>
<td>-</td>
<td>CDS1 – 9,633-9,731</td>
<td>-</td>
</tr>
<tr>
<td>Gene 5</td>
<td>14,960 – 20,200</td>
<td>20,228-20,234</td>
<td>-</td>
<td>CDS1 – 18,739-18,990</td>
<td>14,972-14,946</td>
</tr>
<tr>
<td>Gene 6</td>
<td>20,838 – 21,767</td>
<td>21,797-21,801</td>
<td>22,224-22,228</td>
<td>CDS1 – 21,334-21,504</td>
<td>20,839- 20817</td>
</tr>
<tr>
<td>Gene 10</td>
<td>32,417 – 34,101</td>
<td>34,128- 34,133</td>
<td>-</td>
<td>CDS1 – 33,736-33,855</td>
<td>32,413- 32,422</td>
</tr>
</tbody>
</table>
The predicted exon/intron splice points were confirmed for each of the three predicted genes. CBB4 is 4,049 bp in size and includes 6 exons comprising a 1,428 bp coding sequence (Figure 2.4-2.5D; Table 2.5-2.6).

2.4.2 OAC Rex Gene Locations within OAC Rex Genome Assembly

In attempting to confirm the sequence continuity of the contig1701 and contig1455 (see section 2.4.1.2) sequences, no location within the most recent OAC Rex pseudochromosomes could be found. Therefore, no conclusions about the chromosome locations within OAC Rex could be made.

2.4.2.1 G19833 Location for Contig1701 Identified Genes

When the sequences of the ten genes identified in contig1701 were BLASTed to the G19833 genome sequence, the genes matched to multiple locations (>100) within and between chromosomes, with varying similarities and confidence values. The best result for each CG resulted in genes 1-5 matched to chromosome 6 and genes 6-10 matching to chromosome 8 (Figure 2.5; Table 2.5). However, when BLAST was performed using the entire contig1701 sequence a location within the G19833 sequence assembly was not identified (data not shown).

Putative functions for the genes on contig1701 were assigned according to the annotations present for the G19833 genes they were matched to. Gene 5 matched to the same location within chromosome 6 as an annotation for a multi-copper oxidase/L-ascorbate oxidase gene and for which expression was observed in multiple tissue sources (Figure 2.3 and 2.6; Table 2.7). Gene 6 matched to the same location within chromosome 8 as a hypothetical protein, for which expression was observed in the nodules and roots (Figure 2.3 and 6; Table 2.7). Gene 7 matched to the same location
Figure 2.4 Location, coding sequence and potential function of predicted genes on contig1455 showing candidate gene CBB4 (bright green). Genes with solid fill have promoter and poly A elements, and outlined genes are missing both the promoter and poly A. A BLAST analysis against the G19833 genome sequence associated predicted genes with retroelement (red and candidate genes).

Table 2.5 Locations within the contig1455 sequence assembly and genetic features of contig 1455 gene annotations

<table>
<thead>
<tr>
<th>Gene</th>
<th>Contig 1455 Location</th>
<th>TATA Box</th>
<th>CAAT Core</th>
<th>Coding Sequence</th>
<th>Poly A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gene 1</td>
<td>282-706</td>
<td>283-290</td>
<td>-</td>
<td>CDS1 – 412-549</td>
<td>688-706</td>
</tr>
<tr>
<td>Gene 2 (CBB4)</td>
<td>1,604-6,056</td>
<td>6,002-6,008</td>
<td>-</td>
<td>CDS1 – 5,064-5402</td>
<td>1,604-1,736</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CDS2 – 4,544-4,759</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CDS3 – 4,135-4347</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CDS4 – 3,908-4033</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CDS5 – 2,997-3110</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CDS6 – 2,491-2916</td>
<td></td>
</tr>
<tr>
<td>Gene 3</td>
<td>6,695-8,815</td>
<td>-</td>
<td>8051-8054</td>
<td>CDS1 – 7,619-7969</td>
<td>-</td>
</tr>
</tbody>
</table>
Figure 2.5 Candidate common bacterial blight (CBB) resistance genes. Gene organization for candidate genes CBB1 (A), CBB2 (B), CBB3 (C) and CBB4 (D). Promoters are indicated with black arrow boxes, coding regions are indicated by colored boxes, pathogen resistance gene homologies are indicated by the outline of a rectangle and the poly A tails are indicated with four black lines. The positions (bp) within contig1701 for CBB1, CBB2, and CBB3 and contig1455 for CBB4 are indicated.

Table 2.6 Summary of candidate common bacterial blight resistance gene characteristics. The sizes of the candidate genes, predicted coding sequences, protein sizes and the number of leucine rich repeats for each of the candidate genes CBB1, CBB2, CBB3 and CBB4 are indicated.

<table>
<thead>
<tr>
<th>Candidate Gene</th>
<th>Total Gene Size (Base Pairs)</th>
<th>Coding Sequence Size (Base Pairs)</th>
<th>Protein Size (Amino Acids)</th>
<th>Number of LRRs</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBB1</td>
<td>2,249</td>
<td>804</td>
<td>267</td>
<td>8</td>
</tr>
<tr>
<td>CBB2</td>
<td>3,922</td>
<td>1,452</td>
<td>483</td>
<td>11</td>
</tr>
<tr>
<td>CBB3</td>
<td>2,806</td>
<td>1,020</td>
<td>339</td>
<td>10</td>
</tr>
<tr>
<td>CBB4</td>
<td>4,453</td>
<td>1,428</td>
<td>476</td>
<td>11</td>
</tr>
</tbody>
</table>
within chromosome 8 as an annotation for a β-glucan-like synthase/callose synthase for which expression was not observed in the source tissues used in the RNA-seq study performed by Schmutz et al. (2014) and reported in Phytozome (Figure 2.3 and 6; Table 2.7). The remaining genes, including CBB1-3, matched to regions within chromosomes 6 and 8 which were annotated as retrotransposons (Table 2.7). Some of the retrotransposons had associated RNAseq data indicating expression, but no source tissue locations were given (Table 2.7).

2.4.2.2 G19833 Location for Contig1455 Identified Genes

When the genes identified in contig1455 were BLASTed against the G19833 sequence, they matched to multiple different locations within and between chromosomes. The best results matched genes 1 and 2 on chromosome 8 and gene 3 on chromosome 6 (Figure 2.7; Table 2.8). Each of the three predicted genes matched to locations where retrotransposons had been identified, but no other gene annotations were present. Contig1455 predicted genes 2 and 3 were associated with locations where RNA seq data indicated gene expression to be present. Two of the three predicted genes were located in the same region as contig1701 predicted genes. CBB4 matched to the same location as CBB2 and CBB3. Contig1455 gene 1 was matched to the same location as contig1701 gene 10.
Figure 2.6 Sequence comparisons between OAC Rex contig1701 and syntenic regions in G19833. Contig1701 sections coloured by blue and purple blocks correspond to chromosome 6 (5.817–5.841 Mbp) and green and yellow coloured regions corresponded to chromosome 8 (49.680–49.698 Mbp) when compared (with BLAST) to the G19833 sequence on Phytozome. Predicted genes for G19833 sequences are indicated by grey arrows for genes matching to contig1701 genes and arrow outline for genes not matching contig1701 genes. Gene match locations are indicated with lines and white boxes.
Table 2.7 Predicted locations in G19833 and expression of annotated genes in OAC Rex contig1701. The positions in the G19833 assembly (Phytozome; Schmutz et al. 2014) are indicated by chromosome number and location (bp) of contig1701 gene synteny, and associated gene(s) or element(s), which indicate a possible function.

<table>
<thead>
<tr>
<th>Gene Number</th>
<th>G19833 Chromosome Location (bp)</th>
<th>Similarity to G19833</th>
<th>Associated Genetic Feature</th>
<th>G19833 Expression Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gene 1 (CBB1)</td>
<td>5,817,952 – 5,820,113</td>
<td>98%</td>
<td>pvRetroSF205-2, pvRetroSF112-2, pvRetroSF100-1</td>
<td>No expression</td>
</tr>
<tr>
<td>Gene 2</td>
<td>5,820,266 – 5,823,700</td>
<td>98%</td>
<td>pvRetroSF100-1, pvRetroSF172-1</td>
<td>No expression</td>
</tr>
<tr>
<td>Gene 3</td>
<td>5,826,105 – 5,832,901</td>
<td>98%</td>
<td>pvRetroSF318-1</td>
<td>No expression</td>
</tr>
<tr>
<td>Gene 4</td>
<td>5,825,866 – 5,833,506</td>
<td>97%</td>
<td>pvRetroSF318-1</td>
<td>No expression</td>
</tr>
<tr>
<td>Gene 5</td>
<td>5,835,002 – 5,839,661</td>
<td>98%</td>
<td>Multi-copper oxidase</td>
<td>Expression in flowers, pods, leaves, stem, nodules, roots</td>
</tr>
<tr>
<td>Gene 6</td>
<td>49,682,264 – 49,683,238</td>
<td>97%</td>
<td>Unknown protein transcript</td>
<td>Expression in nodules, roots</td>
</tr>
<tr>
<td>Gene 7</td>
<td>49,683,929 – 49,687,869</td>
<td>92%</td>
<td>Glucan synthase</td>
<td>No expression</td>
</tr>
<tr>
<td>Gene 8 (CBB2)</td>
<td>49,682,319 – 49,693,055</td>
<td>75%</td>
<td>pvRetro12, pvRetroSF188-3</td>
<td>Expression</td>
</tr>
<tr>
<td>Gene 9</td>
<td>49,693,473 – 49,693,864</td>
<td>88%</td>
<td>pvRetroSF188-3</td>
<td>Expression</td>
</tr>
<tr>
<td>Gene 10</td>
<td>49,694,477 – 49,699,799</td>
<td>69%</td>
<td>pvRetro19</td>
<td>No expression</td>
</tr>
<tr>
<td>Gene 13 (CBB3)</td>
<td>49,688,390 – 49,692,057</td>
<td>95%</td>
<td>pvRetro12, pvRetroSF188-3</td>
<td>Expression</td>
</tr>
</tbody>
</table>
Figure 2.7 Alignments of contig1455 and associated match locations within G19833. The locations where contig1455 gene annotations match within the G19833 assembly chromosomes are indicated with lines. Candidate gene CBB4 is indicated by a green arrow, and other gene predictions within contig1455 are indicated with red arrows. The red coloured block indicates similarity to chromosome 6 and the orange coloured block indicates similarity to chromosome 8 within the G19833 assembly.
Table 2.8 Predicted gene locations on OAC Rex contig 1455 in bp with associated G19833 chromosome number, location (in bp), gene or element and possible function.

<table>
<thead>
<tr>
<th>Gene Number</th>
<th>G19833 Chromosome Location</th>
<th>Similarity to G19833</th>
<th>Associated Genetic Feature</th>
<th>G19833 Expression Details</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Chromosome 6</td>
<td>Chromosome 8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene 1</td>
<td>49,693,478 – 49,693,898</td>
<td>99%</td>
<td>pvRetro19</td>
<td>No Expression</td>
</tr>
<tr>
<td>Gene 2 (CBB4)</td>
<td>49,688,126 – 49,691,770</td>
<td>98%</td>
<td>pvRetroSF188-3, pvRetro12</td>
<td>Expression</td>
</tr>
<tr>
<td>Gene 3</td>
<td>2,980,777 – 2,981,831</td>
<td>95%</td>
<td>pvRetro3</td>
<td>Expression</td>
</tr>
</tbody>
</table>
2.4.3 Contig1701 Gene and Protein Characterization

2.4.3.1 Characterization of CGs

2.4.3.1.1 CG Similarities

Alignments of the CG sequences showed that CBB3 had the highest similarity to CBB2 and CBB4 (≥90%), CBB2 was less similar to CBB4 (72%), and as a group CBB2, CBB3, CBB4 (CBB2-4) had low similarity with CBB1 (≤15%) (Table 2.9). When the predicted coding sequences were aligned for CBB1-4 (Figure 2.8), the sequences were less similar than the full gene sequences (Table 2.9). CBB1 was the least similar to CBB2-4 (≤11%), CBB3 had the highest similarity to CBB2 and CBB4 (≥89%), and CBB2 was less similar to CBB4 (65%) (Figure 2.8; Table 2.9). The predicted protein sequences for the CGs were 267, 483, 339 and 476 amino acids long for CBB1, CBB2, CBB3 and CBB4, respectively (Table 2.6). Alignments between amino acid sequences showed there to be low similarity between CBB1 and CBB2-4 (≤17%), the highest similarity was seen between CBB3 and CBB4 (84%) (Figure 2.9; Table 2.9). Large stretches of conserved sequence could be seen in all three comparisons between CBB2-CBB4, especially in the carboxyl end, with stretches of non-conserved, variable amino acids at the amino-terminal end as well as roughly in the middle of the protein sequence (Figure 2.8-9).

2.4.3.1.2 CG Domain Analysis

Domain analysis of CBB1-4 determined that there were no NBS or related motifs present in any of the CG predicted proteins. In addition, no TIR or CC motifs could be identified in their N terminal ends. Further analysis of the N terminal ends indicated that no signal peptides for extracellular transport exist in the putative proteins, indicating a
low likelihood that the proteins function outside the cell. Also, no transmembrane
domains were present in the putative protein sequences. These results indicate that the
CG proteins likely function within the cell.

Although NBS motifs were absent in all four putative CG proteins, they all contained high numbers of leucine and other allopathic amino acids, as well as an LDL motif, with a consensus sequence of LDLLLD. The predicted amino acid sequences contained 8, 11, 10, and 11 LRRs for CBB1, CBB2, CBB3 and CBB4, respectively (Figure 2.9-10; Table 6). These LRRs conformed to the consensus motif sequence consisting of LxxxLxLxLxxL, (Figure 2.9) where L represents a leucine or allopathic amino acid and x represents any amino acid. Some variability in the lengths of LRRs and the spacings between Ls was observed.

2.4.3.1.3 Contig Sequence Retrotransposon Analysis

Different analyses for gene and protein features provided low confidence
evidence that CGs were retrotransposons. BLAST searches using contig1701 and
contig1455 predicted gene sequences against the G19833 genome sequence indicated
that many of the predicted genes on contig 1701 and all contig1455 genes were
retrotransposons (Figure 2.3-4; Table 2.4-5). However, when CGs were BLASTed
against the NCBI database, only the sequence for CBB1 (gene 1) matched to a
retrotransposon (NCBI accession FJ402925.1). NCBI BLAST performed using CG
protein sequences resulted in multiple matches to unknown/hypothetical proteins.
BLAST searches against the UniPro database indicated that all four CGs had low (30-
47%) similarity to retrotransposon proteins including Gag-pol poly protein, POL3-like
Table 2.9 Similarity between candidate common bacterial blight resistance genes. Percent similarity between candidate resistance genes was determined by aligning full gene sequences, coding sequences and protein sequences in ClustalW.

<table>
<thead>
<tr>
<th>Genes Compared</th>
<th>Compared Gene Similarity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Full Gene</td>
</tr>
<tr>
<td>CBB1</td>
<td></td>
</tr>
<tr>
<td>CBB2</td>
<td>14%</td>
</tr>
<tr>
<td>CBB3</td>
<td>15%</td>
</tr>
<tr>
<td>CBB4</td>
<td>14%</td>
</tr>
<tr>
<td>CBB2</td>
<td></td>
</tr>
<tr>
<td>CBB3</td>
<td>90%</td>
</tr>
<tr>
<td>CBB4</td>
<td>72%</td>
</tr>
<tr>
<td>CBB3</td>
<td></td>
</tr>
<tr>
<td>CBB4</td>
<td>93%</td>
</tr>
</tbody>
</table>
Figure 2.8 Alignment of candidate common bacterial blight resistance gene coding sequences. Candidate genes CBB1 (blue), CBB2 (yellow), CBB3 (red) and CBB4 (green) were aligned in ClustalW. Coding sequences are indicated with alternating coloured rectangles.
Figure 2.9 Alignment of candidate gene protein sequences. LRR regions highlighted for candidate genes CBB1 (blue) CBB2 (yellow), CBB3 (red) and CBB4 (green), LDL motif is indicated by a box.
<table>
<thead>
<tr>
<th>A) CBB1</th>
<th>B) CBB2</th>
<th>C) CBB3</th>
<th>D) CBB4</th>
</tr>
</thead>
<tbody>
<tr>
<td>EKTKFGVMSDE</td>
<td>KREDPPTDE</td>
<td>RDKHPPTDEL</td>
<td>LQTSTFTSLYLL</td>
</tr>
<tr>
<td>VDEKVKVTI</td>
<td>LYSGHOQRI</td>
<td>KEHVLHQ</td>
<td>ESTNSHLSL</td>
</tr>
<tr>
<td>VNEVRSEHGL</td>
<td>DGHYFELEH</td>
<td>FATSQSGNTG</td>
<td>RHPDPITPDE</td>
</tr>
<tr>
<td>IATPNAVK</td>
<td>FATOSQLYGH</td>
<td>RGFENTTF</td>
<td>LHSGVDYR</td>
</tr>
<tr>
<td>LKEKPFELAL</td>
<td>LRGFLSLTGQ</td>
<td>TTPKTM</td>
<td>FATSQVESNGN</td>
</tr>
<tr>
<td>YNDDLDFSILY</td>
<td>FTEKLLOISOV</td>
<td>LAQVDSTCH</td>
<td>LRGFNLNTGITY</td>
</tr>
<tr>
<td>LVGHFRIAKTE</td>
<td>LNRSLEAYQOCF</td>
<td>LHQHTLVSHA</td>
<td>LVTPLTEPLTI</td>
</tr>
<tr>
<td>LGRYFGYGFY</td>
<td>LHOHTFVHLH</td>
<td>VWGRQSTYRQ</td>
<td>LQAQYDSTCH</td>
</tr>
<tr>
<td></td>
<td>VSFGQVLLRR</td>
<td>YSGPOYRR</td>
<td>LHQEHTVNHK</td>
</tr>
<tr>
<td></td>
<td>VRCFDVLS</td>
<td>HPHELHSL</td>
<td>VWGRQSYVRG</td>
</tr>
<tr>
<td></td>
<td>VTSDDFLAPK</td>
<td></td>
<td>VSVGQYER</td>
</tr>
</tbody>
</table>

**Figure 2.10** LRR sequences of candidate genes A) CBB1, B) CBB2, C) CBB3 and D) CBB4 with LRR consensus motif indicated.
reverse transcriptase, Gypsy/Ty-3 retro-element polyprotein, as well as many uncharacterized proteins.

Domain analysis with InterPro was unable to identify protein domains within any of the CG protein sequences. However, the program indicated the presence of protein signatures for all four CGs other than CBB2, providing a suggested protein type/function. For CBB1, reverse transcriptase, GAG/POL/ENV polyprotein and DNA/RNA polymerase signatures were found. For CBB2, no protein signatures were found. For CBB3, reverse transcriptase and DNA/RNA polymerase signatures were found. Finally, for CBB4, reverse transcriptase and DNA/polymerase signatures were identified.

LTR_finder was unable to identify long terminal repeats within either of the contig sequences. However, LTRphyler identified multiple retrotransposon poly protein domains associated with CGs. Within CBB1, both an integrase core domain (660-1028 bp) and a Gypsy reverse transcriptase domain (1,908-2,374 bp) were identified which corresponded to the same region as CDS1 and CDS3, although they also associated with intron sequence. Gypsy reverse transcriptase domains were identified in CBB2, CBB3, and CBB4. For CBB2, the reverse transcriptase domain was located between CDS1 and CDS2 (28,086-28,295 bp). For CBB3, the reverse transcriptase domain was associated with the location of CDS2 and the adjacent intron (41,622-41,831 bp). For CBB4, the reverse transcriptase domain corresponded to the location of CDS2 (4,572-4,814).
Figure 2.11 Secondary structure models of predicted proteins for candidate CBB resistance genes CBB1 (A), CBB2 (B), CBB3 (C) and CBB4 (D), using Phyre2 software (Kelley and Sternberg 2009)
Table 2.10 Phyre2 secondary structure modelling (Kelley and Sternberg 2009) for candidate gene proteins CBB1, CBB2, CBB3 and CBB4

<table>
<thead>
<tr>
<th>Candidate Gene</th>
<th>Amino Acids Modelled</th>
<th>Template</th>
<th>Template Characteristic</th>
<th>Similarity to Template</th>
<th>Confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBB1</td>
<td>14-127</td>
<td>D1mu2a2</td>
<td>Reverse transcriptase DNA/RNA polymerase</td>
<td>12%</td>
<td>99.9%</td>
</tr>
<tr>
<td></td>
<td>146-262</td>
<td>C3I2uA_</td>
<td>Prototype foamy virus (pfv)2 integrase</td>
<td>22%</td>
<td>99.7%</td>
</tr>
<tr>
<td>CBB2</td>
<td>5-271</td>
<td>D1mu2a2</td>
<td>Reverse transcriptase DNA/RNA polymerase</td>
<td>14%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>213-430</td>
<td>C3I2uA_</td>
<td>Prototype foamy virus (pfv)2 integrase</td>
<td>19%</td>
<td>99.9%</td>
</tr>
<tr>
<td>CBB3</td>
<td>5-173</td>
<td>D1mu2a2</td>
<td>Reverse transcriptase DNA/RNA polymerase</td>
<td>18%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>193-338</td>
<td>C3I2uA_</td>
<td>Prototype foamy virus (pfv)2 integrase</td>
<td>15%</td>
<td>99.7%</td>
</tr>
<tr>
<td>CBB4</td>
<td>9-68</td>
<td>C3nr6A_</td>
<td>Xenotropic murine leukemia virus-related virus2 (xmrv) protease</td>
<td>21%</td>
<td>97.9%</td>
</tr>
<tr>
<td></td>
<td>75-279</td>
<td>D1mu2a2</td>
<td>Reverse transcriptase DNA/RNA polymerase</td>
<td>15%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>218-476</td>
<td>C3I2uA_</td>
<td>Prototype foamy virus (pfv)2 integrase</td>
<td>12%</td>
<td>99.7%</td>
</tr>
</tbody>
</table>
2.4.3.1.4 CG Secondary Structure Prediction

Modeling of protein secondary structure using Phyre2 (Kelley and Sternberg 2009) resulted in models for each of the CG proteins (Figure 2.11). CGs CBB1-4 were modelled with the same reverse transcriptase (D1mu2a2) and integrase (C312uA_) proteins, and CBB4 had an additional section modelled with a protease protein (C3nr6A_) (Table 2.11). The model for CBB2 was produced using two viral integrase proteins (Table 2.11). The presence of the manually modelled LRR motifs were unconfirmed by the secondary structure models produced for the CGs, except for the small section of β-sheet structure in CBB4 (Figure 2.11).

2.4.3.2 Characterization of Genes with Putative Functions

2.4.3.2.1 Ascorbate Oxidase Gene Homologs

Gene 5 matched to a Multicopper oxidase/L-ascorbat oxidase gene at 5,835,002 – 5,839,661 bp on chromosome 6 (Figure 2.6; Table 2.7). When the full gene sequence for gene 5 was BLASTed at NCBI, putative conserved Cupredoxin superfamily domains, and L-ascorbate oxidase and multicopper oxidase multi-domains were predicted (data not shown). When the amino acid sequence was BLASTed against the PvGEA Atlas (O’Rourke et al. 2014), matches were identified in several chromosomes to several L-ascorbate oxidase annotations, including the L-ascorbate oxidase annotation gene 5 matched to within G19833 chromosome 6 (Phvul.006G011700). When the sequence was BLASTed to the G. max genome (Schmutz et al. 2010) in Phytozome it identified homology to an ascorbate oxidase annotation at 11,769,051 – 11,776,969 bp on chromosome 20 (Glyma.20G051900).
Table 2.11 Similarity between contig1701 genes of interest and identified homologues. Percent similarity between genes identified in *P. vulgaris* varieties OAC Rex contig1701 (gene 5), G19833 annotation Phvul.006G011700, G. max annotation Glyma.20G051900 was determined by aligning full gene sequences, coding sequences and protein sequences in ClustalW. Zucchini ascorbate oxidase protein sequence (1ASP_A; Messerschmidt *et al.* 1993) was also aligned to the protein sequences to indicate conservation of protein domains.

<table>
<thead>
<tr>
<th>Genes Compared</th>
<th>Similarity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Full Gene</td>
</tr>
<tr>
<td>Gene 5</td>
<td>91</td>
</tr>
<tr>
<td>Phvul.006G011700</td>
<td>Glyma.20G051900</td>
</tr>
<tr>
<td>Glyma.20G051900</td>
<td>19</td>
</tr>
<tr>
<td>1ASP_A</td>
<td>-</td>
</tr>
<tr>
<td>Phvul.006G011700</td>
<td>Glyma.20G051900</td>
</tr>
<tr>
<td>Glyma.20G051900</td>
<td>1ASP_A</td>
</tr>
<tr>
<td>Glyma.20G051900</td>
<td>1ASP_A</td>
</tr>
<tr>
<td>Gene 7</td>
<td>G2</td>
</tr>
<tr>
<td>Glyma.08G361500</td>
<td>2</td>
</tr>
<tr>
<td>G2</td>
<td>Glyma.08G361500</td>
</tr>
<tr>
<td>Gene 6</td>
<td>Phvul.008G191100</td>
</tr>
</tbody>
</table>
Alignment of the three gene sequences (gene 5, Phvul.006G011700 and Glyma.20G051900) showed high similarity between gene 5 and Phvul.006G011700 (91%), but the similarity was low between Glyma.20G051900 and the other two genes (Table 2.1). When the coding sequences for the three genes were aligned, higher similarities were observed, with similarity between the coding sequences >80% (Figure 2.12; Table 2.1). Alignments of the protein sequences for gene 5, Phvul.006G011700, Glyma.20G051900, and X-ray modelled Zucchini ascorbate oxidase (IASP_A), obtained from NCBI indicated a number of regions with high similarities (Figure 2.13; Table 2.1).

When InterPro was used to identify domains within gene 5, Phvul.006G011700, Glyma.20G051900 and IASP_A, multicopper oxidase type 1, 2 and 3 domains were identified (Figure 2.13). The amino acid sequences of these domains were highly conserved between the aligned genes. However, the domains indicated by InterPro in gene 5 were less conserved (Figure 2.13). Phyre2 secondary structure modelling of gene 5, Phvul.006G011700 and Glyma.20G051900 used the X-ray structure of the Zucchini ascorbate oxidase 1ASP_A (Messerschmidt et al. 1993) (Figure 2.14). Secondary structure modelling of the putative ascorbate oxidase proteins resulted in structures that consisted of three similar motifs (Figure 2.14).

2.4.3.2.2 β-Glucan Synthase-Like Homologs

Gene 7 matched to 49,683,929 – 49,687,869 bp in chromosome 8 of G19833, which is associated with an EST, and multiple annotations for similarity to β-glucan synthase-like or callose synthase genes from various other species. When the gene 7 sequence was BLASTed against NCBI, multiple matches to P. vulgaris hypothetical proteins were
<table>
<thead>
<tr>
<th>Gene 5</th>
<th>Phvul.006G011700</th>
<th>Glyma.20G051900</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ATGGGGTTGGAAGCAGTTTTGGTTTGGTGC- - - ATATGGTTGGGGCTGATACAATATTCA</td>
<td>ATGGGGTTGGAAGCAGTTTTGGTTTGGTGC- - - ATATGGTTGGGGCTGATACAATATTCA</td>
</tr>
<tr>
<td></td>
<td>Glyma.20G051900</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TTGTTGGAAACCTTGTTGAGGGAAATCAAATGGGCTTTCCAGGGCCACACTTATTAGGGCT</td>
<td>TTGCTTGGAAACCTTGTTGAGGGAAATCAAATGGGCTTTCCAGGGCCACACTTATTAGGGCT</td>
</tr>
<tr>
<td></td>
<td>Phvul.006G011700</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TTGGAGGAAGGGTGAGGCACTACAGGTTTGATGTGGAGTACATGATGAGAAAGCCAGAT</td>
<td>CTTGGAGGAAGGGTGAGGCACTACAGGTTTGATGTGGAGTACATGATGAGAAAGCCAGAT</td>
</tr>
<tr>
<td></td>
<td>Glyma.20G051900</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TGCTTGGAACACGTTGTGATGGGAATCAACGGCCAGTTTCCAGGCCCAACTTATTAGGGCT</td>
<td>TGCTTGGAACACGTTGTGATGGGAATCAACGGCCAGTTTCCAGGCCCAACTTATTAGGGCT</td>
</tr>
<tr>
<td></td>
<td>Phvul.006G011700</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GAAGTTGGTGACATTCTTGACATTGCTCTCACCAACAAGCTTTTCACTGAGGGAACTGTT</td>
<td>GAAGTTGGTGACATTCTTGACATTGCTCTCACCAACAAGCTTTTCACTGAGGGAACTGTT</td>
</tr>
<tr>
<td></td>
<td>Glyma.20G051900</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ATTCACTGGCATGGAATCAGACAGGTTGGAACTCCTTGGGCAGATGGAACTGCTGCTATC</td>
<td>ATTCACTGGCATGGAATCAGACAGGTTGGAACTCCTTGGGCAGATGGAACTGCTGCTATC</td>
</tr>
<tr>
<td></td>
<td>Phvul.006G011700</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ATTCACTGGCATGGAATCAGACAGGTTGGAACTCCTTGGGCAGATGGAACTGCTGCCATC</td>
<td>ATTCACTGGCATGGAATCAGACAGGTTGGAACTCCTTGGGCAGATGGAACTGCTGCCATC</td>
</tr>
<tr>
<td></td>
<td>Glyma.20G051900</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GGAACGTATTTCTACCATGGACACTATGGTATGCAAAGATCAGCAGGGTTGTATGGTTCA</td>
<td>GGAACGTATTTCTACCATGGACACTATGGTATGCAAAGATCAGCAGGGTTGTATGGTTCA</td>
</tr>
<tr>
<td></td>
<td>Phvul.006G011700</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CTGATAGTAGATTTGCCAAAGGGACAAAACGAACCGTTTCACTATGATGATGAGTTCAAC</td>
<td>CTGATAGTAGATTTGCCAAAGGGACAAAACGAACCGTTTCACTATGATGATGAGTTCAAC</td>
</tr>
<tr>
<td></td>
<td>Glyma.20G051900</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CTGATAGTAGATTTGCCAAAGGGACAAAACGAACCGTTTCACTATGATGATGAGTTCAAC</td>
<td>CTGATAGTAGATTTGCCAAAGGGACAAAACGAACCGTTTCACTATGATGATGAGTTCAAC</td>
</tr>
<tr>
<td></td>
<td>Phvul.006G011700</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CTTCTTCTAAGTGATTTGTGGCACACCAGCTCACATGAACAAGAAGTTGGTCTCTCTTCC</td>
<td>CTTCTTCTAAGTGATTTGTGGCACACCAGCTCACATGAACAAGAAGTTGGTCTCTCTTCC</td>
</tr>
<tr>
<td></td>
<td>Glyma.20G051900</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CTTCTTCTAAGTGATTTGTGGCACACCAGCTCACATGAACAAGAAGTTGGTCTCTCTTCC</td>
<td>CTTCTTCTAAGTGATTTGTGGCACACCAGCTCACATGAACAAGAAGTTGGTCTCTCTTCC</td>
</tr>
<tr>
<td></td>
<td>Phvul.006G011700</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CTACCATTCAAATGGATTTGGTGAAACCTCACAGCTCTCTTGTCATGAAATTGGGCAATTC</td>
<td>CTACCATTCAAATGGATTTGGTGAAACCTCACAGCTCTCTTGTCATGAAATTGGGCAATTC</td>
</tr>
<tr>
<td></td>
<td>Glyma.20G051900</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CTACCATTCAAATGGATTTGGTGAAACCTCACAGCTCTCTTGTCATGAAATTGGGCAATTC</td>
<td>CTACCATTCAAATGGATTTGGTGAAACCTCACAGCTCTCTTGTCATGAAATTGGGCAATTC</td>
</tr>
<tr>
<td></td>
<td>Phvul.006G011700</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AATTGTTCCCTTGCATCTAAATTCATCAACACAACCCTACCCCAATGCCAACTTAAAGG</td>
<td>AATTGTTCCCTTGCATCTAAATTCATCAACACAACCCTACCCCAATGCCAACTTAAAGG</td>
</tr>
<tr>
<td></td>
<td>Glyma.20G051900</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AATTGTTCCCTTGCATCTAAATTCATCAACACAACCCTACCCCAATGCCAACTTAAAGG</td>
<td>AATTGTTCCCTTGCATCTAAATTCATCAACACAACCCTACCCCAATGCCAACTTAAAGG</td>
</tr>
<tr>
<td></td>
<td>Phvul.006G011700</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GGTGAAGAATGTGCACCTCAAATTCTTCATGTAGAGCCAAACAAGACCTACAGGATCAGA</td>
<td>GGTGAAGAATGTGCACCTCAAATTCTTCATGTAGAGCCAAACAAGACCTACAGGATCAGA</td>
</tr>
<tr>
<td></td>
<td>Glyma.20G051900</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GGTGAAGAATGTGCACCTCAAATTCTTCATGTAGAGCCAAACAAGACCTACAGGATCAGA</td>
<td>GGTGAAGAATGTGCACCTCAAATTCTTCATGTAGAGCCAAACAAGACCTACAGGATCAGA</td>
</tr>
<tr>
<td></td>
<td>Phvul.006G011700</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AATTGTTCTCTGGCAGCTAAATTCATTAACACCACTTTACCCGAGTGCCAATTCAGAGG</td>
<td>AATTGTTCTCTGGCAGCTAAATTCATTAACACCACTTTACCCGAGTGCCAATTCAGAGG</td>
</tr>
<tr>
<td></td>
<td>Glyma.20G051900</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AATTGTTCTCTGGCAGCTAAATTCATTAACACCACTTTACCCGAGTGCCAATTCAGAGG</td>
<td>AATTGTTCTCTGGCAGCTAAATTCATTAACACCACTTTACCCGAGTGCCAATTCAGAGG</td>
</tr>
<tr>
<td></td>
<td>Phvul.006G011700</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GGTGAAGAATGTGCACCTCAAATTCTTCATGTAGAGCCAAACAAGACCTACAGGATCAGA</td>
<td>GGTGAAGAATGTGCACCTCAAATTCTTCATGTAGAGCCAAACAAGACCTACAGGATCAGA</td>
</tr>
<tr>
<td></td>
<td>Glyma.20G051900</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GGTGAAGAATGTGCACCTCAAATTCTTCATGTAGAGCCAAACAAGACCTACAGGATCAGA</td>
<td>GGTGAAGAATGTGCACCTCAAATTCTTCATGTAGAGCCAAACAAGACCTACAGGATCAGA</td>
</tr>
<tr>
<td></td>
<td>Phvul.006G011700</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AATTGTTCTCTGGCAGCTAAATTCATTAACACCACTTTACCCGAGTGCCAATTCAGAGG</td>
<td>AATTGTTCTCTGGCAGCTAAATTCATTAACACCACTTTACCCGAGTGCCAATTCAGAGG</td>
</tr>
<tr>
<td></td>
<td>Glyma.20G051900</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AATTGTTCTCTGGCAGCTAAATTCATTAACACCACTTTACCCGAGTGCCAATTCAGAGG</td>
<td>AATTGTTCTCTGGCAGCTAAATTCATTAACACCACTTTACCCGAGTGCCAATTCAGAGG</td>
</tr>
<tr>
<td></td>
<td>Phvul.006G011700</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GGTGAAGAATGTGCACCTCAAATTCTTCATGTAGAGCCAAACAAGACCTACAGGATCAGA</td>
<td>GGTGAAGAATGTGCACCTCAAATTCTTCATGTAGAGCCAAACAAGACCTACAGGATCAGA</td>
</tr>
<tr>
<td></td>
<td>Glyma.20G051900</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GGTGAAGAATGTGCACCTCAAATTCTTCATGTAGAGCCAAACAAGACCTACAGGATCAGA</td>
<td>GGTGAAGAATGTGCACCTCAAATTCTTCATGTAGAGCCAAACAAGACCTACAGGATCAGA</td>
</tr>
<tr>
<td></td>
<td>Phvul.006G011700</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AATTGTTCTCTGGCAGCTAAATTCATTAACACCACTTTACCCGAGTGCCAATTCAGAGG</td>
<td>AATTGTTCTCTGGCAGCTAAATTCATTAACACCACTTTACCCGAGTGCCAATTCAGAGG</td>
</tr>
<tr>
<td></td>
<td>Glyma.20G051900</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AATTGTTCTCTGGCAGCTAAATTCATTAACACCACTTTACCCGAGTGCCAATTCAGAGG</td>
<td>AATTGTTCTCTGGCAGCTAAATTCATTAACACCACTTTACCCGAGTGCCAATTCAGAGG</td>
</tr>
</tbody>
</table>
Figure 2.12 Alignment and annotation of ascorbate oxidase coding sequences. Exons are indicated in alternating dark and light coloured boxes. Gene 5 is indicated in green, Gene 16 is indicated in purple, Phvul.006G011700 is indicated in red, and Glyma.20G051900 is indicated in blue.
<table>
<thead>
<tr>
<th>Protein</th>
<th>Sequence</th>
<th>Protein</th>
<th>Sequence</th>
<th>Protein</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gene5</td>
<td>MGLKAVLWCINLGLIQYSVGGRVRHYPKEVYMRRKPCCLNHVINGQFFPGPTIRA</td>
<td>Phvul.006G011700</td>
<td>MGLKAVLWCINLGLIQYSVGGRVRHYPKEVYMRRKPCCLNHVINGQFFPGPTIRA</td>
<td>Glyma.20G051900</td>
<td>MGLKAVLWCINLGLIQYSVGGRVRHYPKEVYMRRKPCCLNHVINGQFFPGPTIRA</td>
</tr>
<tr>
<td>1ASP_A</td>
<td>SQRHYKKWEVYMFWAPNCNENIVMGINGQFPGPTIRA</td>
<td>Phvul.006G011700</td>
<td>EVGDLNALTNLKNTVGVHIWHRGVRQVTPWADGTAISGCAINPGETTFHYRFIVDPR</td>
<td>Glyma.20G051900</td>
<td>EVGDLNALTNLKNTVGVHIWHRGVRQVTPWADGTAISGCAINPGETTFHYRFIVDPR</td>
</tr>
<tr>
<td></td>
<td>**</td>
<td>1ASP_A</td>
<td>NADDVVTELNLKHTEGVHIWHRGILQRGTPWADGTAISGCAINPGETTFHYRFIVDPR</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>Gene5</td>
<td>GTYFYHGHHYGMQRSAGLYGSLIVDLPKGQNEPFHYDDEFNLLLSDLWHTSSHEQEVGLSS</td>
<td>Phvul.006G011700</td>
<td>GTYFYHGHHYGMQRSAGLYGSLIVDLPKGQNEPFHYDDEFNLLLSDLWHTSSHEQEVGLSS</td>
<td>Glyma.20G051900</td>
<td>GTYFYHGHHYGMQRSAGLYGSLIVDLPKGQNEPFHYDDEFNLLLSDLWHTSSHEQEVGLSS</td>
</tr>
<tr>
<td>1ASP_A</td>
<td>GTYYFHYGHHYGMQRSAGLYGSLIVDLPKGQNEPFHYDDEFNLLLSDLWHTSSHEQEVGLSS</td>
<td>Phvul.006G011700</td>
<td>GTYYFHYGHHYGMQRSAGLYGSLIVDLPKGQNEPFHYDDEFNLLLSDLWHTSSHEQEVGLSS</td>
<td>Glyma.20G051900</td>
<td>GTYYFHYGHHYGMQRSAGLYGSLIVDLPKGQNEPFHYDDEFNLLLSDLWHTSSHEQEVGLSS</td>
</tr>
<tr>
<td></td>
<td>**</td>
<td>1ASP_A</td>
<td>GTYYFHYGHHYGMQRSAGLYGSLIVDLPKGQNEPFHYDDEFNLLLSDLWHTSSHEQEVGLSS</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>Gene5</td>
<td>LPFHKWIGEFPQSLINGRQVENSLAASFINTTLPCEFQFEGGCAFPQLLHVEMPNKYIR</td>
<td>Phvul.006G011700</td>
<td>LPFHKWIGEFPQSLINGRQVENSLAASFINTTLPCEFQFEGGCAFPQLLHVEMPNKYIR</td>
<td>Glyma.20G051900</td>
<td>LPFHKWIGEFPQSLINGRQVENSLAASFINTTLPCEFQFEGGCAFPQLLHVEMPNKYIR</td>
</tr>
<tr>
<td>1ASP_A</td>
<td>LPFHKWIGEFPQSLINGRQVENSLAASFINTTLPCEFQFEGGCAFPQLLHVEMPNKYIR</td>
<td>Phvul.006G011700</td>
<td>LPFHKWIGEFPQSLINGRQVENSLAASFINTTLPCEFQFEGGCAFPQLLHVEMPNKYIR</td>
<td>Glyma.20G051900</td>
<td>LPFHKWIGEFPQSLINGRQVENSLAASFINTTLPCEFQFEGGCAFPQLLHVEMPNKYIR</td>
</tr>
<tr>
<td></td>
<td>**</td>
<td>1ASP_A</td>
<td>LPFHKWIGEFPQSLINGRQVENSLAASFINTTLPCEFQFEGGCAFPQLLHVEMPNKYIR</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>Gene5</td>
<td>GNYVTPFVVDIYSGYESYSUVRRTDQPKEKTGQPQQLD</td>
<td>Phvul.006G011700</td>
<td>GNYVTPFVVDIYSGYESYSUVRRTDQPKEKTGQPQQLD</td>
<td>Glyma.20G051900</td>
<td>GNYVTPFVVDIYSGYESYSUVRRTDQPKEKTGQPQQLD</td>
</tr>
<tr>
<td>1ASP_A</td>
<td>GNYVTPFVVDIYSGYESYSUVRRTDQPKEKTGQPQQLD</td>
<td>Phvul.006G011700</td>
<td>GNYVTPFVVDIYSGYESYSUVRRTDQPKEKTGQPQQLD</td>
<td>Glyma.20G051900</td>
<td>GNYVTPFVVDIYSGYESYSUVRRTDQPKEKTGQPQQLD</td>
</tr>
<tr>
<td></td>
<td>**</td>
<td>1ASP_A</td>
<td>GNYVTPFVVDIYSGYESYSUVRRTDQPKEKTGQPQQLD</td>
<td>**</td>
<td>**</td>
</tr>
</tbody>
</table>

**Type III**

<table>
<thead>
<tr>
<th>Protein</th>
<th>Sequence</th>
<th>Protein</th>
<th>Sequence</th>
<th>Protein</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gene5</td>
<td>GTYFYHGHHYGMQRSAGLYGSLIVDLPKGQNEPFHYDDEFNLLLSDLWHTSSHEQEVGLSS</td>
<td>Phvul.006G011700</td>
<td>GTYFYHGHHYGMQRSAGLYGSLIVDLPKGQNEPFHYDDEFNLLLSDLWHTSSHEQEVGLSS</td>
<td>Glyma.20G051900</td>
<td>GTYFYHGHHYGMQRSAGLYGSLIVDLPKGQNEPFHYDDEFNLLLSDLWHTSSHEQEVGLSS</td>
</tr>
<tr>
<td>1ASP_A</td>
<td>GTYFYHGHHYGMQRSAGLYGSLIVDLPKGQNEPFHYDDEFNLLLSDLWHTSSHEQEVGLSS</td>
<td>Phvul.006G011700</td>
<td>GTYFYHGHHYGMQRSAGLYGSLIVDLPKGQNEPFHYDDEFNLLLSDLWHTSSHEQEVGLSS</td>
<td>Glyma.20G051900</td>
<td>GTYFYHGHHYGMQRSAGLYGSLIVDLPKGQNEPFHYDDEFNLLLSDLWHTSSHEQEVGLSS</td>
</tr>
<tr>
<td></td>
<td>**</td>
<td>1ASP_A</td>
<td>GTYFYHGHHYGMQRSAGLYGSLIVDLPKGQNEPFHYDDEFNLLLSDLWHTSSHEQEVGLSS</td>
<td>**</td>
<td>**</td>
</tr>
</tbody>
</table>

**Type I**

<table>
<thead>
<tr>
<th>Protein</th>
<th>Sequence</th>
<th>Protein</th>
<th>Sequence</th>
<th>Protein</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gene5</td>
<td>ISSTTSLASLNLAISRSTSVINGQYQLSFFSSECTYVIAICCSSNHKLIVEAD</td>
<td>Phvul.006G011700</td>
<td>ISSTTSLASLNLAISRSTSVINGQYQLSFFSSECTYVIAICCSSNHKLIVEAD</td>
<td>Glyma.20G051900</td>
<td>ISSTTSLASLNLAISRSTSVINGQYQLSFFSSECTYVIAICCSSNHKLIVEAD</td>
</tr>
<tr>
<td>1ASP_A</td>
<td>ISSTTSLASLNLAISRSTSVINGQYQLSFFSSECTYVIAICCSSNHKLIVEAD</td>
<td>Phvul.006G011700</td>
<td>ISSTTSLASLNLAISRSTSVINGQYQLSFFSSECTYVIAICCSSNHKLIVEAD</td>
<td>Glyma.20G051900</td>
<td>ISSTTSLASLNLAISRSTSVINGQYQLSFFSSECTYVIAICCSSNHKLIVEAD</td>
</tr>
<tr>
<td></td>
<td>**</td>
<td>1ASP_A</td>
<td>ISSTTSLASLNLAISRSTSVINGQYQLSFFSSECTYVIAICCSSNHKLIVEAD</td>
<td>**</td>
<td>**</td>
</tr>
</tbody>
</table>

**Type II**

<table>
<thead>
<tr>
<th>Protein</th>
<th>Sequence</th>
<th>Protein</th>
<th>Sequence</th>
<th>Protein</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gene5</td>
<td>GNYVTPFVVDIYSGYESYSUVRRTDQPKEKTGQPQQLD</td>
<td>Phvul.006G011700</td>
<td>GNYVTPFVVDIYSGYESYSUVRRTDQPKEKTGQPQQLD</td>
<td>Glyma.20G051900</td>
<td>GNYVTPFVVDIYSGYESYSUVRRTDQPKEKTGQPQQLD</td>
</tr>
<tr>
<td>1ASP_A</td>
<td>GNYVTPFVVDIYSGYESYSUVRRTDQPKEKTGQPQQLD</td>
<td>Phvul.006G011700</td>
<td>GNYVTPFVVDIYSGYESYSUVRRTDQPKEKTGQPQQLD</td>
<td>Glyma.20G051900</td>
<td>GNYVTPFVVDIYSGYESYSUVRRTDQPKEKTGQPQQLD</td>
</tr>
<tr>
<td></td>
<td>**</td>
<td>1ASP_A</td>
<td>GNYVTPFVVDIYSGYESYSUVRRTDQPKEKTGQPQQLD</td>
<td>**</td>
<td>**</td>
</tr>
</tbody>
</table>
Figure 2.13 Alignment and annotation of L-ascorbate oxidase amino acid sequences. Amino acid sequences were aligned for predicted ascorbate oxidase like genes from *P. vulgaris* varieties OAC Rex (gene 5) and G19833 (Phvul.006G011700; Schmutz et al. 2014), a predicted *G. max* ascorbate oxidase gene (Glyma.20G051900; Schmutz et al. 2010) and an ascorbate oxidase protein from *Cucurbita pepo var. melopepo*, for which the secondary structured was resolved by X-ray (1ASP_A; Messerschmidt et al. 1992). Type 1, 2 and 3 multicopper oxidase domains are outlined and labelled. A copper binding site is highlighted in grey.
Figure 2.14 Secondary structure models of OAC Rex and G19833 L-ascorbate oxidase proteins. Secondary structure models for Gene 5 (A) from contig1701, G19833 gene annotation Phvul.006G011700 (B) and G. max gene annotation Glyma.20G051900 (C) were produced using Phyre2 (Kelley and Sternberg 2009). Ascorbate oxidase type I, II and III domains are indicated by circles. All models were created using the X-ray structure of a zucchini ascorbate oxidase (D) as a template (Messerschmidt et al. 1993)
**Table 2.12** Phyre2 secondary structure modelling (Kelley and Sternberg 2009) for gene 5, Phvul.006G011700, Glyma.20G051900 and 1ASP_A

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Amino Acids</th>
<th>Template</th>
<th>Template Characteristic</th>
<th>Similarity to Template</th>
<th>Confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gene 5</td>
<td>11-408</td>
<td>C1asqB_</td>
<td>Ascorbate oxidase</td>
<td>57%</td>
<td>100%</td>
</tr>
<tr>
<td>Phvul.006G011700</td>
<td>22-570</td>
<td>C1asqB_</td>
<td>Ascorbate oxidase</td>
<td>69%</td>
<td>100%</td>
</tr>
<tr>
<td>Glyma.20G051900</td>
<td>23-572</td>
<td>C1asqB_</td>
<td>Ascorbate oxidase</td>
<td>70%</td>
<td>100%</td>
</tr>
<tr>
<td>1ASP_A</td>
<td>1-552</td>
<td>C1asqB_</td>
<td>Ascorbate oxidase</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>
identified, as well as to callose synthase 3-like predictions from other species, including *Medicago truncatula* (data not shown). BLAST against the pvGEA EST database identified several β-glucan synthase-like gene matches on several chromosomes, including chromosome 8 (data not shown). However, none of these matches corresponded to 49,683,929 – 49,687,869 bp location in chromosome 8.

Since there was no transcript or gene annotation within G19833 where gene 7 matched, a ~12 kb sequence (49,679,901 – 49,691,730 bp on chromosome 8) was selected surrounding the match location and FGeneSH analysis identified 4 genes. One gene (G2) corresponded to the gene 7 match location within the G19833 sequence (49683929 – 49687948 bp; Figure 2.6). Gene 7 and G2 gene sequences were then BLASTed against the *G. max* genome sequence in Phytozome (Schmutz *et al.* 2010), which identified homology to gene annotation Glyma.08GG361500 on chromosome 8 (47,300,299 – 47,324,252 bp).

Alignment of gene 7, G2 and Glyma08G361500 DNA sequences showed high similarity between gene 7 and G2, but low similarity to Glyma08G361500 (Table 2.12). Alignment of the coding sequences indicated low similarity between all three genes (Table 2.11). When the protein sequences were aligned, the highest similarity to Glyma08G361500 was seen for both genes, however, they were still <50% similar (Appendix 1; Table 2.11).

When the protein sequences were BLASTed at InterPro, callose synthase family related signatures were identified in G2 and Glyma08G361500. G2 also contained three transmembrane domains which corresponded to the location of similar domains in Glyma08G361500 (Appendix 1). Glyma08G361500 was classified as a callose
Figure 2.15 Secondary structure model for predicted β glucan synthase-like genes. Phyre 2 (Kelley and Sernberg 2009) secondary structure models were created for *P. vulgaris* gene 7 (A) from OAC Rex contig1701, and G2 (B) from G19833, and *G. max* gene annotation Glyma.08G361500 (C).

Table 2.13 Phyre2 secondary structure modelling (Kelley and Sternberg 2009) for gene 7, G2, and Glyma.08G361500

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Amino Acids Modelled</th>
<th>Template</th>
<th>Template Characteristic</th>
<th>Similarity to Template</th>
<th>Confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gene 7</td>
<td>36-98</td>
<td>D1h2vc3</td>
<td>Alpha-alpha superhelix</td>
<td>17%</td>
<td>52.3%</td>
</tr>
<tr>
<td>G2</td>
<td>105-225</td>
<td>C4gmjE_</td>
<td>RNA binding protein</td>
<td>17%</td>
<td>54.4%</td>
</tr>
<tr>
<td>Glyma.08G361500</td>
<td>66-449</td>
<td>C4hg6A_</td>
<td>Cellulose synthase subunit a</td>
<td>14%</td>
<td>96.8%</td>
</tr>
</tbody>
</table>
synthase based on the whole peptide sequence, and as a glycosyl transferase family 48 (glucan synthase) member based on amino acids 1,059-1,815. Domains identified in Glyma08G361500 included a vacuolar protein sorting-associated protein Vta1/calllose synthase N-terminal domain (amino acids 45-175), a 1,3-β-glucan synthase subunit FKS1-like domain (amino acids 321-436), two large and one small cytoplasmic domains and two transmembrane domains (Appendix 1). No family or domain signatures were predicted for gene 7 (Appendix 1).

When Phyre2 was used to predict protein secondary structure for gene 7, G2 and Glyma08G361500, there was very low similarity to previously resolved proteins (Figure 2.15; Table 2.13). Secondary structure modelling for gene 7 and G2 produced low confidence (<55%) models for small sections of the protein sequences (Figure 2.15A-B; Table 2.12). Glyma.08G361500 was modelled using a cellulose synthase subunit α, and was a 96.8% confident model, with 14% similarity to the model template (c4hg6A_) (Figure 2.15C; Table 2.13).

2.4.3.2.3 Hypothetical Gene Annotations

Gene 6 matched to a hypothetical protein annotation at 49,682,264 – 49,683,238 bp on chromosome 8 (Figure 2.5-6; Table 2.6). The matched region in G19833 was associated with RNA seq data obtained from the nodules and roots, EST and transcript information related to the hypothetical gene annotation Phvul.008G191100.1. The coding sequence predicted in FGeneSH for gene 6 corresponded to a region beside this annotation, but still corresponded to low level RNA seq expression data (49,682,789 – 49,682,959 bp). Neither the coding sequence associated with Phvul.008G191100.1,
nor the coding sequence of gene 6 had high similarity to known protein domains when BLASTed against the NCBI database.

When the gene 6 and Phvul.008G191100 DNA sequences were BLASTed against the G. max sequence in Phytozome (Schmutz et al. 2010), no homologs were found. Alignment of the DNA sequences for gene 6 and Phvul.008G191100.1 indicated that gene 6 was 97% similar to Phvul.008G191100 (Table 2.1). However, alignment of coding and protein sequences indicated low similarity (8% and 5% respectively) between gene 6 and Phvul.008G191100 (Appendix 2; Table 2.11).

Analysis with InterPro identified no protein family or domain signatures within the protein sequences. Phyre2 modelling (Appendix 3) produced a low confidence (15.2%) model that had low similarity (33%) to an iron regulated transcription activator (c4ImgD) for gene 5 (Appendix 4). The model produced for Phvul.008G191100 also had low confidence (32%) (Appendix 4).

2.4.4 Isolation and Amplification of CGs

Amplification of bands from genomic OAC Rex DNA using the designed primers for CGs CBB2 and CBB3 (Table 2.14) produced multiple bands ranging from 850 – 5,000 bp (Figure 2.16 A-D). To reduce amplification of non-specific bands, BiBAC2 library clone 21 DNA was used as the template, which produced a single band when the same primers were used (Figure 2.16 E-H). The PCR fragments were cloned into E. coli (Figure 2.17) and 1 kb of the forward (5’) and reverse (3’) ends of the clones were sequenced. BLAST of the 5’ and 3’ end sequences for all sequenced clones matched to contig1455 at 2,236-3,401 bp and 4,526-5,566 bp respectively (Figure 2.18). When the contig1455 sequence surrounding the 5’ and 3’ 1 kb sequences was analyzed using
FGeneSH and GeneMark, three genes were predicted (discussed in section 2.4.1.3.2; Figure 2.4; Table 2.5). Additionally, when contig1455 was compared to the region surrounding CGs CBB2 and CBB3, it was 99% similar to each region, with small (<100 bp) sequence differences, so the gene was included as a CG (CBB4).
Table 2.14 Primers used to amplify candidate genes CBB2 and CBB3 from OAC Rex DNA

<table>
<thead>
<tr>
<th>Primer</th>
<th>Forward Sequence</th>
<th>Reverse Sequence</th>
<th>Band Size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBBS2-14</td>
<td>CATTTCTCACCACCTCACCTACCCC</td>
<td>GAAAGTCCTTTGAGGTAGTGAGTG</td>
<td>2,553</td>
</tr>
<tr>
<td>CBBS2-15</td>
<td>CACCTATAATACCTATGCGCAACCC</td>
<td>GTGAGTGAAGTACTGAGTACTGAGTACTGG</td>
<td>2,808</td>
</tr>
<tr>
<td>CBBS3-8</td>
<td>CACCTATAATACCTATGCGCAACCCC</td>
<td>GTTGGAGAATTGTTAGGGTGTGAGGTTTTT</td>
<td>2,391</td>
</tr>
<tr>
<td>CBBS3-9</td>
<td>CACATTTCTCACCACCTACCTCACCC</td>
<td>GTTTTGGGGTTGGGAATCTTTGGTG</td>
<td>2,204</td>
</tr>
</tbody>
</table>

Figure 2.16 Candidate gene PCR amplification from OAC Rex genomic DNA and BiBAC2 library DNA sources. Amplification of candidate genes from OAC Rex genomic DNA exhibited non-specific binding A) CBBS3-8, B) CBBS3-9, C) CBBS2-14, and D) CBBS2-15. Amplification of candidate genes from OAC Rex BiBAC2 library clone 021 E) CBB3-8, F) CBB3-9, G) CBB2-14, and H) CBB2-15 produced single bands. No bands were produced when water was used as a negative control (NC)
Figure 2.17 Common bacterial blight resistance candidate gene clones. Amplification of ~3.5 kb band using candidate gene primers for CBBS3-8 (A), CBBS3-9 (B), CBBS2-14 (C), CBBS2-15 from *Escherichia coli* clones. *E. coli* clones contained DNA amplified from OAC Rex BiBAC2 clone 021 using primers CBBS3-8 and CBBS3-9 designed for candidate gene CBB3, and CBBS2-14 and CBBS2-15 designed for candidate gene CBB2.

Figure 2.18 Location of clone sequence match within contig1455. Sequence of 1kb flanking ends of CBB candidate gene *Escherichia coli* clones matched to CBB4 5’ and 3’ ends when BLAST search was performed. The 1 kb sequences matched to 2,236-3,401 bp and 4,526-5,566 bp within the contig1455 sequence, and are indicated with a blue box.
2.5 Discussion

2.5.1 CG Characteristics

2.5.1.1 R Gene Characteristics of CGs

Several characteristics that are exhibited by CBB resistance CGs CBB1, CBB2, CBB3 and CBB4 have previously been reported in literature. The organization of multiple CGs and genes with potential disease resistance functions on contig1701 (Figure 2.2B and 2.3; Table 2.4, Table 2.7) are typical of resistance gene family clusters (Michelmore and Meyers 1998; Graham et al. 2002; Meyers et al. 2003; Ameline-Torregrosa et al. 2008). This clustering was accompanied by high similarity between two of the CGs on contig1701 (CBB2 and CBB3) and the single CG identified on contig1455 (CBB4) (Figure 2.2B). High similarity between resistance genes has been frequently reported, especially in gene families where genes are present in close proximity (Anderson et al. 1997; Meyers et al. 2003; Song et al. 2003). Both characteristics are thought to be the result of frequent recombination events, which result in high variability in the pathogen-interaction region (LRR) and may change pathogen recognition capabilities (Anderson et al. 1997; Meyers et al. 1998; Song et al. 2003; Leister, 2004; Bakker et al. 2006).

Another important feature of the CGs is the presence of LRR domains, which are found in many resistance related protein classes such as NBS-LRR, eLRR, RLKs and RLPs (Meyers et al. 2003; Wang et. al. 2008b). Although it was not reflected in the secondary structure model (Figure 2.11), LRR motifs were identified in the protein sequence for each of the CG proteins (Figure 2.10). In addition, CBB2-4 contained an
LDL motif, but didn’t contain end motifs similar to those described for CC-NBS-LRR and TIR-NBS-LRR proteins in Arabidopsis (Meyers et al. 2003).

Although there are similarities between the CBB2, CBB3 and CBB4 gene sequences (Table 2.9), they were sufficiently variable to encode proteins which were even more variable (Figure 2.9). Both the overall protein sizes and the numbers of LRR motifs identified within the sequences varied (Figure 2.10; Table 2.6), resulting in similarities between proteins as low as 11%, in the case of CBB1 and CBB2 (Table 2.9). However, an overall consensus sequence (LxxxLxLxxLxL) was still identified, that was applicable to all four predicted CG proteins (Figure 2.10). The finding that the motif is not an exact match to LRR motifs reported previously (LxxLxLxxL), is consistent with reports of high variability in LRR motif organization in NBS-LRR genes, within and between species (Kajava 1998; Monosi et al. 2004; Zhou et al. 2004b)

2.5.1.2 Unique/Unusual Characteristics of CGs

Although the characteristics discussed above provide evidence to support the hypothesis that they play a role in disease resistance, there are some differences which indicate otherwise. When the protein sequences were analyzed both manually and via predictive programs, NBS, CC, TIR, kinase, transmembrane and eLRR motifs were not identified, and no signal peptide was indicated for any of the CGs. The sizes of the loops between LRR motifs were more variable than expected when compared to the LRR domains of NBS-LRRs and PGIPs (Di Matteo et al. 2003; Meyer et al. 2003). Additionally, the secondary structure predicted for the CGs failed to support the manual prediction of LRR domains (Figure 2.11). It was expected that the secondary structure of the CGs would resemble the resolved secondary structures for eLRR protein PGIP,
and receptor like kinase RPK2 (Di Matteo et al. 2003; Song et al. 2014). These proteins exhibit LRRs that align to form a stacked backbone, with the inter-LRR sequence forming loops. Since there are few plant pathogen-resistance LRR proteins modelled, it might be assumed that the failure to model CBB1-4 protein sequences into the expected LRR secondary structure model is due to the lack of a good reference structure for the proteins analysed with the modelling program. However, when a PGIP was used as the model template, the expected secondary structure model still could not be obtained. This evidence indicates that the CBB1-4 may not be LRR proteins, and originate from another classification of protein.

2.5.1.3 Retrotransposon Characteristics of CGs

Retrotransposons are characterized by the ability to make copies of themselves using RNA as the transposition medium, reverse transcribe and re-insert into the genome (Wicker et al. 2007). BLAST search and secondary modelling of the CG proteins identified low similarities to viral integrases, reverse transcriptase and DNA/RNA polymerase proteins and GAG/POL/ENV polyproteins (see section 2.4.3.1.3; Table 2.11). These are all proteins found in various types of retrotransposons, although the similarity to the templates was low (Table 2.11), with percent similarity reaching no higher than 22% for an 116 amino acid sequence for CBB4. This is likely to be an indication of the lack of information about the secondary structure of plant retrotransposon polyproteins.

When the CGs and other predicted genes from the area surrounding the CGs were BLASTed against the G19833 sequence, many of the genes matched to areas within the G19833 genome that were marked as retrotransposons (Figure 2.3; Table
In several cases, including for CBB2 and CBB3, there was also RNASeq data associated with the retrotransposons, indicating that they were active (Table 2.7). Retrotransposons have been shown to integrate preferentially into genic regions (Miyao et al. 2003; Le et al. 2007). It is therefore not surprising that multiple gene predictions associated with retrotransposons were present in a contig1701. Since the contig was selected based on its presumed association with disease resistance marker PV-ctt001, the presence of multiple possible resistance-related genes and retrotransposons indicate the likelihood that part of a multigene family is represented in the sequence.

Retrotransposons have previously been shown to be involved in disease resistance by interacting with resistance genes. When transposable elements transpose into the location of a resistance gene, they may disrupt gene function which might lead to reduced or abolished resistance (Song et al. 1997; Hernández-Pinzón et al. 2009). Gene 7 may be an example of this, since it matched to the site of a putative callose synthase gene on Chromosome 8 in the G19833 sequence assembly, but was shown to be similar to only a portion of the large gene expressed in G. max (Figure 2.6, 2.15 and 2.16; Table 2.7 and 2.12). Studies have also shown that transposable elements transpose in front of or between two genes and cause a chimeric protein to be formed (Kashkush et al. 2003; Frost et al. 2004). Transposition can also change promoter function and epigenetic controls (Hayashi et al. 2009; McDowell and Meyers 2013; Tsuchiya and Eulgem 2013). Considering that transposable elements were reported to comprise 45% of the G19833 sequence (Schmutz et al. 2014), the likelihood of an interaction between transposable elements and resistance genes is high, and is an area to expand research into pathogen resistance.
2.5.2 Additional CGs on Contig1701

Since the predicted CGs CBB1-4 may be retrotransposons, the characterization of three other genes within contig1701, which have potential disease resistance roles, were explored. The genes were classified as ascorbate oxidase (gene 5), callose synthase (gene 7) and hypothetical gene (gene 6) (see section 2.4.3.2; Appendices 1-4; Figure 2.3 and 2.12-2.15; Table 2.7 and 2.11). Only gene 5, which matched an ascorbate oxidase gene in G19833 chromosome 6, was likely to be a functional gene based on the gene and protein analyses performed (see section 2.4.3.2).

The other two potential genes (callose synthase, hypothetical) either were not very similar in sequence or secondary structure to previously identified genes of their classification (Appendix 1-4; Figure 2.15; Table 2.11 and 2.13), or they were similar to other uncharacterized genes and so were uninformative (Appendices 2-4). Gene 6 and gene 7 may be fragments of genes that were produced by recombination events that are often associated with the evolution of resistance genes (Michelmore and Meyers 1998; Meyers et al. 2005; Liu et al. 2007; Joshi and Nayak 2013). The association between multiple predicted genes within the contig1701 with retro-element annotations within the G19833 (Figure 2.3 and 2.5; Table 2.7) sequence suggests that transposable element activity caused gene fragmentation in this stretch of the bean genome.

2.5.2.1 Ascorbate Oxidases and Disease Resistance

Ascorbate oxidases (AOs) are cell wall bound proteins which, are members of the multi-copper oxidase family and are involved in binding and oxidizing ascorbic acid (AA) in the apoplast (Messerschmidt et al. 1989). AA is involved primarily in providing protection against reactive oxygen species (ROS) which are involved in the modulation
of stress (Smirnoff 2000) and defense responses (Pastori et al. 2003; Pavet et al. 2005) have been shown to have a role in stress and pathogen resistance in various plant species (Pavet et al. 2005; Pignocchi et al. 2006; Sanmartin et al. 2007; Fotopoulos et al. 2008).

The presence of a predicted ascorbate oxidase gene within contig1701 is suggestive of a potential role in resistance to CBB. The involvement of AA in protection against ROS indicates that gene 5, as an AO, would regulate PCD associated with cell death (Pavet et al. 2005), which is not a characteristic of the CBB resistance reaction. However, the truncated domain predictions for gene 5 (Figure 2.13) may cause incomplete binding and oxidation to occur, which has the potential to affect the accumulation of ROS and the resultant hypersensitive response. Since this gene was characterized at the bioinformatics level only, further studies would show more definitively whether gene 5 is a functional ascorbate oxidase.

2.5.3 CG Location within Bean Genome

When the genes identified in the contig1701 sequence assembly were BLASTed against the G19833 sequence, there were high numbers of significant matches (> e -63) to many chromosomes for most of the predicted genes. The best matches (Figure 2.6; Table 2.7) were to chromosomes 6 (5,817,952 – 5,839,661 bp) and 8 (49,682,264 – 49,698,171 bp). Since the contig1701 sequence was produced from BiBAC clones that were selected based on the presence of disease resistance marker PV-cttt001, it was expected that the contig and the genes predicted on it would match best to the area surrounding its location (517,577-517,741 bp) on chromosome 4 (Tar’an et al. 2001; Perry 2010; Perry et al. 2013). This is also the location of a large cluster of NBS-LRR
genes at the end of chromosome 4, which is associated with a variety of resistance-related genes (Geffroy et al. 2009; Schmutz et al. 2014)

An explanation for this discrepancy is that the contig1701 sequence was assembled using sequence data which originates from both Chromosome 6 and 8, and assembled into one continuous sequence by human or program error. However, the junction between the section that matched to chromosome 6 (contig1701-1) and the section that matched to chromosome 8 (contig1701-2) was confirmed by PacBio reads. Additionally, the creation of a new reference assembly which had read coverage of 8-2,037 per 1 kb, indicating that it is likely that the sequence is continuous for the 1-35 kkb length of contig1701 (see section 2.4.1.2.1; Figure 2.2A). The presence of the PV-ctt001 marker within the BiBAC2 library clone from which CG CBB4 was isolated (Rex021F10; Figure 2.1, 2.16 and 2.17) indicates that the contig1701 sequence is from chromosome 4. This suggests the possibility that sequence rearrangement has relocated sections of chromosomes 6 and 8 to chromosome 4 in the process of integrating P. acutifolius DNA into the common bean genome. Considering the apparent presence of retrotransposons in both the contig1455 and contig1701 sequence assemblies discussed above, this is likely to have been facilitated by transposable element action, however, further studies are required for confirmation.

The difficulty in identifying a location for the full contig1701 sequence within the OAC Rex genome assembly is that it is located within a region that has yet to be completely assembled. Perry et al. (2013) compared the location of the PV-ctt001 marker and its surrounding area within OAC Rex with the G19833 sequence (Schmutz et al. 2014). Although there was high conservation in the genomic region between
varieties, there was a 12 bp sequence difference between PV-ctt001 in OAC Rex, and the corresponding location in G19833. There was also a large section (~200 kb) of the G19833 sequence which did not match any of the assembled contigs within the OAC Rex assembly. Since CBB resistance in OAC Rex resulted from an interspecific cross between *P. vulgaris* and *P. acutifolius* (Parker 1985; Scott and Michaels 1992; Michaels *et al.* 2006), it is possible that the area in question has higher amounts of the *P. acutifolius* DNA, which is not present in the G19833 genome. However, the genomic comparisons performed by Perry *et al.* (2013) indicated that differences between OAC Rex and G19833 sequences were limited to single genes or up to 3 genes, as opposed to a large stretch of highly different sequence.

The matching patterns observed by Perry *et al.* (2013) are similar to what was observed in the current study for the comparison between section contig1701-1 and the the G19833 sequence. The genes found in contig1701-1 (genes 1-5) matched to a segment of G19833 chromosome 6 that contained all 5 genes in the same orientations. However, an additional gene was predicted in the G19833 sequence between contig1701 genes 3 and 4 (Figure 2.6). In contrast, there were multiple differences in the number and size of genes identified in section contig1701-2 compared to those identified in the associated region within the G19833 sequence assembly (genes 6-10). This may be a reflection of the presence of multiple retrotransposon like elements in this sequence, with gene predictions identifying multiple sections of transposable element polyproteins.
2.5.3.1 CBB Resistance Markers and CG G19833 Genome Location

Two CBB resistance QTLs, SU91 and UBC420, have previously been reported on chromosomes 6 and 8 (Pedraza et al. 1997; Yu et al. 2000b; Vandemark et al. 2008; Shi et al. 2011). The QTLs associated with these markers have also been shown to have epistatic effects which lead to increased disease resistance when both markers are present (Shi et al. 2011; Durham et al. 2013). SU91 was located initially to chromosome 8 by linkage mapping (Pedraza et al. 1997). Subsequently, genome sequence analysis by Perry et al. (2013) physically located the marker sequence to a contig within the current OAC Rex sequence. Although the marker itself was absent from the G19833 sequence assembly, the sequence surrounding SU91 in the OAC Rex contig matched to the segment between 58,994,870-59,444,870 bp in chromosome 8 of G19833. UBC420 was located to chromosome 6 by linkage mapping in various populations (Yu et al. 2000b; Vandemark et al. 2008; Durham et al. 2013), as well as by association mapping performed by Shi et al. (2011). BLAST of the UBC420 marker sequence (EF553635.1) against G19833 results in a match from 8,124,578-8125483 bp, and alignment between the two indicates 77% similarity.

The genes predicted in section contig1701-1 matched to ~5.8 Mbp which is approximately 2.3 Mbp away from the UBC420 location on chromosome 6. In the case of contig1701-2 genes, they matched to ~49.7 Mbp which is approximately 10 Mbp away from SU91 on chromosome 8 (Perry et al. 2013). Schmutz et al. (2014) reported the recombination rate of euchromatic arms to be 220 kb/cM, which indicates that the regions to which the contig1701-1 and contig1701-2 genes have similarity to are 10.45 cM and 45.45 cM from the location of UBC420 on chromosome 6 and SU91 on
chromosome 8. Thus, the likelihood that CGs CBB1, CBB2, CBB3 and CBB4 are the resistance genes associated with these CBB resistance markers is very small. However, it is interesting that similar sequences are found in chromosomes 4, 6 and 8; all of which have been associated with CBB resistance in various studies.
2.6 Conclusion

Candidate common bacterial blight (CBB) resistance genes CBB1-4 were identified in the sequence of OAC Rex BiBAC2 library clones associated with CBB resistance QTL marker PV-ctt001. Characterization of the candidate genes resulted in the prediction of amino acid sequences which had leucine rich repeat (LRR) characteristics including a consensus LRR motif sequence (LxxxLxLxLxxL). BLAST comparisons of CG sequences to CBB susceptible common bean variety G19833 genome sequence identified similarities to retrotransposons in chromosome 6 and 8. Protein sequence BLAST and secondary structure modelling indicated low similarity between CGs and retrotransposon proteins. Characterization of gene predictions in the sequence surrounding CGs CBB1-4 identified gene 5 as an ascorbate oxidase, which has potential pathogen resistance implications.
Chapter 3: Detached leaf Xap susceptibility assay in *Arabidopsis thaliana*

### 3.1 Abstract

Efforts are currently underway to identify CGs for CBB resistance in common bean and test them. As an alternative to transformation of common bean to observe CG efficacy, a testing method using a model system would be beneficial. Here we describe methods for inoculating detached Arabidopsis rosette leaves with different Xap isolates, maintenance of the leaves in a tissue culture system over 192 hours, and assaying disease progression in the inoculated tissue. Symptom progression and assays of colony forming units (CFUs)/unit of leaf tissue from inoculated leaf tissue samples indicated that one isolate (*Xap18*) is more aggressive on Arabidopsis leaves than other isolates. However, significant variation in establishing a disease state with this system suggests that optimization of the protocol is needed to obtain consistent and reproducible results.
3.2 Introduction

3.2.1 Common Bean Improvement

Breeding for the improvement crop species is a widely used tactic, especially in the case of disease resistance. A major focus of common bean improvement has been resistance to the bacterial disease common bacterial blight (CBB). CBB affects worldwide production of *P. vulgaris* and is caused by the bacterial pathogen *Xap* and fuscans variant *Xff* (Vauterin et al. 1995). Since little natural CBB resistance is exhibited in common bean, introgression of resistance has occurred primarily by interspecific crosses with resistant lines of *P. acutifolius* (Parker, 1985) and *P. coccineus* (Park and Dhanvantari 1987; Miklas et al. 1994). Introgression of loci associated with resistance into common bean lines has been achieved through backcrossing, disease screening and marker assisted selection (MAS) and has resulted in the development of CBB-resistant common bean varieties, including the navy bean OAC Rex (Michaels et al. 2006). However, the use of markers for selecting for resistance is limited by the possibility of the gene for resistance segregating from the marker, creating the potential for false positives. The identification of the gene(s) that confers resistance to CBB in common bean and subsequent development of a gene-based marker would reduce the testing time and the frequency of false positives involved in developing new disease resistant varieties.

3.2.2 CBB in *A. thaliana*

To test the involvement of a particular candidate gene in CBB resistance, a testing method for the gene(s) *in planta* is desired. However, common bean has proven to be a difficult species to transform (Gepts et al. 2008). This makes the use of gene
transformation to test CGs in susceptible varieties a difficult, if not impossible strategy, and necessitates the use of other means of testing CGs for efficacy. Although common bean is the primary host of CBB, a study by Perry (2010) indicated that the model plant *A. thaliana* can be manually inoculated with *Xap* to produce CBB-like symptoms. Chlorotic lesions appeared on Arabidopsis leaves inoculated with a mixture of *Xap* isolates by 96 hours post inoculation (HPI). The lesions expanded and caused leaf curling and eventually considerable tissue damage by 192 HPI. Symptoms also began to appear around the margins of non-inoculated leaves in proximity to inoculated leaves, which developed into symptoms similar to those of inoculated leaves. The demonstration of Arabidopsis as a host makes it possible to test candidate CBB resistance genes in the model system. By first transforming Arabidopsis with genes of interest, the transformed plants can then be inoculated with *Xap*, and tested for symptoms.

Although inoculation of leaves while still on the plant was shown to produce CBB-like symptoms on Arabidopsis, for the purpose of testing transformed Arabidopsis plants, a more precise, sterile and contained environment is preferred. The goal of this study was to develop a tissue culture system to maintain detached rosette leaves post-inoculation for the purpose of observing disease progression. This approach has the potential to reduce the damage to the leaves and plant, as well as the likelihood of introducing contaminating diseases during inoculation. This will help to improve the ability to observe CBB symptom progression in Arabidopsis plants transformed with candidate CBB resistance genes.
3.3 Materials and Methods

3.3.1 Bacterial and Plant Growth

*Arabidopsis thaliana* ecotype Columbia seeds were suspended in distilled water at approximately 50 seeds/mL and dispersed into 24-cell flats containing LA4 (Green Island Distributors, Riverhead, N.Y.) that had been pre-soaked with deionized water and drained. Flats were covered with clear plastic domes and maintained in a growth chamber providing 16 h light (150 µmol/m²/s), 8 h dark on a daily cycle at 22 °C.

Seeds from *Phaseolus vulgaris* varieties Nautica (susceptible) and OAC Rex (resistant) were inserted in 6" round pots filled with LA4 and covered with a 1" layer of medium vermiculite (Green Island Distributors, Riverhead, N.Y.). Pots were soaked with deionized water and maintained in a growth chamber providing 16 h light and 8 h dark cycle at 22°C.

Bacterial inoculum was generated by spreading 50 mL of *Xap* isolates 12, 18, 98 and 118, separately, on yeast salt 1% agar media plates and maintaining them at 28 °C for 48 h. *Xap* cultures were sub-cultured as streak plates for each isolate, and were incubated for an additional 48 h at 28°C. The resultant colonies were then washed off the plates into a 50 mL Falcon tube with 20 mL of distilled sterile water. The cell density of the plate rinse was determined and diluted with distilled sterile water to an OD₆₀₀ of 0.6 (1x10⁷ colony forming units (CFU)/g) as needed.

3.3.2 Plant Wounding and Inoculation

Healthy rosette leaves were detached from 3-4 week old *Arabidopsis* plants using a scalpel so that approximately 2 cm of petiole remained attached to the leaf. Healthy *P. vulgaris* leaflets were detached from 3-4 week old plants. Inoculation was
performed by wounding each leaf or leaflet with a 10 µl plastic pipette tip once, close to the mid-vein, and ejecting 10 µl of inoculum onto the wound site. Wound sites were inoculated with 10 µl of either sterile distilled water or one of the four *Xap* isolates (*Xap12, Xap18, Xap98, Xap118*) collected in Ontario (Park and Dhanvantari 1987). The liquid inoculum was left on the wound site of each inoculated leaf to ensure infection. Leaves representative of the 0 HPI were carefully transferred to a light blue background, pictures were taken for IA, and then a sample of 1 cm² of the tissue surrounding the inoculation site was cut from the leaf.

Additional inoculated Arabidopsis, Nautica and OAC Rex leaves were placed in tissue culture for further observations at 0, 48, 96 and 144 HPI. The Arabidopsis leaves were placed in petri plates containing 0.7% water agar media. Both bean varieties were placed in Phytatray™ II boxes (Sigma-Aldrich Col. LLC) containing 0.7% water agar. For both *Arabidopsis* and *P. vulgaris*, the leaves were kept upright at a 45° angle by inserting the petiole into the solid media (Chen *et al.* 2006) to prevent the leaf blade from contacting the agar. Phytatrays and petri plates were wrapped with Parafilm to maintain a high relative humidity and placed into a 28 °C tissue culture growth cabinet set to 16 h light (150 µmol/m²/s) and 8 h dark in a randomized complete block design where each experiment was a block. At 48 hour intervals post inoculation, pictures and ~1cm leaf samples were taken as described above, for assessment of disease symptom progression over time.
3.3.3 Assessment of Disease Symptoms

3.3.3.1 Image Collection and Analysis

Disease symptom assessment was accomplished by performing Image Analysis (IA) on inoculated leaves using the APS Assess 2.0 computer program (The American Phytopathological Society, St. Paul, MN). Image collection of the inoculated common bean leaves and subsequent IA was performed as described by Xie et al. (2012) to determine the percent of the leaf area showing symptoms. The same procedure was followed for inoculated Arabidopsis leaves. However, the distance of the camera from the Arabidopsis leaves was adjusted to 20 cm, to reflect the smaller size of the Arabidopsis leaves. Manual IA was performed to determine the area of chlorosis and/or necrosis for inoculated leaves at 0, 48, 96, 144 and 192 HPI. Although there were no symptoms at 0 HPI, pictures were taken to provide a visual comparison for later time points when the symptoms had developed.

3.3.3.2 CFU Analysis

Samples of approximately 1 cm² of the tissue surrounding the inoculation site were taken from each leaf and homogenized in 0.9 ml of sterile distilled water (Perry, 2010). Serial dilutions of $10^3$-fold and $10^5$-fold were made for samples taken at 0 HPI for all inoculum types. The dilutions were plated on XCP media ($10$ g/L KBr, 0.25 g/L CaCl₂, 10 g/L potato starch, 10 g/L peptone, 10 ml/L Tween 80, 3 mg/L fluorouracil and 0.16 mg/L tobramycin pH 7.0; Remeeus and Sheppard 2006) for a final dilution of $10^5$ and $10^7$, and plates were incubated at 28°C for 48 hours. For samples taken at 48, 96, 144 and 192 HPI, $10^5$ and $10^7$ fold serial dilutions were made for all inoculum types. The dilutions were plated on XCP media plates for a final dilution of $10^7$ and $10^9$ and
incubated at 28°C for 48 hours. Visible colonies >2 mm in diameter were marked at 48
hours, and plates were incubated for an additional 96 hours. Colonies with visible
media clearing were then counted to provide the CFU value for the corresponding
inoculum type and time post inoculation.

3.3.4 Statistical Analyses

Variance analysis, lsmeans estimates and comparisons were performed on IA
lesion percent values and CFUs from a single replication of the experiment using
PROC MIXED in SAS 9.4 (SAS institute Inc., Cary, NC). Isolate, genotype and time
were all considered fixed effects. Tests for normality prompted log transformation of the
data, however further analysis was performed using untransformed data since
transformation did not improve the normality of the residuals. The data was sliced to
separate the effects of the different isolates, plant genotypes and time points post
inoculation on the interaction of all factors. ANOVA was performed to determine the
significance of the sliced effects. Letter grouping to show the significant differences
between lsmeans was performed using the PDMIX800 macro (Saxton 2000) with data
sliced for time to make comparisons between all lsmeans at the different time points.
For all analyses, an interaction was considered significant when the p-value was less
than 0.05.
3.4 Results

3.4.1 Analysis of Variance

Analysis of variance for OAC Rex and Nautica CFU and IA results showed that the inoculum, genotype effects were significant (Appendix 5). Interactions between inoculum and genotype, inoculum and time and genotype and time were also significant (Appendix 5). Interaction between inoculum, genotype and time was not significant (Appendix 5). Analysis of variance for Arabidopsis CFU and IA results showed that inoculum and time effects were significant, and that the interaction between inoculum and time was significant (Appendix 6).

3.4.2 Visual Observations and IA

3.4.2.1 Common Bean Visible Leaf Symptoms

Nautica and OAC Rex leaves inoculated with water showed no visible disease symptoms at all sampling time-points, and lsmeans of percent of leaf with disease symptoms (hereafter known as symptoms) showed no significant differences between time points or variety (Appendix 7-9; Figure 3.1A; Tables 3.1-3.2 and 3.8-3.9). *Xap* inoculated leaves from both common bean varieties were not significantly different for inoculation with any of the four isolates or water at 0, 48 and 96 HPI (Appendix 7-9; Figure 3.1B-E). At 144 HPI, significant differences were seen between the symptoms seen in Nautica and OAC Rex leaves when inoculated with *Xap* isolates (Figure 3.1; Table 3.1). OAC Rex leaves showed no significant difference between water and *Xap* inoculation (Figure 3.1; Table 3.1). *Xap* and water inoculated Nautica leaf symptoms were significantly different when isolates *Xap98* and *Xap118* were used, with the most severe symptom being observed in the *Xap98* inoculated sample (36.21%) (Figure 3.1)
Figure 3.1. Symptom progression in inoculated *P. vulgaris* leaves. Percent leaf area showing symptoms at 0, 48, 96, 144 and 192 hours post inoculation (HPI). Inoculations were performed on leaves of susceptible (Nautica – dark colour) and resistant (OAC Rex - light) varieties with water (A - Black) or one of four *X. axonopodis* pv. *phaseoli* (*Xap*) isolates; *Xap12* (B - Red), *Xap18* (C - Purple), *Xap98* (D - Blue) and *Xap118* (E - Green). Error bars represent the standard error of the values of the lsmeans (± 6.18).
A, D, E; Table 3.1). At the last time point (192 HPI) there were significant differences between Nautica and OAC Rex for inoculation with each Xap isolate (Figure 3.1; Table 3.2). Inoculation of OAC Rex leaves, resulted in symptoms that were significantly different from water inoculation only when isolate Xap118 was used (31.65%) (Figure 3.1 A, E; Table 3.2). All isolates produced symptoms greater than water inoculation in Nautica leaves, with the highest percentage of affected leaf being observed in the samples inoculated with Xap118 (66.71%) (Figure 3.1 A-E; Table 3.2).

### 3.4.2.2 Arabidopsis Visible Leaf Symptoms

Arabidopsis leaves inoculated with water showed no visible disease symptoms at all sampling time-points, and lsmeans of symptoms showed no significant differences between time points (Appendix 7-9; Figure 3.2A; Tables 3.1-3.2 and 3.7). Xap inoculated Arabidopsis leaves were not significantly different than the water controls or each other with any of the four isolates at 0, 48 and 96 HPI (Appendix 7-9; Figure 3.2; Table 3.7). At 144 HPI, all Xap isolates produced symptoms that were significantly higher than the water control in Arabidopsis leaves, with the most severe symptoms being produced by Xap98 (24.50%) (Figure 3.2; Table 3.1). However, at 192 HPI, the symptoms in Arabidopsis leaves that were induced by Xap98 (2.10%) were not significantly different than in the water control (Figure 3.2; Table 3.2). Symptoms induced by the other isolates at 192 HPI were significantly higher than the water control for the other isolates with the highest mean symptoms being produced by Xap18 (19.30%) (Figure 3.2; Table 3.2).
Figure 3.2. Symptom progression in inoculated *A. thaliana* rosette leaves. Percent leaf area showing symptoms for Arabidopsis rosette leaves at 0, 48, 96, 144 and 192 hours post inoculation (HPI). Inoculations were performed using water (black) as a control, and one of four *X. axonopodis pv. phaseoli* (Xap) isolates Xap12 (red) Xap18 (purple) Xap98 (blue) and Xap118 (green). Error bars represent the standard error of the values of the lsmeans (± 3.06).
Table 3.1. LSmeans for common bacterial blight symptoms on OAC Rex, Nautica and Arabidopsis leaves at 144 hours post inoculation (HPI). Percent of leaf showing disease symptoms for OAC Rex, Nautica and Arabidopsis inoculated with water, Xap12, Xap18, Xap98 and Xap118 at 144 HPI

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Symptoms (%)</th>
<th>Inoculum</th>
<th>Water</th>
<th>Xap12</th>
<th>Xap18</th>
<th>Xap98</th>
<th>Xap118</th>
<th>Genotype Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>OAC Rex</td>
<td>1.87 ^c</td>
<td>2.72 ^c</td>
<td>3.14 ^c</td>
<td>4.72 ^c</td>
<td>7.90 ^c</td>
<td>4.07 ^w</td>
<td></td>
<td>4.07 ^W B</td>
</tr>
<tr>
<td>Nautica</td>
<td>1.88 ^c</td>
<td>3.78 ^c</td>
<td>6.38 ^c</td>
<td>36.2 ^a</td>
<td>14.8 ^bc</td>
<td>12.6 ^A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arabidopsis</td>
<td>1.68 ^y x</td>
<td>16.6 ^yz</td>
<td>13.9 ^y</td>
<td>24.5 ^z</td>
<td>9.79 ^xz</td>
<td>13.3 ^u</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

^c LSmeans followed by the same lower case letter were not significantly different at p=0.05
^y LSmeans followed by the same upper case letter were not significantly different at p=0.05
^X Denotes lsmeans with a standard error of ±127.80
^W Denotes lsmeans with a standard error of ±57.1532
^V Denotes lsmeans with a standard error of ±3.06
^U Denotes lsmeans with a standard error of ±1.37

Table 3.2. LSmeans for common bacterial blight symptoms on OAC Rex, Nautica and Arabidopsis leaves at 192 hours post inoculation (HPI). Percent of leaf showing disease symptoms for OAC Rex, Nautica and Arabidopsis inoculated with water, Xap12, Xap18, Xap98 and Xap118 at 192 HPI

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Symptoms (%)</th>
<th>Inoculum</th>
<th>Water</th>
<th>Xap12</th>
<th>Xap18</th>
<th>Xap98</th>
<th>Xap118</th>
<th>Genotype Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>OAC Rex</td>
<td>1.62 ^f</td>
<td>10.7 ^def</td>
<td>2.69 ^f</td>
<td>5.27 ^ef</td>
<td>31.7 ^bc</td>
<td>10.4 ^B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nautica</td>
<td>7.07 ^def</td>
<td>20.4 ^cd</td>
<td>57.1 ^a</td>
<td>38.7 ^b</td>
<td>66.7 ^a</td>
<td>38.0 ^A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arabidopsis</td>
<td>1.04 ^x</td>
<td>9.3 ^xy</td>
<td>19.3 ^y</td>
<td>2.10 ^x</td>
<td>12.6 ^yz</td>
<td>8.87 ^v</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

^f LSmeans followed by the same lower case letter were not significantly different at p=0.05
^y LSmeans followed by the same upper case letter were not significantly different at p=0.05
^X Denotes lsmeans with a standard error of ±127.80
^W Denotes lsmeans with a standard error of ±57.1532
^V Denotes lsmeans with a standard error of ±3.06
^U Denotes lsmeans with a standard error of ±1.37
3.4.3 Leaf Sample CFUs

3.4.3.1 Common Bean CFUs

CFUs produced from leaf samples taken from water-inoculated Nautica and OAC Rex leaves were not significantly different at all time-points (Appendix 10; Figure 3.3 A; Table 3.3-3.6). There were also no significant differences between CFU produced from either bean genotype inoculated with water or the four Xap isolates at 0 HPI (Appendix 10; Figure 3.3 B-E; Table 3.3). At 48, 96, 144 and 192 HPI, there were significant differences between the CFUs for OAC Rex and Nautica leaf samples inoculated Xap isolates (Figure 3.3; Table 3.3-3.6). At 48 HPI, OAC Rex CFU means were significantly different from water when Xap18 and Xap98 were used, with the highest mean being produced by Xap18 (7.64 x 10^7 CFU) (Figure 3.3 C-D; Table 3.3). All isolates produced CFU means that were significantly different from water in Nautica leaves, with the highest being produced by inoculation with Xap18 (1.07 x 10^8 CFU) (Figure 3.3; Table 3.3). At 96 HPI, OAC Rex CFU means were significantly different from water inoculation when Xap18 and Xap98 were used, with the highest mean being produced by Xap18 (7.03 x 10^7 CFU) (Figure 3.3 C-D; Table 3.4). Inoculation of Nautica leaves by all Xap isolates produced CFUs that were significantly different from water inoculation, with the highest mean being produced by Xap118 (1.10 x 10^8 CFU) (Figure 3.3; Table 3.4). At 144 HPI, OAC Rex leaf samples produced significantly more CFU than the water control only when inoculated with Xap18 (4.57 x 10^7 CFU) (Figure 3.3 C; Table 3.5). All isolates produced more CFUs than the water inoculation in the Nautica leaf samples, with the highest mean being produced by Xap18 (1.18 x 10^8 CFU) (Figure 3.3; Table 3.5). At 192 HPI all isolates produced CFUs that were significantly higher.
Figure 3.3. Pathogen growth in inoculated *P. vulgaris* leaf samples. Colony forming units (CFU) for leaf samples taken from susceptible (Nautica - dark) and resistant (OAC Rex - light) varieties, respectively at 0, 48, 96, 144 and 192 hours post inoculation (HPI). Inoculation was performed with water (A - Black) or one of four *X. axonopodis pv. phaseoli* (Xap) isolates; Xap12 (B - Red) Xap18 (C - Purple) Xap98 (D - Blue) and Xap118 (E - Green). Error bars represent the standard error of the values of the lsmeans (± 142).
than observed for the water inoculated leaf samples for both OAC Rex and Nautica (Figure 3.3; Table 3.6). The highest mean CFUs produced were produced by inoculation with *Xap18*, which had $4.57 \times 10^7$ CFU and $1.18 \times 10^8$ CFU for OAC Rex and Nautica, respectively (Figure 3.3 C; Table 3.6).

### 3.4.3.2 Arabidopsis CFUs

Arabidopsis leaves inoculated with water showed no visible disease symptoms at all sampling time-points, and lsmeans of symptoms showed no significant differences between time points (Appendix 10; Figure 3.4; Table 3.3-3.6). At 0 HPI, none of the *Xap* isolates produced CFU means that were significantly different from water inoculation (Appendix 10; Figure 3.4 B-E). At 48 HPI, only *Xap98* produced a CFU mean that was significantly higher than the water sample ($4.76 \times 10^7$ CFU) (Figure 3.4 D; Table 3.3). At 96 HPI, all *Xap* isolates leaves produced CFU means that were significantly higher than water inoculated leaves, with the highest mean being produced by *Xap18* ($5.76 \times 10^7$ CFU) (Figure 3.4; Table 3.4). At 144 HPI, all *Xap* isolates produced CFU means that were significantly higher than water inoculated leaves, with the highest mean being produced by *Xap18* ($6.76 \times 10^7$ CFU) (Figure 3.4; Table 3.5). At 192 HPI, all *Xap* isolates leaves produced CFU means that were significantly higher from water inoculated leaves, with the highest mean being produced by *Xap18* ($1.04 \times 10^8$ CFU) (Figure 3.4; Table 3.6).
Figure 3.4. Pathogen growth in inoculated *A. thaliana* leaf samples. Colony forming units (CFU) for Arabidopsis leaves 0, 48, 96, 144 and 192 hours post inoculation (HPI). Inoculation was performed with water (black), Xap12 (red), Xap18 (purple), Xap98 (blue), or Xap118 (green). Error bars represent the standard error of the values of the lsmeans (+ 96.3).
Table 3.3. LSmeans for colony forming units for *X. axonopodis pv. phaseoli* (*Xap*) bacteria at 48 hours post inoculation (HPI). CFUs were determined by incubating $10^5$ dilutions of ground leaf samples taken from inoculated leaves from resistant (OAC Rex) and susceptible (Nautica) *P. vulgaris* varieties, and Arabidopsis. Samples were taken at 48 HPI after inoculation with Water, *Xap12, Xap18, Xap98* or *Xap118*.

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Inoculum</th>
<th>CFU ($10^5$)</th>
<th>Genotype Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Water</td>
<td>Xap12</td>
<td>Xap18</td>
</tr>
<tr>
<td>OAC Rex</td>
<td>108 $^{de}$</td>
<td>348 $^{cde}$</td>
<td>764 $^{ab}$</td>
</tr>
<tr>
<td>Nautica</td>
<td>107 $^{de}$</td>
<td>676 $^{abc}$</td>
<td>1070 $^a$</td>
</tr>
<tr>
<td>Arabidopsis</td>
<td>49.7 $^{y}$</td>
<td>261 $^{yz}$</td>
<td>307 $^{yz}$</td>
</tr>
</tbody>
</table>

*LSmeans followed by the same lower case letter were not significantly different at p=0.05

Table 3.4. LSmeans for colony forming units for *X. axonopodis pv. phaseoli* (*Xap*) bacteria at 96 hours post inoculation (HPI). CFUs were determined by incubating $10^5$ dilutions of ground leaf samples taken from inoculated leaves from resistant (OAC Rex) and susceptible (Nautica) *P. vulgaris* varieties, and Arabidopsis. Samples were taken at 96 HPI after inoculation with Water, *Xap12, Xap18, Xap98* or *Xap118*.

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Inoculum</th>
<th>CFU ($10^5$)</th>
<th>Genotype Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Water</td>
<td>Xap12</td>
<td>Xap18</td>
</tr>
<tr>
<td>OAC Rex</td>
<td>48.7 $^{de}$</td>
<td>101 $^{de}$</td>
<td>703 $^{abc}$</td>
</tr>
<tr>
<td>Nautica</td>
<td>3.00 $^{e}$</td>
<td>667 $^{bc}$</td>
<td>1070 $^a$</td>
</tr>
<tr>
<td>Arabidopsis</td>
<td>26.7 $^{x}$</td>
<td>206 $^{yz}$</td>
<td>567 $^z$</td>
</tr>
</tbody>
</table>

*LSmeans followed by the same lower case letter were not significantly different at p=0.05

- $^{de}$ Denotes lsmeans with a standard error of $\pm 142$
- $^{bcd}$ Denotes lsmeans with a standard error of $\pm 63.6$
- $^{yz}$ Denotes lsmeans with a standard error of $\pm 96.3$
- $^{x}$ Denotes lsmeans with a standard error of $\pm 43.1$
Table 3.5. LSmeans for colony forming units for *X. axonopodis pv. phaseoli* (*Xap*) bacteria at 144 hours post inoculation (HPI). CFUs were determined by incubating $10^5$ dilutions of ground leaf samples taken from inoculated leaves from resistant (OAC Rex) and susceptible (Nautica) *P. vulgaris* varieties, and Arabidopsis. Samples were taken at 144 HPI after inoculation with Water, *Xap*12, *Xap*18, *Xap*98 or *Xap*118.

<table>
<thead>
<tr>
<th>Inoculum</th>
<th>OAC Rex</th>
<th>Genotype Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td>57.7 $^{bc}$ c</td>
<td>222 $^{wB}$</td>
</tr>
<tr>
<td><em>Xap</em>12</td>
<td>183 $^{bc}$ x</td>
<td></td>
</tr>
<tr>
<td><em>Xap</em>18</td>
<td>457 $^{b}$ a</td>
<td></td>
</tr>
<tr>
<td><em>Xap</em>98</td>
<td>207 $^{bc}$ b</td>
<td></td>
</tr>
<tr>
<td><em>Xap</em>118</td>
<td>204 $^{bc}$ b</td>
<td></td>
</tr>
<tr>
<td>Nautica</td>
<td>0.333 $^{bc}$ x</td>
<td>756 $^{wA}$</td>
</tr>
<tr>
<td>Arabidopsis</td>
<td>363 $^{u}$</td>
<td></td>
</tr>
</tbody>
</table>

$^x$ Lsmeans followed by the same lower case letter were not significantly different at $p=0.05$

$^y$ Lsmeans followed by the same upper case letter were not significantly different at $p=0.05$

$x$ Denotes lsmeans with a standard error of $\pm 142$

$y$ Denotes lsmeans with a standard error of $\pm 63.6$

$z$ Denotes lsmeans with a standard error of $\pm 96.3$

$u$ Denotes lsmeans with a standard error of $\pm 43.1$

---

Table 3.6. LSmeans for colony forming units for *X. axonopodis pv. phaseoli* (*Xap*) bacteria at 192 hours post inoculation (HPI). CFUs were determined by incubating $10^5$ dilutions of ground leaf samples taken from inoculated leaves from resistant (OAC Rex) and susceptible (Nautica) *P. vulgaris* varieties, and Arabidopsis. Samples were taken at 192 HPI after inoculation with Water, *Xap*12, *Xap*18, *Xap*98 or *Xap*118.

<table>
<thead>
<tr>
<th>Inoculum</th>
<th>OAC Rex</th>
<th>Genotype Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td>34.7 $^{bc}$ c</td>
<td>617 $^{wA}$</td>
</tr>
<tr>
<td><em>Xap</em>12</td>
<td>789 $^{ab}$ a</td>
<td></td>
</tr>
<tr>
<td><em>Xap</em>18</td>
<td>863 $^{ab}$ a</td>
<td></td>
</tr>
<tr>
<td><em>Xap</em>98</td>
<td>804 $^{ab}$ a</td>
<td></td>
</tr>
<tr>
<td><em>Xap</em>118</td>
<td>592 $^{b}$ b</td>
<td></td>
</tr>
<tr>
<td>Nautica</td>
<td>0 $^{x}$ a</td>
<td>794 $^{A}$</td>
</tr>
<tr>
<td>Arabidopsis</td>
<td>467 $^{u}$</td>
<td></td>
</tr>
<tr>
<td>Arabidopsis</td>
<td>319 $^{y}$ x</td>
<td></td>
</tr>
<tr>
<td>Arabidopsis</td>
<td>1040 $^{y}$ x</td>
<td></td>
</tr>
</tbody>
</table>

$^x$ Lsmeans followed by the same lower case letter were not significantly different at $p=0.05$

$^y$ Lsmeans followed by the same upper case letter were not significantly different at $p=0.05$

$x$ Denotes lsmeans with a standard error of $\pm 142$

$y$ Denotes lsmeans with a standard error of $\pm 63.6$

$z$ Denotes lsmeans with a standard error of $\pm 96.3$
Table 3.7. Inoculated *A. thaliana* leaf symptom progression. Arabidopsis rosette leaves were inoculated with sterile distilled water control and *X. axonopodis pv. phaseoli* (*Xap*) isolates 12, 18, 98 and 118. Pictures were taken at 0, 48, 96, 144 and 192 hours post inoculation (HPI).

<table>
<thead>
<tr>
<th>Inoculum</th>
<th>0</th>
<th>48</th>
<th>96</th>
<th>144</th>
<th>192</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td><img src="image1" alt="Image" /></td>
<td><img src="image2" alt="Image" /></td>
<td><img src="image3" alt="Image" /></td>
<td><img src="image4" alt="Image" /></td>
<td><img src="image5" alt="Image" /></td>
</tr>
<tr>
<td>Xap12</td>
<td><img src="image6" alt="Image" /></td>
<td><img src="image7" alt="Image" /></td>
<td><img src="image8" alt="Image" /></td>
<td><img src="image9" alt="Image" /></td>
<td><img src="image10" alt="Image" /></td>
</tr>
<tr>
<td>Xap18</td>
<td><img src="image11" alt="Image" /></td>
<td><img src="image12" alt="Image" /></td>
<td><img src="image13" alt="Image" /></td>
<td><img src="image14" alt="Image" /></td>
<td><img src="image15" alt="Image" /></td>
</tr>
<tr>
<td>Xap98</td>
<td><img src="image16" alt="Image" /></td>
<td><img src="image17" alt="Image" /></td>
<td><img src="image18" alt="Image" /></td>
<td><img src="image19" alt="Image" /></td>
<td><img src="image20" alt="Image" /></td>
</tr>
<tr>
<td>Xap118</td>
<td><img src="image21" alt="Image" /></td>
<td><img src="image22" alt="Image" /></td>
<td><img src="image23" alt="Image" /></td>
<td><img src="image24" alt="Image" /></td>
<td><img src="image25" alt="Image" /></td>
</tr>
</tbody>
</table>
**Table 3.8.** Inoculated common bacterial blight susceptible *P. vulgaris* variety leaf symptom progression. Trifoliate Nautica leaves were inoculated with sterile distilled water control and *X. axonopodis* pv. *phaseoli* (*Xap*) isolates 12, 18, 98 and 118. Pictures were taken at 0, 48, 96, 144 and 192 hours post inoculation (HPI).

<table>
<thead>
<tr>
<th>Inoculum</th>
<th>0</th>
<th>48</th>
<th>96</th>
<th>144</th>
<th>192</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td><img src="image" alt="Leaf" /></td>
<td><img src="image" alt="Leaf" /></td>
<td><img src="image" alt="Leaf" /></td>
<td><img src="image" alt="Leaf" /></td>
<td><img src="image" alt="Leaf" /></td>
</tr>
<tr>
<td>Xap12</td>
<td><img src="image" alt="Leaf" /></td>
<td><img src="image" alt="Leaf" /></td>
<td><img src="image" alt="Leaf" /></td>
<td><img src="image" alt="Leaf" /></td>
<td><img src="image" alt="Leaf" /></td>
</tr>
<tr>
<td>Xap18</td>
<td><img src="image" alt="Leaf" /></td>
<td><img src="image" alt="Leaf" /></td>
<td><img src="image" alt="Leaf" /></td>
<td><img src="image" alt="Leaf" /></td>
<td><img src="image" alt="Leaf" /></td>
</tr>
<tr>
<td>Xap98</td>
<td><img src="image" alt="Leaf" /></td>
<td><img src="image" alt="Leaf" /></td>
<td><img src="image" alt="Leaf" /></td>
<td><img src="image" alt="Leaf" /></td>
<td><img src="image" alt="Leaf" /></td>
</tr>
<tr>
<td>Xap118</td>
<td><img src="image" alt="Leaf" /></td>
<td><img src="image" alt="Leaf" /></td>
<td><img src="image" alt="Leaf" /></td>
<td><img src="image" alt="Leaf" /></td>
<td><img src="image" alt="Leaf" /></td>
</tr>
</tbody>
</table>
Table 3.9. Inoculated common bacterial blight resistant *P. vulgaris* variety leaf symptom progression. Trifoliate OAC Rex leaves were inoculated with sterile distilled water control and *X. axonopodis* pv. *phaseoli* (*Xap*) isolates 12, 18, 98 and 118. Pictures were taken at 0, 48, 96, 144 and 192 hours post inoculation (HPI)
3.5 Discussion

3.5.1 Symptom Development

3.5.1.1 Common Bean Symptom Evaluation by IA

The highest disease symptoms were generally produced in Nautica, with significant variability caused by the isolate used. Nautica was included in this study as a *Xap* susceptible common bean check variety. Therefore, the increased symptoms observed in Nautica leaves compared to OAC Rex leaves inoculated with the same *Xap* isolate was consistent with the expectation that this variety would show more severe symptoms in a shorter time after inoculation. Although a direct comparison is not possible due to the differences between experimental parameters and the resistant varieties used, a comparison of the symptom severity scores observed in the current study with detached leaves to those reported for a similar growth chamber study (Xie *et al.* 2012) with intact plants, suggests that the disease reaction was more severe in the current detached leaf assay. In the previous study, the percent lesion area was 6.03% for Nautica and 0.22-0.29% for the resistant varieties (MBE7, Rexeter and Apex) 14-18 days post inoculation (336-432 HPI) (Xie *et al.*2012). In contrast, Nautica and OAC Rex symptoms covered 38.0% and 10.4% of the leaves, respectively, at 192 HPI in the current study (Table 3.2).

There are several potential explanations for the observed differences between the results reported by Xie *et al.* (2012) and those obtained in the current study. The most immediate explanation is that the environment provided by the tissue culture is more suited for symptom development. However, it is also possible that there are differences in the reaction to pathogens in detached versus attached leaves. This was
observed previously in Arabidopsis by Liu et al. (2007b), when attached and detached leaves were inoculated with either Colletotrichum liniola A1 or C. higginsianum, resulting in opposite responses. Inoculation with C. liniola A1 did not affect attached leaves, but it did infect detached leaves. In contrast, inoculation with C. higginsianum did affect attached leaves, but not when they were detached.

Other possible causes for the increased symptom severity in the detached leaf assay are differences in the sizes and developmental stages of the leaves. In both the common bean study performed by Xie et al. (2012) and the Arabidopsis study performed by Perry (2010), inoculated leaves remained on the plant for the duration of the experiment. This allowed the leaves to continue to grow, whereas, the detached leaves in the current study remained approximately the same size, which would affect the relative symptom size. Since the media used in this study lacked any additional nutrients, the detached leaves may have had an impaired ability to produce a resistance response due to lack of resources.

In the current study, trifoliate leaves with different sizes and developmental stages were used, whereas, Xie et al. (2012) used only unifoliate leaves. Not only were the trifoliate leaves smaller than the unifoliate leaves, some of the leaves used for inoculation were quite new, which may have made them more susceptible to Xap infection than the unifoliate leaves. This is consistent with the study performed by Jung et al. (1997), which found that there was a significant strain by organ interaction, and that certain QTL only contributed to resistance in specific organs and developmental stages. The Random Amplified Polymorphic DNA (RAPD) marker D13.1000, in
particular, was only associated with resistance to Xap isolate DR-7 in the first trifoliate leaves, and not in later trifoliate leaves, pods or seeds.

3.5.1.2 Arabidopsis Symptom Evaluation by IA

Xap induced symptoms on Arabidopsis leaves were variable compared to those induced on common bean leaves of both varieties. The mean symptom percent for Arabidopsis was higher than the means for both Nautica and OAC Rex at 144 HPI, indicating that it’s response to Xap inoculation was similar to the susceptible Nautica (Table 3.1). This contrasted to the symptom severity observed in Arabidopsis leaves at 192 HPI, which was lower than that observed for Nautica or OAC Rex (Table 3.2). This change from a susceptible symptom mean to a resistant one, over time, is likely caused by the changes observed in two isolates in particular. The symptoms produced by isolates Xap12 and Xap98 were the highest at 144 HPI, and then dropped significantly at 192 HPI (Figure 3.1 B, D; Table 3.1-3.2). However, the symptoms produced by Xap18 and Xap118 increased from 144 to 192 HPI (Figure 3.1 C, E; Table 3.1-3.2).

This change from high to low symptoms in Xap12 and Xap98 inoculated Arabidopsis leaves may be a reflection of the ability of these isolates to infect Arabidopsis. It is also possible that the conditions of those detached leaves were particularly conducive to disease progression such as increased humidity or light availability, or that contamination from another pathogen occurred. However, since only one set of assays was performed, and the results are not entirely consistent with those observed in the study by Perry (2010), more repetitions of the experiment are needed to determine if this is a significant trend, or a random occurrence.
3.5.1.3 Symptom Evaluation of CFUs

When compared to the collected IA data, the CFU means were more variable than the data based on symptom development. However, all genotypes showed significant differences between *Xap*- inoculation means and water inoculation means as early as 48 HPI. Additionally, Nautica generally had the highest CFU values compared to OAC Rex, which is consistent with the trends observed by IA. This supports the assumption that the symptoms are caused by the proliferation of the pathogen in the tissue.

Nevertheless, some inconsistencies like the observation that the highest CFU values were induced by *Xap18* on all three leaf types. Although the highest symptoms observed in Arabidopsis were also produced by inoculation with isolate *Xap18*, the highest values induced in leaves from both common bean varieties were from inoculation with isolate *Xap118*. The difference in common bean and Arabidopsis symptom development in response to inoculation with *Xap18* is further emphasized by the lack of difference between the symptoms produced in OAC Rex leaves inoculated by water and *Xap18*. It is also interesting to note that leaves from all genotypes inoculated with *Xap18* tended to be a lighter green than leaves inoculated with other isolates (Figures 3.7-3.9), which indicates that *Xap18* affects a larger area within the leaf, instead of causing localized lesions. This suggests that *Xap18* symptoms included more than localized lesions, and that IA didn’t entirely capture *Xap18* symptoms because the IA parameters measured lesions (chlorosis and necrosis), and the lighter green colour was not taken into account. These symptom and CFU differences
suggest that differences in the interaction between the genotype and Xap isolate are at play.

The variability in Xanthomonas host specificity, both within and between species and pathovars, has been well documented (Rademaker et al. 2005; Sun et al. 2006; Bogdanove et al. 2011), as well as specifically for the Xap-P. vulgaris interaction (Schuster et al. 1983; Mkandawire et al. 2004; Duncan et al. 2010; Fourie and Herselman 2011). Within the four Xap isolates used in Ontario breeding efforts, isolates Xap98 and Xap118 have been observed to be the most virulent on common bean varieties. It is likely that human error is responsible for the discrepancy between the Xap18 IA and CFU results observed in both common bean genotypes. Just as the IA results may be lower than they should be (discussed above), the CFUs may be higher than actual due to inexperience in counting. Since the IA and CFU means for Xap18 inoculated Arabidopsis leaves progressed in a similar way (Figure 3.2; Figure 3.4), it is most likely that the suggested higher virulence of this strain on Arabidopsis is reliable. However, replication of the study is required to determine the significance of these results.

3.5.2 Suitability of Detached Arabidopsis Leaf CBB Assay

3.5.2.1 Variability in CFU Observations

Considering the variability of the CFU counts over the course of the study, it must be concluded that CFU count cannot be taken as an independent measure of isolate virulence. However, the study has limited scope due to the failure to replicate the results using a single isolate (Xap12) (data not shown). Additional studies with different parameters might reduce the variability in the results by providing more data, and
reducing human error. Regardless of the limitations of the CFU method to compare virulence between isolates within the current study, it does demonstrate that Xap accumulates in inoculated leaf samples over time. Therefore, when combined with IA values, CFU values can be used as an indicator of proliferation within the inoculated sample.

The failure of the experiments with detached Arabidopsis leaves and a single Xap isolate (Xap12) to replicate the symptoms observed with attached Arabidopsis leaves may have been caused by the isolate selection. This isolate initially produced symptoms in Arabidopsis rosette leaves that were visibly similar to those observed by Perry (2010) at 144 HPI (Table 7). However, these symptoms have not been observed in subsequent experiment repetitions, and statistical analysis of the initial results showed that Xap12 produces low IA symptom means and CFU counts in Arabidopsis, as well as OAC Rex and Nautica, in comparison to other isolates (Figure 3.1-3.4).

A more appropriate choice of Xap isolates for further replications of this study would be Xap18, which was shown to produce the highest symptoms in Arabidopsis at 192 HPI (19.3%). This isolate also had a more consistent increase in symptoms and CFU means over time compared to Xap12. Symptoms produced by Xap12 increased quickly and then dropped, whereas Xap18 inoculation produced a more steady increase in symptoms (Figure 3.2). This was even more evident with the CFU means, where Xap18 CFU means increased at a relatively steady rate over time, whereas Xap12 CFU means fluctuated around an apparent plateau from 48 to 192 HPI (Figure 3.4).
3.5.2.2 Detached Leaf vs. Attached Leaf Symptom Progression

The symptom progression in detached Arabidopsis leaves did not follow the pattern of symptom development that was observed in inoculated leaves on the plant, where chlorotic symptoms developed around 96 HPI and increased to the point of considerable damage by 192 HPI (Perry 2010). However, the much slower development of symptoms in the detached leaves, which were not significantly different from water until 144 HPI, and varied in severity at 192 HPI depending on which isolate was used, may have been due to the amount of inoculum and delivery method used in this study compared to the one performed by Perry (2010). This study inoculated the detached leaves with 10 µl of inoculum, with a directed injury to the adaxial side of the leaf, where the inoculum was left on top of the injury location. This contrasts with the ~200 µl of inoculum used in the previous study, which was forced into the intercellular matrix by infiltration on the abaxial side of the leaf. The use of the infiltration method would likely have caused injury not only to the site of inoculation, but the region surrounding it, causing artificially high symptoms compared to the single small injury location used in this study.

The increased damage potential and resultant effect on symptom development of the infiltration method on attached leaves was one of the motivations for the development of the detached leaf tissue culture method. It also provided a more controllable environment, which was hypothesized to decrease the variability and increase reproducibility of the results, as was observed by Chen et al. (2006) in their detached leaf assay for observing Fusarium graminearum infection in Arabidopsis. However, Chen et al. (2006) required the use of an external disease factor
deoxynivalenol (DON) to provide consistency to the symptoms caused by wounding and inoculation with *F. graminearum*, and a similar option is not available for *Xap* inoculation.

A possible combination of the infiltration and detached leaf methods would be to inoculate detached Arabidopsis leaves with *Xap* isolates using the infiltration method. A similar technique was reported for peach detached leaves by Randhawa and Civerolo (1985). In their study, they used paper towels to cushion and support detached peach leaves while they were infiltrated with *Xanthomonas campestris pv. pruni*. They reported that this resulted in little damage to the leaves, suggesting that infiltration of detached Arabidopsis leaves may be possible with the use of cushioning such as sterilized filter paper.

Another possible explanation of the differences in symptoms seen in the previous study (Perry 2010) and the current study is that the latter used a mixture of all four isolates to inoculate the leaves on the plant, whereas, the current study used individual isolates. The use of the bacterial mixture may have resulted in exposure of the leaf tissue to several effectors, simultaneously, thus enabling infection of Arabidopsis, even though it is not a natural host. This has been documented before in literature, such as the infection of Arabidopsis by non-host *Pseudomonas syringae pv. tomato* DC3000 (Li *et al.* 2005), where the tomato pathovar overcomes non-host defense by the secretion of multiple effectors.
3.6 Conclusion

The use of a detached leaf assay indicate the susceptibility of A. thalianarosette leaves inoculated with Xap isolates could be an important tool when testing CBB resistance genes in the model plant system. This study demonstrated that CBB symptoms and increased Xap CFUs accompany the inoculation of detached Arabidopsis and common bean leaves with Xap isolates at 192 HPI. However, variability in disease response compared to previous common bean and Arabidopsis Xap inoculation studies prevent conclusions to be made about the efficacy of this bioassay. More work needs to be performed to optimize this method.
Chapter 4: Candidate Common Bacterial Blight Resistance Gene Expression in a Resistant (OAC Rex) and Susceptible (Seaforth) Variety of *Phaseolus vulgaris*

4.1 Abstract

Resistance to common bacterial blight, caused by *Xanthomonas axonopodis* pv. *phaseoli* (Xap) is a quantitative trait in common bean (*P. vulgaris*), introduced from tepary bean (*P. acutifolius*) and selected for in common bean to produce resistant varieties such as OAC Rex (Michaels et al. 2006). Although selection for QTLs by their associated markers has been successful in producing resistant varieties of common bean, direct selection for a gene underlying the QTL would provide greater speed and power of selection. This study aimed to measure the expression of candidate CBB R genes CBB2, CBB3 and CBB4 in resistant (OAC Rex) and susceptible (Seaforth) varieties after inoculation with water (control) or Xap, for up to 168 HPI. To measure expression, leaf samples were collected at 0, 12, 24, 36, 48, 72, 96, 120, and 168 HPI, RNA was isolated and cDNA synthesized via RT-PCR was examined by gel electrophoresis. Expression of CGs was not observed consistently, suggesting that none of the CGs are involved in resistance to CBB, or that there were some limitations to the experimental protocol that prevented expression of the CGs to be measured. Further studies are needed to confirm these results, and to identify and test more CGs.
4.2 Introduction

4.2.1 Significance of Common Bean

Common dry bean (P. vulgaris) is a widely grown legume crop species, which is produced in Canada for its dry grain and fresh pods. Edible dry beans are a recommended part of a healthy human diet (Health Canada, www.hc-sc.gc.ca) because they contain high levels of protein (Yañez et al. 1995), dietary fibre and complex carbohydrates, vitamins such as folate, and minerals (Tosh and Yada 2010). They are also low in fats and high in resistant starches. In Canada, common bean is grown in Ontario, Manitoba, and Alberta (Goodwin 2005). Most of the production is primarily exported to the US and Europe, as only a small portion is consumed in Canada.

Common bean production is restricted due to a variety of abiotic factors such as frost, humidity and soil quality as well as biotic factors, in particular weeds and diseases (Goodwin 2005). A combination of management practices and variety selection are employed to reduce the effect of these factors on yield quality and quantity, however there is room for improvement, especially for increased disease resistance.

4.2.2 CBB in Common Bean

One of the major diseases of common dry bean, which affects producers worldwide is CBB, caused by the bacterial pathogen Xap and its fuscans variety Xff (Vauterin et al. 1995). The disease affects producers yearly, with severity dependent on the susceptibility of the variety and environmental factors, such as humidity and temperature, with yield reduction between 10-45% being observed (Saettler 1989; Gillard et al. 2009). Since the symptoms may develop on the leaves, stems, pods and
seeds of infected plants (Vidaver 1993), CBB not only affects yield potential, but the
goodness of the marketable portion of the plant – the seeds and pods.

4.2.3 CG Testing in Common Bean

*P. vulgaris* has little natural resistance to *Xap*. However, close relatives of the
species, *P. acutifolius* (tepary bean) and *P. coccineus* (scarlet runner bean) have
resistant genotypes (Mohan 1982; Drijfhout and Blok 1987; Singh and Muñoz 1999).
These relatives have been used to introduce resistance to CBB into *P. vulgaris* lines,
and have resulted in the production of varieties which are considered resistant, as they
exhibit delayed and/or reduced symptoms. CBB resistance is a quantitative trait in
common bean, with multiple major and minor quantitative trait loci (QTLs) being
mapped in many populations (Miklas *et al.* 2000b; Tar’an *et al.* 2001; Blair *et al.* 2003;
Cordoba *et al.* 2010; Shi *et al.* 2011). However, it has been shown that *P. acutifolius*
has a high level of leaf and pod resistance which is controlled by one dominant gene
and displays a characteristic hypersensitive response (Drijfhout and Blok 1987).

Although mapping of CBB resistance QTLs has helped to identify regions of the
genome that are associated with the trait, specific genes and their expression patterns
in response to *Xap* inoculation have yet to be identified. The aim of this study was to
measure expression from candidate CBB R genes (predicted in Chapter 2) in response
to the inoculation by *Xap* in resistant and susceptible bean cultivars. The study tested
the hypothesis that CBB R genes would be upregulated in leaf samples inoculated with
*Xap* in resistant *P. vulgaris* compared to water inoculated resistant, and water and *Xap-
inoculated susceptible *P. vulgaris*. 
4.3 Materials and Methods

4.3.1 Plant and Inoculum Preparation

4.3.1.1 Plant Growth

A total of 50 plants each for *P. vulgaris* resistant variety (OAC Rex) and susceptible variety (OAC Seaforth) were potted in 5” round pots filled with pre-moistened LA4 mix (Green Island Distributors, Riverhead, N.Y.) at 5 seeds per pot, and covered with medium vermiculite. The pots were soaked with deionized water with fertilizer (20-20-20 N-P-K), and maintained in a growth room providing 16h light (150 µmol/m²/s), 8h dark cycle at 22°C until the uni-foliate stage (~10 days). At this point, all plants from each variety were inoculated, placed in a misting chamber, and the temperature was raised to 28°C daytime temperature and 20°C night temperature, with 80% humidity. A second replication was performed which included 90 plants to accommodate additional time points.

4.3.1.2 Inoculum Preparation

Bacterial inoculum was generated by spreading 50 mL of *Xap* isolate 98 (*Xap*98) on 5 yeast salt media plates with 1% agar, and maintaining them at 28°C for 48 hours. To generate a liquid inoculum, the colonies were washed off the plates with distilled sterile water and diluted with distilled sterile water to an OD₆₀₀ of 0.3. Control inoculum was prepared by autoclaving distilled water.

4.3.1.3 Plant Inoculation

Plant inoculations were performed using the multiple needle method (Yu *et al.* 2000b). Both uni-foliate leaves of each plant were pierced by needles into a sterile
sponge soaked in inoculum. The needles were removed and inoculum soaked sponges were squeezed gently against either side of the punctured leaf.

4.3.2 Sample Collection and Analysis

4.3.2.1 Sample Collection and Preparation

Single whole-leaf samples were taken from three plants per treatment (Xap98 and water control) per time point for OAC Rex and Seaforth varieties respectively. Samples were removed from the plant and immediately frozen in liquid nitrogen to preserve RNA present in the leaf. Samples were taken at 0, 12, 24, 48, 72, 96, 120 HPI for the first replication (January 2013) and 0, 12, 24, 36, 48, 72, 96, 120 and 168 HPI for the second replication (February 2013). Frozen leaf samples were maintained in a -80°C freezer until grinding and RNA isolation.

Whole leaf samples were ground to a fine powder using a mortar and pestle chilled to -80°C by liquid nitrogen. Ground tissue samples were transferred to liquid nitrogen chilled 1.5 mL centrifuge tubes in ~100 mg aliquots and stored in a -80°C freezer until further analysis.

4.3.2.2 RNA Isolation

RNA was isolated from ~100 mg samples of ground leaf tissue using Qiagen RNeasy Plant RNA isolation kit. Aliquots of ground leaf tissue were placed in liquid nitrogen to maintain -80°C temperature and added to 450 µl of RLT buffer – β-mercaptoethanol mixture (1 mL RLT buffer and 10 µL β-mercaptoethanol) then vortexed for ~5 seconds. The mixture was then added to the QIAshredder column and centrifuged for 2 minutes at 15,000 RPM. The supernatant was then mixed with ~225 µL 100% ethanol, the mixture was added to an RNeasy Mini spin column and
centrifuged for 15 seconds at 10,000 RPM. The flow-through was discarded, 700 µL of RW1 buffer was added to the RNeasy spin column and centrifuged for 15 seconds at 10,000 RPM. The flow-through was discarded and 500 µL of RPE buffer was added to the spin column and centrifuged at 15 seconds at 10,000 RPM. This step was repeated with centrifugation lasting for 2 minutes. The spin column was then transferred to a 1.5 µL tube, 50 µL of RNase-free water was added to the column membrane and the tube was centrifuged for 1 minute at 10,000 RPM to elute the RNA.

The RNA concentrations of the eluted samples were measured by applying 1 µL of the RNA sample to the Nanodrop ND-1000 Spectrophotometer (Thermo Fisher Scientific Inc. © 2013, Wilmington DE) pedestal. RNA quality was determined by both the 260/280 (>1.90) and 260/340 (> 1.80) absorbance ratios, and by visualization of band patterns when ~1 µg of each sample was run in 1% agarose gel electrophoresis at 100 V. The sample was considered to be good quality when the most intense band was present at ~1100 bp, with the second most intense band present at ~700 bp.

DNAse digestion was performed using the Ambion® Turbo DNA free (Life Technologies Inc., Burlington, Ontario) kit and protocol. A total of 10 µg of RNA for each sample was mixed with 5 µL of DNAse buffer and 1 µL of Turbo DNAse, mixed and incubated for 25 minutes at 37°C in a water bath. Post-incubation, 5 µL of DNAse inactivation buffer was added to the sample and incubated for 2 minutes at room temperature, then centrifuged for 1.5 minutes at 10,000 RPM. The supernatant was transferred to a new tube, and the concentration and quality of the RNA was determined by applying 1 µL of the sample to the Nanodrop pedestal. Only samples with
absorbance ratios higher than 1.90 at 260/280 and 1.80 at 260/230 were used for cDNA synthesis.

4.3.2.3 cDNA Transcription

Isolated RNA was transcribed to cDNA using the qScript™ cDNA SuperMix. A standard of 1 µg of RNA template was used in each reaction and incubated as per kit instructions in a PCR cycler. Transcription of cDNA from RNA was confirmed by using 2 µL of the cDNA product as the nucleotide template in a PCR reaction using CG specific primers and actin primers. The products were visualized via 1% agarose gel electrophoresis, which was run for 1.25 hours at 100 volts.

4.3.2.4 Primer Design and PCR Amplification

CG primers were designed using PRIMER3 software (http://frodo.wi.mit.edu/primer3) to amplify sections of the predicted coding region for each of the CGs. Primers were designed to amplify sequences either between or within predicted exons of CGs (Figure 4.1). The primers that produced the most intense and distinct bands from genomic DNA were used for RNA expression analysis (Table 4.1). PCR reactions consisted of 2.5 µl 10x PCR Buffer without MgCl₂, 2.5 µl MgCl₂, 1 µl 10 mM dNTP, 1 µl each for reverse and forward primers, 0.5 µl Jumpstart Taq Polymerase (Sigma-Alderich Co.), 14.5 µl sterile distilled H₂O and 2 µl DNA template for a total reaction size of 25 µl. The PCR reaction cycling parameters were initial denaturation at 94°C for 1 minute, 25 cycles of denaturation (94°C for 30 seconds), annealing (temperature dependent on primer, 30 seconds) and extension (72°C for 1 minute), a final extension cycle for 10 minutes at 72°C and held at 4°C until gel electrophoresis. PCR reaction products were loaded with a 100 bp ladder as size reference into a 1%
agarose 1x TBE gel stained with ethidium bromide and run at 100 volts for 1.5 hours in 1x TBE buffer. Visualization was performed using UV light.
Figure 4.1. Locations of RNA expression primers within CGs. For CG CBB2, yellow arrows indicate the locations of reverse and forward primers. For CG CBB3, red arrows indicate the locations of reverse and forward primers. For CG CBB4, green arrows indicate the locations of reverse and forward primers. The red and yellow arrow indicates a shared primer location between CBB2 and CBB3.
Table 4.1. Candidate common bacterial blight R gene primer sequences for real time quantitative RT-PCR. Expected genomic and CDS template product sizes and annealing locations within the CG sequences are indicated.

<table>
<thead>
<tr>
<th>Target Gene</th>
<th>Primer Name</th>
<th>Reverse Sequence</th>
<th>Forward Sequence</th>
<th>Genomic Band Size</th>
<th>CDS Band Size</th>
<th>CG Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBB2</td>
<td>CBB2R1</td>
<td>CCTAAACTAGCCTTCACC</td>
<td>GGCTCAACTACAATTTTCAC</td>
<td>1,627 bp</td>
<td>-</td>
<td>CDS2-CDS6</td>
</tr>
<tr>
<td></td>
<td>CBB2CS</td>
<td>CGAAGCTCTTTAAGTCGC</td>
<td>CGACTCATCCCTAGAAAC</td>
<td>1,635 bp</td>
<td>-</td>
<td>CDS2-CDS6</td>
</tr>
<tr>
<td>CBB3</td>
<td>CBB3R1</td>
<td>GAAACCTTTARATTTAAGG</td>
<td>GTTTATCACCACCTCTTCAT</td>
<td>582 bp</td>
<td>-</td>
<td>CDS3-CDS4</td>
</tr>
<tr>
<td></td>
<td>CBB3CS</td>
<td>CGAAGCTCTTTAAGTCGC</td>
<td>TGCTCGATGAACTAGGGCTCA</td>
<td>1,470 bp</td>
<td>719 bp</td>
<td>CDS2-CDS5</td>
</tr>
<tr>
<td>CBB4</td>
<td>CBB4R3</td>
<td>CAACCTCAAACCTCTTAG</td>
<td>GTTTTGCACTTCTCAAG</td>
<td>1,503 bp</td>
<td>-</td>
<td>CDS3-CDS6</td>
</tr>
<tr>
<td></td>
<td>CBB4CS</td>
<td>TATGTCGAGCTGACCTCCT</td>
<td>GCACCTTGGCCTACTCAG</td>
<td>2,662 bp</td>
<td>1,320 bp</td>
<td>CDS1-CDS6</td>
</tr>
<tr>
<td>ACTIN</td>
<td>Act-R/F (Chen et al. 2008)</td>
<td>TTTCTTGCTTATTCTGTCCG</td>
<td>GAAGTTCTTCCTCAACCATC</td>
<td>-</td>
<td>180 bp</td>
<td>-</td>
</tr>
</tbody>
</table>
4.4 Results

4.4.1 CG Expression

4.4.1.1 CBB Susceptible Genotype (Seaforth)

Initial RNA isolation and cDNA synthesis from Seaforth ground leaf samples from the first experiment replication showed amplification of actin, CBB2, CBB3 and CBB4 for samples taken at 12HPI and 24HPI with primer sets CBB2R1, CBB3R1, CBB4R2 (Figure 4.2). Amplification of a 200bp band by actin primers was observed for both water and *Xap*-inoculated cDNA as well as for the OAC Rex control, and there was no amplification of actin from the water control (Figure 4.2). Amplification of a 550bp band by the CBB2 primers was observed at 12 and 24HPI for *Xap* inoculated samples, but not for water inoculated samples (Figure 4.2). Amplification of CBB2 was observed in the OAC Rex genomic DNA control but not the water control (Figure 4.2). CBB3 primers amplified a 300 bp band from samples taken at both 12 and 24 HPI, when inoculated by water and *Xap98*, as well as the OAC Rex genomic DNA (Figure 4.2). There was no amplification from the water control (Figure 4.2). Amplification of a 200 bp band by CBB4 primers was observed in samples taken at 12 and 24 HPI when inoculated by *Xap98* and water (Figure 4.2). In the case of the water-inoculated 24 HPI sample, the band was indistinct, and covered a larger range of sizes (100-200 bp). A 200 bp band was amplified from OAC Rex genomic DNA, but not the water control (Figure 4.2).

When isolation of RNA and synthesis of cDNA was performed using leaf samples from the first experiment replication for time points 0HPI, 48HPI, 72HPI, 96HPI, and 120 HPI, no amplification of CGs CBB2, CBB3 and CBB4 was observed when
Figure 4.2. Initial CG expression in inoculated *P. vulgaris* leaves. Expression of CGs CBB2, CBB3 and CBB4, were compared to expression of actin check in resistant (OAC Rex) and susceptible (Seaforth) common bean varieties. The inoculum used was *X. axonopodis pv. phaseoli* isolate 98 (X) and sterile distilled water (C). RNA was isolated from leaf samples taken from water and *Xanthomonas*-inoculated OAC Rex and Seaforth plants at 12 (12HPI) and 24 (24HPI) hours post-inoculation. OAC Rex genomic DNA and sterile distilled water were used as positive (+) and negative (-) PCR controls, respectively. CG bands were produced using primers CBB2R1, CBB3R1 and CBB4R2 for CBB2, CBB3 and CBB4, respectively. Band size is indicated by a white arrow for actin, yellow arrow for CBB2, red arrow for CBB3 and green arrow for CBB4.
Figure 4.3. Subsequent CG expression in second experiment replication of *P. vulgaris* leaf inoculation. Expression of CGs CBB2, CBB3 and CBB4 were compared to expression of actin check in resistant (OAC Rex) and susceptible (Seafort) common bean varieties. The inoculum used was *X. axonopodis* pv. *phaseoli* isolate 98 (X) and sterile distilled water (C). RNA was isolated from leaf samples taken from water and *Xanthomonas*-inoculated OAC Rex and Seafort plants at 0, 12, 24, 36, and 48 hours post-inoculation (HPI). OAC Rex genomic DNA and sterile distilled water were used as positive (+) and negative (-) PCR controls, respectively. CG bands were produced using primers CBB2CS, CBB3CS and CBB4CS for CBB2, CBB3 and CBB4, respectively. Band size is indicated by a white arrow for actin, yellow arrow for CBB2, red arrow for CBB3 and green arrow for CBB4.
primers CBB2R1, CBB3R1 and CBB4R2 were used (data not shown). This was consistent for both water and Xap- inoculated samples. Multiple repetitions of the RNA isolation and cDNA synthesis procedures for 12 and 24 HPI samples from both the first and second experiment replications produced no PCR no amplification of CGs CBB2, CBB3 and CBB4 when primers CBB2R1, CBB3R1 and CBB4R2 were used (data not shown). CG primers amplified bands of the previously observed size for CBB2 (550bp), CBB3 (300bp) and CBB4 (200bp) from Seaforth genomic DNA (data not shown).

The bands produced by actin primers with cDNA and genomic DNA templates were either of inconsistent intensity, or not present (data not shown). Therefore, multiple repetitions of cDNA synthesis were performed using samples from all time points from both experiments until consistent quality was obtained. The most consistent cDNA quality was achieved when absorbance ratios above 1.90 (260/280) and 1.80 (260/340) were observed for isolated RNA post DNAse treatment. Consistency of cDNA quality was identified by the presence of actin bands for all samples, although variability in the intensity was still observed (data not shown).

To determine if primers from different locations within the CG sequence were capable of amplifying bands from cDNA, new primers were designed (Table 4.1; Figure 4.1). Newly designed primers (CBB2CS, CBB3CS and CBB4CS) produced no amplification of CGs from synthesized cDNA from samples taken at 0, 12, 24, 48, 72, 96, and 120 HPI for the first experiment replication (Appendix 11) and 0, 12, 24, 36, 48, 72, 96, 120, and 168 HPI for the second experiment replication (Figure 4.3, samples from 0-48 HPI). These primers produced bands of the expected size for CBB2 (1,635 bp), CBB3 (1,470 bp) and CBB4 (2,662 bp) from Seaforth genomic DNA (Figure 4.3).
Amplification of actin was observed for all time points and inoculum types, as well as Seaforth genomic DNA control, but no amplification was observed with the water control (Figure 4.2-4.3).

### 4.4.1.2 CBB Resistant Genotype (OAC Rex)

Initial RNA isolation and cDNA synthesis from OAC Rex ground leaf samples from the first replication showed amplification of actin, CBB3 and CBB4 for samples taken at 12HPI and 24HPI with primer sets CBB2R1, CBB3R1, CBB4R2 (Figure 4.2). Amplification of a 200bp band by actin primers was observed for both water and Xap-inoculated cDNA as well as for the OAC Rex control, and there was no amplification of actin from the water control (Figure 4.2). There was no amplification of a 550bp band by the CBB2 primers at either the 12 or 24HPI for water and Xap-inoculated OAC Rex leaf samples (Figure 4.2). Amplification of CBB2 was observed in the OAC Rex genomic DNA control but not the water control (Figure 4.2). CBB3 primers amplified a 300 bp band from samples taken at both 12 and 24 HPI, when inoculated by Xap98, as well as from the OAC Rex genomic DNA (Figure 4.2). There was faint amplification of a 300 bp band from the water inoculated sample at 24HPI, and no amplification at 12HPI or from the water control (Figure 4.2). Amplification of a 200 bp band by CBB4 primers was observed in samples taken at 12 and 24 HPI when inoculated by Xap98 as well as the OAC Rex genomic DNA control (Figure 4.2). Water-inoculated samples and the water control produced no observed amplification of the 200 bp band (Figure 4.2).

When isolation of RNA and synthesis of cDNA was performed using leaf samples from the first experiment replication for time points 0HPI, 48HPI, 72HPI, 96HPI, and 120 HPI for OAC Rex leaf samples, no amplification of CGs CBB2, CBB3 and CBB4 was
observed when primer sets CBB2R1, CBB3R1, CBB4R2 were used (data not shown). This was consistent for both water and Xap- inoculated samples. Repetition of RNA isolation and cDNA synthesis for 12 and 24 HPI samples from both the first and second replicates produced no PCR amplification of CGs CBB2, CBB3 and CBB4 with the same primers (data not shown). CG primers amplified bands of the previously observed size for CBB2 (550bp), CBB3 (300bp) and CBB4 (200bp) from OAC Rex genomic DNA (data not shown).

The bands produced by actin primers with cDNA and genomic DNA templates were either of inconsistent intensity, or not present (data not shown). Therefore, multiple repetitions of cDNA synthesis were performed using samples from all time points from both experiments until consistent quality was obtained. The most consistent cDNA quality was achieved when absorbance ratios above 1.90 (260/280) and 1.80 (260/340) were observed for isolated RNA post DNAse treatment. Consistency of cDNA quality was identified by the presence of actin bands for all samples, although variability in the intensity was still observed (data not shown).

Since the CGs primers (CBB2R1, CBB3R1, and CBB4R3) were still unable to amplify bands from the cDNA, new primers were designed (CBB2CS, CBB3CS and CBB4CS) (Figure 4.1; Table 4.1). When new primers were used (CBB2CS, CBB3CS and CBB4CS), no amplification was observed for all time points for both water and Xap-inoculated leaves over both replicates (Appendix 11; Figure 4.3; Table 4.1). These primers produced bands of the expected size for CBB2 (1,635 bp), CBB3 (1,470 bp) and CBB4 (2,662 bp) from OAC Rex genomic DNA (Figure 4.3). Amplification of actin was observed from OAC Rex cDNA samples at all time points and inoculation types, as
well as from OAC Rex genomic DNA, but no amplification was observed from water control (Figure 4.3).
4.5 Discussion

4.5.1 Observed CG Expression

Despite initial results at 12 and 24 HPI suggesting that there were differences in expression of candidate CBB CGs CBB2, CBB3 and CBB4 between resistant (OAC Rex) and susceptible (Seaforth) varieties, these results could not be repeated both with the initial (CBB2R1, CBB3R1, CBB4R2) and subsequent primers (CBB2CS, CBB3CS, CBB4CS) (Figures 4.2-4.3). The inability of these primers to show CG expression may be caused by the location of CGs. Since the design and use of the primers used in this experiment (Table 4.1), predictions for exon locations within CBB2, CBB3 and CBB4 have changed, based on the same program used previously to identify exons (FgeneSH). This re-analysis also caused the incorporation of a previously disregarded CG (CBB1) into bioinformatics analyses due to improved gene feature prediction (see Chapter 2). It was also shown that the contig1701 sequence was not continuous, which resulted in the ~20 kb sequence where CBB3 was predicted to be discarded (see Chapter 2).

In the case of primers CBB2R1, CBB2CS, CBB3R1 and CBB4R3, one of the pair of primers is located outside of predicted coding regions, which is the likely cause of the inability to amplify from cDNA (Figure 4.1). This corresponds to the observed differences between expected and actual band sizes produced (Figure 4.2). The initial bands produced by CBB2R1 (~550 bp), CBB3R1 (~300 bp) and CBB4R3 (~170 bp) were the same regardless of whether cDNA or genomic DNA were used. These bands were also significantly smaller than the than the expected genomic DNA size (Figure
This indicates the likelihood that the initial results were produced by DNA contamination, and that the primers were amplifying an unexpected target.

It is also important to note that the both primer sets designed for CBB2 have at least one primer located outside the exon (Figure 4.2). Consequently, these primers would not successfully indicate the presence of CBB2 coding sequence within the synthesized cDNA (Table 4.1). This raises the possibility that although current results indicate that CBB3 and CBB4 are not expressed as a result of Xap inoculation, CBB2 may have been expressed, but the designed primers lacked the ability to show this.

### 4.5.2 Expected CG Expression Patterns

Disease resistance gene expression patterns have been previously characterized in multiple plants and in different plant-pathogen models, with examples including common bean, coffee, orange, Arabidopsis and tobacco (Anderson et al. 2004; Ganesh et al. 2005; Ahn et al. 2011; Borges et al. 2012; Rodríguez et al. 2014). Expression patterns can take a variety of forms and may be different depending on the plant, tissue, pathogen and genes involved as was evidenced by the study performed by Borges et al. (2012). Different expression patterns were observed for multiple genes in the 96 HPI with *Colletotrichum lindemuthianum* (anthracnose). These responses to pathogen recognition include constant expression (PR16a), up-regulation from low expression (PR2), rapid induction of unexpressed genes (PR1b), and down-regulation (PGIP). Another example of resistance gene expression patterns includes the measurement of PAL in *P. vulgaris* inoculated with the same pathogen, *C. lindemuthianum* (Fraire-Velázquez and Lozoya-Gloria 2003). In the resistant reaction, PAL expression was strong for the first 2 HPI, decreased until 16 HPI and then increased more significantly.
Large scale mRNA expression profiling performed by Tao \textit{et al.} (2003), demonstrated that in the Arabidopsis-\textit{Pseudomonas syringae} patho-system, increased gene expression of \~2000 genes occurred at 6-9 HPI in the resistant interaction as opposed to 30 HPI in the susceptible one. Although the CBB resistant source for OAC Rex, \textit{P. acutifolius} accession (P1440795), exhibits a hypersensitive response to CBB (Parker 1985; Drijfhout and Blok 1987), the response to CBB in OAC Rex is characterized by an increased time to development of CBB symptoms, and decreased severity of symptoms. Considering that the resistant reaction in OAC Rex causes retardation in disease progression and not complete immunity, it is hypothesized that this is due to slower pathogen recognition and/or response. Based on this hypothesis, it is therefore expected that the R gene(s) involved would be expressed less strongly or at a later time post-inoculation when compared to the same response in \textit{P. acutifolius}. 
4.6 Conclusion

This study aimed to observe three candidate CBB resistance genes for changes in their expression over time in common bean when resistant (OAC Rex) and susceptible (OAC Seaforth) varieties were inoculated with either water or Xap. Expression of the CGs was not observed within the parameters of the experiment. Further studies of these CGs and further identification and testing of CGs in the area of the PV-ctt001 QTL region will help to progress current knowledge about the mechanisms for CBB resistance and the genes involved.
Chapter 5: Future Directions and Recommendations

5.1 OAC Rex Genome Sequence Assembly

Since a location within the OAC Rex genome has not been identified for contig1701, it is important to continue to search for this. This is highly dependent on the completion of the full OAC Rex genome sequence. The scaffolds and contigs identified in BLAST searches discussed above have yet to be included in the pseudochromosome assembly, making it difficult to identify the location and surrounding region of the candidate genes reported in this study. Once a location for the candidate genes is determined, it will be possible to explore further their position relative to the PV-ctt001 marker as well as other markers and disease resistance genes in the surrounding area. Despite the possibility that this marker was associated to CBB resistance as the result of a false positive (see Chapter 1), the marker provides a targeted region to look at genome details such as the repetitive nature surrounding identified disease resistance genes.

5.2 Candidate CBB R Gene Analysis

5.2.1 Identification and Characterization of New CGs

The evidence indicating the possible retrotransposon characteristics of candidate genes CBB1-4 (see Chapter 2), and the lack of candidate gene expression (see Chapter 3) suggest that identification of new candidate genes would further the aims of identifying CBB resistance genes more effectively. Several potential locations for the identification of candidate resistance genes exist, including the sequence in the regions surrounding CBB resistance markers that was analyzed by Perry et al. (2013).
The sequence characterized by Perry et al. (2013) surrounding CBB resistance QTLs PV-ctt001 (Chromosome 4), SU91 (Chromosome 8) and SAP6 (Chromosome 10) provide a source for identification of additional candidate resistance genes based on predicted function. A total of 114 genes reported within 5 OAC Rex contigs that matched to the PV-ctt001 region of chromosome 4. Of the identified genes, several categories have potential as the source of disease resistance, including defense/signalling/stress (21.9%), unknown (7.9%) and hypothetical (27.5%) genes. The defense/signalling/stress gene predictions especially would be an excellent target group, considering the large percent of the predicted genes it comprises, as well as the fact that they are easily comparable to previously characterized genes. Testing of multiple candidate genes identified in the region surrounding PV-ctt001 to observe expression changes could be performed by real time PCR. Alternatively, genome wide gene expression observations could be performed using RNAseq.

In a previous study, Shi et al. (2012) identified 16 candidate genes associated with CBB resistance marker SU91, which is present in OAC Rex Pv08, to test for increased marker precision. Although two candidate gene markers were reported to have improved association to CBB resistance, this was only related to disease ratings, and not gene expression. The use of these candidate gene markers in RNA expression studies would have more power in confirming the involvement of the associated candidate resistance genes in resistance to CBB. An interesting line of inquiry could explore candidate resistance gene expression identified in both the PV-ctt001 and SU91 QTL regions.
5.2.2 CG Expression Analysis

The inclusion of additional collection time points and tissue sources would increase the ability to make observations about candidate gene expression patterns. Multiple studies have shown the first 24 hours post-inoculation to be a critical period for the induction of disease resistance-related genes, with peaks of induction occurring at a variety of different time points within that critical period (Fraire-Velázquez and Lozoya-Gloria 2003; Tao et al. 2003; Ahn et al. 2011; Borges et al. 2012). However, the expected increased time to resistance gene induction, for the reasons discussed above, suggests that sample collection after a longer time period (perhaps up to 36 HPI), with higher frequency (perhaps every 2-6 hours) would provide the best coverage and increase the likelihood of detecting changes in expression.

5.3 Model Plant CBB Inoculation Procedure

Disease progression was shown in Xap-inoculated Arabidopsis leaves (see Chapter 3), showing that Arabidopsis is a host when manually inoculated. However, repetitions of the initial experiment have been unsuccessful. It is therefore imperative that the study be repeated, to decrease variability and improve the power of any conclusions made. Since the subsequent repetitions used only isolate Xap12, which was shown to be a less virulent isolate, future replications should use Xap18, which produced the highest symptoms and CFU count in Arabidopsis. Future inoculations of detached leaves should also be performed using a mixture of the inoculation method described (see Chapter 3), and the infiltration method used by Randhawa and Civerolo (1985) in peach. The use of a mixture of isolates as an additional inoculum to compare to inoculation with a single isolate (i.e. Xap18) would also indicate whether a mixture of
effectors contributed from the four isolates is what is required for non-host resistance to be overcome.
Chapter 6: Summary

Disease susceptibility is one of the primary limitations to yield for all plant species produced across the globe. One of the major diseases affecting common bean (*P. vulgaris* L.) yield quantity and quality is the disease CBB, caused by bacteria pathogen *Xap* and its fuscans variant *Xff*. The resistant variety OAC Rex provides reduced disease symptoms and increased yield under disease pressure (Michaels *et al.* 2006; Gillard *et al.* 2009). The source of its resistance is a wild relative *P. acutifolius* (tepary bean) accession P1440795 (Parker 1985). Although resistance QTLs have been identified and associated with markers in OAC Rex, such as PV-ctt001 (Tar’an *et al.* 2001; Perry, 2010; Perry *et al.* 2013) and SU91 (Perry *et al.* 2013), no resistance genes have been identified to date. It is important to determine the genes involved in resistance to CBB to increase the current understanding of the resistance reaction, and also to improve selection capabilities for breeding new resistant varieties.

The current study aimed to identify candidate CBB resistance genes within contigs for BAC clone Rex021F10 of OAC Rex selected with the PV-ctt001 marker, and test them in a variety of ways. The characterization of genes identified through homology to previously reported candidate genes revealed that retrotransposon elements comprised a high proportion of the sequence of the selected contig sequences (contig1701 and contig1455). Retrotransposons integrate most often in genic regions (Miyao *et al.* 2003; Le *et al.* 2007), and have been shown to interact with resistance genes by disrupting gene function (Song *et al.* 1997; Hernández-Pinzón *et al.* 2009), causing chimeric proteins (Kashkush *et al.* 2003; Frost *et al.* 2004) changing promoter
function and modifying the epigenetome (Hayashi et al. 2009; McDowell and Meyers, 2013; Tsuchiya and Eulgem, 2013). Characterization of the identified candidate genes revealed retrotransposon features including similarity to GAG/POL poly proteins, reverse transcriptases, and Gypsy/Ty-3 retro-element proteins. Although the similarities to these proteins were not high (30-47%), when combined with the presence of multiple other retrotransposons identified within the same contig indicates the likelihood the CGs are retrotransposons. The presence of several genes in the associated contig sequence with potential disease resistance functions (i.e. ascorbate oxidase, callose synthase, hypothetical protein), provides additional areas for testing additional candidate CBB resistance genes.

Candidate CBB resistance gene expression was tested in susceptible and resistant common bean genotypes. However, no gene expression was observed in cDNA at multiple time points for candidate genes CBB2-CBB4. This could be due to human error during primer design. However, the subsequent characterization of the candidate genes as retrotransposons indicates that expression is not likely to be measured, even with experimental parameter alterations. A more promising target is the putative ascorbate oxidase gene identified in contig1701, since this gene family has previously been shown to be involved in regulating reactive oxygen species levels, which are associated with stress and disease resistance responses (Pavet et al. 2005; Pignocchi et al. 2006; Sanmartin et al. 2007; Fotopoulos et al. 2008).

Since genetic manipulation of common bean is difficult (i.e. genetic transformation), testing candidate CBB resistance gene function within resistant common bean genotypes is an unattractive option. The development of a detached leaf
A. *thaliana* Xap-inoculation method was aimed at providing an alternative method of testing candidate gene function in a susceptible model. Although optimization is still required, symptom progression was shown in the model plant, and provides a more controlled environment to test candidate genes in a more targeted way.
8.0 References


Chan SL, Mukasa T, Santelli E, Low LY and Pascual J. 2010. The crystal structure of a TIR domain from Arabidopsis thaliana reveals a conserved helical region unique to plants. Protein Sci. 19:155–161


Cook, B. J., Clay, R. P., Bergmann, C. W., Albersheim, P. and Darvill, A. G. 1999. Fungal polygalacturonases exhibit different substrate degradation patterns and differ in their susceptibilities to polygalacturonase-inhibiting proteins


170


Goodwin, M. 2005. Crop profile for dry bean in Canada. Agriculture and Agri-Food Canada


Huang, Q and Allen, C. 2000. Polygalacturonases are required for rapid colonization and full virulence of *Ralstonia solanacereaum* on tomato plants. Physiol. Mol. Plant Pathol. 57:77-83


Li, B., and Xu, X. 2002. Infection and development of apple scab (Venturia inaequalis) on old leaves. J. of Phytopath. 150: 687-691


Lu, D., Mu, S., Gao, Zhang, Y., Shan, L. and He, P. 2010. A receptor-like cytoplasmic kinase, BIK1, associates with a flagellin receptor complex to initiate plant innate immunity. PNAS. **107**(1): 496-501


McDowell, J.M. and Meyers, B.C. 2013. A transposable element is domesticated for service in the plant immune system. PNAS. 110(37): 14821-14822


181


Saxton, A.M. 2000. PDMIX800. University of Tennessee, Knoxville TN.


Singh, S. P. 1982. Key for identification of different growth habits of Phaseolus vulgaris L. Centro Internacional de Agricultura Tropical, Apartado Aéreo 6713, Cali Colombia


Tsuchiya, T. and Eulgem, T. 2013. Mutations in EDM2 selectively affect silencing states of transposons and induce plant developmental plasticity. Nature Scientific Reports. 3: 1701 DOI: 10.1038/srep01701


Wang, L., Rong, W. and He, C. 2008a. Two Xanthomonas extracellular polygalacturonases, PghAxc and PghBxc, are regulated by type III secretion regulators HrpX and HrpG and are required for virulence. MPMI 21(5): 555-563


7.0 Appendices

Gene7
G2
Glyma.08G361500 MSSSRGGAGPSSEAPPPRRIMRTQTAGNLGESVIDSEVVPSLVLIAPILKVANEVEKTH

N-Terminal

Gene7
G2
Glyma.08G361500 FRVAYLRCRYAFEAKEAHRLDPNSSGRGVRQFKTALLQDRERENDPTLGRVKKSDAREMQS

Gene7
G2
Glyma.08G361500 MAAWQATFTKPSRNPYATFTN

Gene7
G2
Glyma.08G361500 VAEKTEILVPNIILPDPSANQA1MRPEFIEIQAAVYALRNRGLPWPDKFKKKKDEDILD

Gene7
G2
Glyma.08G361500 PRVAYLRCRYAFEAKEAHRLDPNSSGRGVRQFKTALLQDRERENDPTLGRVKKSDAREMQS

Gene7
G2
Glyma.08G361500 WLGSMFGFQKHKVANQREHLILLLANVTIRFQFPDKQLDERSLMKLYSKW

Gene7
G2
Glyma.08G361500 CKYLGRKSSLWLPTIQEVCQRLLYMGLLHIEAANLRFMPEICLICYIYHMAFELYG

1,3-β Glucan Synthase Subunit FKS1-like

Gene7
G2
Glyma.08G361500 MLAGNVPMTGENVKPACYDEAFRLKVVTPYVIAAEAAORSRKGRXSQWRYNDL

Gene7
G2
Glyma.08G361500 NEYFWSADCFRLGWPMEADAFCLPAEKLVFDKSNDKPPRSRDRVGVNVFVEIRSh

Gene7
G2
Glyma.08G361500 CPSAIVVFNVKVLNVFIC-AILKLGQ-

Gene7
G2
Glyma.08G361500 MFRSFDRWKSFILLCQAMIVVAWNGSDDPSAIFNGDFKKSIFITAAILKFGQAVLD

Gene7
G2
Glyma.08G361500 VILSRKQWSMSMVHKLGYILKVVSANVIVLVSYSYATWENPPRFAQTIQSWFGSN

Gene7
G2
Glyma.08G361500 VILSRKQWSMSMVHKLGYILKVVSANVIVLVSYSYATWENPPRFAQTIQSWFGS

193
Appendix 1. Predicted β-glucan synthase amino acid sequence alignment. ClustalW peptide sequence alignment for contig1701 gene 7, predicted gene in G19833 (G2), and G. max gene annotation (Glyma08G361500). Callose synthase family domains are indicated by boxes. Transmembrane domains are highlighted in grey.
Appendix 2. Alignment of *P. vulgaris* hypothetical proteins. ClustalW alignment of *P. vulgaris* hypothetical gene predictions were identified from OAC Rex and G19833 genome sequence assemblies.

Appendix 3. Protein secondary structure models for hypothetical proteins. Secondary structure models were created in Phyre2 for *P. vulgaris* hypothetical gene 6 (A) identified in OAC Rex contig1701, and Phvul.008G191100 (B) identified in G19833 chromosome 8.

Appendix 4. Phyre2 secondary structure modelling (Kelley and Sternberg 2009) for gene 6 and Phvul.008G191100

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Amino Acids Modelled</th>
<th>Template</th>
<th>Template Characteristic</th>
<th>Similarity to Template</th>
<th>Confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gene 6</td>
<td>23-49</td>
<td>C4ImgD_</td>
<td>Transcriptional activator</td>
<td>33%</td>
<td>15.2%</td>
</tr>
<tr>
<td></td>
<td>11-33</td>
<td>D2f9ca1</td>
<td>Single-stranded left-handed beta-helix fold</td>
<td>29%</td>
<td>32.0%</td>
</tr>
<tr>
<td>Phvul.008G191100</td>
<td>29-65</td>
<td>D1m0da_</td>
<td>Restriction endonuclease-like</td>
<td>27%</td>
<td>12.5%</td>
</tr>
</tbody>
</table>
Appendix 5. Analysis of variance for inoculated detached common bean (P. vulgaris) leaves maintained in a water agar tissue culture system. Significance ($p < 0.05$) of fixed treatment, treatment interaction and error effects are shown. Genotypes include resistant (OAC Rex) and susceptible (Nautica) common bean varieties. Leaves were inoculated with one of four Xanthomonas axonopodis pv. phaseoli isolates (Xap12, Xap18, Xap98 or Xap118) or water control. Leaves were maintained in a water agar tissue culture system at 28°C for up to 192 hours post inoculation (HPI) and samples were taken at different time points post inoculation (0, 48, 96, 144 and 192 HPI).

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>F-value</th>
<th>P &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inoculum</td>
<td>4</td>
<td>36.55</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Genotype</td>
<td>1</td>
<td>47.42</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Time</td>
<td>4</td>
<td>14.73</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Inoculum x Genotype</td>
<td>4</td>
<td>3.89</td>
<td>0.0059</td>
</tr>
<tr>
<td>Inoculum x Time</td>
<td>16</td>
<td>2.57</td>
<td>0.0025</td>
</tr>
<tr>
<td>Genotype x Time</td>
<td>4</td>
<td>4.72</td>
<td>0.0017</td>
</tr>
<tr>
<td>Inoculum x Genotype x Time</td>
<td>16</td>
<td>1.41</td>
<td>0.1530</td>
</tr>
<tr>
<td>Error</td>
<td>41</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Appendix 6. Analysis of variance for inoculated detached Arabidopsis thaliana rosette leaves maintained in a water agar tissue culture system. Significance ($p < 0.05$) of fixed treatment, treatment interaction and error effects are shown. Leaves were inoculated with one of four Xanthomonas axonopodis pv. phaseoli isolates (Xap12, Xap18, Xap98 or Xap118) or water control. Leaves were maintained in a water agar tissue culture system at 28°C for up to 192 hours post inoculation (HPI) and samples were taken at different time points post inoculation (0, 48, 96, 144 and 192 HPI).

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>F-value</th>
<th>P &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inoculum</td>
<td>4</td>
<td>21.57</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Time</td>
<td>4</td>
<td>11.21</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Inoculum x Time</td>
<td>16</td>
<td>2.58</td>
<td>0.0055</td>
</tr>
<tr>
<td>Error</td>
<td>26</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Appendix 7

LSmeans for common bacterial blight symptoms on OAC Rex, Nautica and Arabidopsis leaves at 0 hours post inoculation (HPI). Percent of leaf showing disease symptoms for OAC Rex, Nautica and Arabidopsis leaves inoculated with Water, Xap12, Xap18, Xap98 and Xap118 at 0 HPI.

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Symptoms (%)</th>
<th>Inoculum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Water</td>
</tr>
<tr>
<td>OAC Rex</td>
<td></td>
<td>4.26 x10^{-2} YZ a</td>
</tr>
<tr>
<td>Nautica</td>
<td></td>
<td>0.120 Y a</td>
</tr>
<tr>
<td>Arabidopsis</td>
<td></td>
<td>0.297 Y z</td>
</tr>
</tbody>
</table>

Lsmeans followed by the same lower case letter were not significantly different at p=0.05. Lsmeans followed by the same upper case letter were not significantly different at p=0.05. Denotes lsmeans with a standard error of ±127.80. Denotes lsmeans with a standard error of ±57.1532. Denotes lsmeans with a standard error of ±3.06. Denotes lsmeans with a standard error of ±1.37.

### Appendix 8

LSmeans for common bacterial blight symptoms on OAC Rex, Nautica and Arabidopsis leaves at 48 hours post inoculation (HPI). Percent of leaf showing disease symptoms for OAC Rex, Nautica and Arabidopsis leaves inoculated with Water, Xap12, Xap18, Xap98 and Xap118 at 48 HPI.

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Symptoms (%)</th>
<th>Inoculum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Water</td>
</tr>
<tr>
<td>OAC Rex</td>
<td></td>
<td>0.138 Y a</td>
</tr>
<tr>
<td>Nautica</td>
<td></td>
<td>0.156 Y a</td>
</tr>
<tr>
<td>Arabidopsis</td>
<td></td>
<td>0.255 Y z</td>
</tr>
</tbody>
</table>

Lsmeans followed by the same lower case letter were not significantly different at p=0.05. Lsmeans followed by the same upper case letter were not significantly different at p=0.05. Denotes lsmeans with a standard error of ±127.80. Denotes lsmeans with a standard error of ±57.1532. Denotes lsmeans with a standard error of ±3.06. Denotes lsmeans with a standard error of ±1.37.
### Appendix 9. LSmeans for common bacterial blight symptoms on OAC Rex, Nautica and Arabidopsis leaves at 96 hours post inoculation (HPI). Percent of leaf showing disease symptoms for OAC Rex, Nautica and Arabidopsis leaves inoculated with Water, Xap12, Xap18, Xap98 and Xap118 at 96 HPI.

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Symptoms (%)</th>
<th>Inoculum</th>
<th>Genotype Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Water</td>
<td>Xap12</td>
<td>Xap18</td>
</tr>
<tr>
<td>OAC Rex</td>
<td>0.627^Za</td>
<td>1.42^Ya</td>
<td>3.56^Ya</td>
</tr>
<tr>
<td>Nautica</td>
<td>2.10^Ya</td>
<td>3.20^Ya</td>
<td>4.74^Ya</td>
</tr>
<tr>
<td>Arabidopsis</td>
<td>0.460^Yz</td>
<td>4.44^Yz</td>
<td>5.46^Yz</td>
</tr>
</tbody>
</table>

*LSmeans followed by the same lower case letter were not significantly different at p=0.05.
*LSmeans followed by the same upper case letter were not significantly different at p=0.05.
*XDenotes lsmeans with a standard error of ±127.80.
*WDenotes lsmeans with a standard error of ±57.1532.
*yDenotes lsmeans with a standard error of ±3.06.
*zDenotes lsmeans with a standard error of ±1.37.

### Appendix 10. LSmeans for colony forming units for X. axonopodis pv. phaseoli (Xap) bacteria at 0 hours post inoculation (HPI). CFU was determined by incubating 10^5 dilutions of ground leaf samples taken from inoculated leaves from resistant (OAC Rex) and susceptible (Nautica) P. vulgaris varieties, and Arabidopsis. Samples were taken at 0 HPI after inoculation with Water, Xap12, Xap18, Xap98 or Xap118.

<table>
<thead>
<tr>
<th>Genotype</th>
<th>CFU (x 10^5)</th>
<th>Inoculum</th>
<th>Genotype Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Water</td>
<td>Xap12</td>
<td>Xap18</td>
</tr>
<tr>
<td>OAC Rex</td>
<td>0^Za</td>
<td>173^Xa</td>
<td>0^Ya</td>
</tr>
<tr>
<td>Nautica</td>
<td>0^Xa</td>
<td>280^Xa</td>
<td>430^Xa</td>
</tr>
<tr>
<td>Arabidopsis</td>
<td>0^Yz</td>
<td>0^Yz</td>
<td>217^Yz</td>
</tr>
</tbody>
</table>

*LSmeans followed by the same lower case letter were not significantly different at p=0.05.
*LSmeans followed by the same upper case letter were not significantly different at p=0.05.
*XDenotes lsmeans with a standard error of ±174.
*WDenotes lsmeans with a standard error of ±77.9.
*yDenotes lsmeans with a standard error of ±96.3.
*zDenotes lsmeans with a standard error of ±43.1.
Appendix 11. Candidate gene expression in first experiment replication of *Phaseolus vulgaris* leaf inoculation. Expression of candidate genes CBB2, CBB3 and CBB4 were compared to expression of actin check in resistant (OAC Rex) and susceptible (Seaforth) common bean varieties. The inoculum used was *X. axonopodis pv. phaseoli* isolate 98 (X) and sterile distilled water (C). RNA was isolated from leaf samples taken from water and *Xanthomonas*-inoculated OAC Rex and Seaforth plants at 0, 12, 24, and 48 hours post-inoculation (HPI). OAC Rex genomic DNA and sterile distilled water were used as positive (+) and negative (-) PCR controls, respectively. Candidate gene bands were produced using primers CBB2CS, CBB3CS and CBB4CS for CBB2, CBB3 and CBB4, respectively. Band size is indicated by a white arrow for actin, yellow arrow for CBB2, red arrow for CBB3 and green arrow for CBB4.