Identification and Genetic Mapping of Clubroot Resistance in Two *Brassica nigra* Lines

A Thesis Submitted to the College of

Graduate Studies and Research

In Partial Fulfillment of the Requirements

For the Degree of Master of Science

In the Department of Biology

University of Saskatchewan

Saskatoon

By

Adrian Chang

PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a postgraduate degree from the University of Saskatchewan, I agree the libraries of this university may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Biology

University of Saskatchewan

112 Science Place

Saskatoon, Saskatchewan S7N 5E2

Canada

ABSTRACT

Clubroot, caused by the protist *Plasmodiophora brassicae* Woronin, is one of the most serious diseases to affect members of the plant family Brassicaceae. This biotrophic soil-borne pathogen is an emerging threat to canola and mustard production in western Canada. To manage the disease, it is important to identify and use new sources of clubroot resistance (CR). The purpose of this research is to map CR genes in the mustard species, *Brassica nigra*, and to analyze host differential gene expression during *P. brassicae* infection.

Lines of *Brassica nigra* with a broad spectrum of resistance to clubroot were recently identified. Plant materials utilized in this study include resistant (R) lines BRA192/78 and PI 219576; and susceptible (S) line CR2748. The two resistant lines of *B. nigra* were crossed with CR2748 female susceptible line to produce the F_1. F_1 plants were self-pollinated to produce F_2. F_1 plants from CR2748 x PI 219576 and CR2748 x BRA192/78 were backcrossed with the CR2748 plants to produce BC_1 populations. Genetic mapping of CR genes was carried out using bulked segregant RNA-Sequencing (RNA-Seq). Validation and genotyping of single nucleotide polymorphism (SNP) markers were carried out using the Kompetitive Allele Specific PCR (KASP) method. Complete resistance to clubroot was found in all F_1 plants derived from crosses of CR2748 with PI 219576 or BRA192/78. Evaluation for resistance to clubroot showed ratios of 1R:1S in BC_1 and 3R:1S in F_2 for both R genotypes, indicating that CR is controlled by a single dominant gene in both PI 219576 and BRA192/78. Short reads from R and S bulked RNA-Seq samples in the BC_1 population derived from PI 219576 and BRA192/78 were assembled into the *B. rapa, B. oleracea* and *B. nigra* reference genomes. Transcriptome analysis was conducted, with 11 differentially expressed genes (DEGs) identified in PI 219576 and 382 DEGs identified in BRA192/78 at 95% confidence. Several of the DEGs annotated were involved in CR, such as

genes encoding Toll-Interleukin-1 receptor / nucleotide-binding site / leucine-rich-repeat (TIR-NBS-LRR)-class disease resistance proteins and pathogenesis-related (PR) transcriptional factors. An ethylene-responsive gene was up-regulated, whereas auxin-responsive genes and a gene associated with cell growth/development were down-regulated in resistant plants. A CR gene designated *Rcr6a* in PI 219576 was mapped to the region between 14.36 Mb and 14.84 Mb on chromosome B3, in a region homologous to one on chromosome A08 of *B. rapa*. SNP markers closely linked to *Rcr6a* were developed. Plants in the BC_1 population from BRA192/78 were also analyzed with the SNP markers linked to *Rcr6a* and results showed that resistance in BRA192/78 was linked on B3 chromosome, in the region between 12.76 Mb and 14.84 Mb, indicating that the CR gene namely *Rcr6b* in BRA192/78 was likely in the *Rcr6a* region. This is the first report on mapping of two single dominant CR alleles, *Rcr6a* and *Rcr6b*, in *B. nigra*. In addition, several SNP markers closely linked to the genes were developed and these markers will be useful in marker-assisted breeding for clubroot resistant canola cultivars.

There has been limited research done on the *B. nigra* genome compared to the vegetable species, *B. rapa* and *B. oleracea*, and its potential for CR is poorly understood. However, some *B. nigra* accessions are highly resistant to clubroot, and CR genes in *B. nigra* may be transferred into canola (*B. napus*) as well as mustard species, *B. juncea* and *B. carinata*. The identification of DEGs is a significant step in better understanding CR mechanisms so CR genes with potentially different modes of action against clubroot can be utilized.

ACKNOWLEDGEMENTS

I am very grateful to have worked through my M.Sc. program under the guidance of my supervisors; ***Dr. Yangdou Wei****, Department of Biology, University of Saskatchewan, Saskatoon and* ***Dr. Fengqun Yu****, Agriculture and Agri-Food Canada (AAFC), Saskatoon.*

I would like to express my sincere gratitude for the time Dr. Fengqun Yu devoted to this project and for her inspiring guidance, encouragement and invaluable advice throughout the program. Her motivation and immense patience for introducing me to the world of research and thesis writing is very much appreciated. I would also like to express my sincere gratitude to Dr. Yangdou Wei for his constructive advices, support and scientific discussions throughout my time as a graduate student.

I am very thankful to my advisory committee members ***Dr. Art Davis*** *and* ***Dr. Jose Andres****, Department of Biology, U of S, Saskatoon for their cooperation, guidance and kind assistance to successfully complete my courses and research work. I genuinely acknowledge Dr. Gary Peng and Mrs. Linda McGregor, AAFC, Saskatoon for their substantial support and guidance. I would also like to acknowledge Dr. Isobel Parkin, Dr. Kevin C. Falk and Dr. Markus Hecker for providing valuable resources to my research work.*

I would like to extend my uppermost gratitude to all the great colleagues from AAFC and U of S who have kindly shared their invaluable knowledge and the fulfilling friendships who made all this time exceptionally worthwhile - Dr. Hao Hu, Dr. Tao Song, Wali Soomro, Dr. Xunjia Liu, Dr. Masud Karim, Dr. Abdulsalam Dakouri, Dr. Yan Zhang, Dr. Qilin Chen, Dr. Mingguang Chu, Dr. Zhen Huang, Dr. Xingguo Zhang, James Bush, Yuhao Wu and many others. Lastly, I want to extend my full gratitude to AAFC for funding this project.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

X^2	Chi-square test statistics
°C	Degrees celsius
µL	Microlitre
µM	Micromolar
ABI	Applied Biosystems, Inc.
AFLP	Amplified fragment length polymorphism
BC_1	First back-cross generation
BC_2	Second back-cross generation
bp	Base pair
BRAD	*Brassica* database
BSA	Bulked segregant analysis
BSR-Seq	Bulked segregant RNA-Seq
CAPS	Cleaved amplified polymorphic sequences
cDNA	Complementary deoxyribonucleic acid
cm	Centimetre
cM	CentiMorgan
CNV	Copy number variation
CR	Clubroot resistance
CTAB	Cetyl trimethyl ammonium bromide
DEGs	Differentially expressed genes
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate

dpi	Days post-inoculation
EDTA	Ethylenediaminetetraacetic acid
ET	Ethylene
F_1	First filial generation
F_2	Second filial generation
g	Gram
Gb	Gigabyte
GBS	Genotyping-by-sequencing
GC	Growth chamber
GH	Greenhouse
GO	Gene ontology
HR	Hypersensitive reaction
IAA	Indole-3-acetic acid
InDel	Insertion deletion
ISSR	Inter-simple sequence repeat
JA	Jasmonic acid
K	Thousand
KASP	Kompetitive allele specific PCR
L	Litre
M	Million
MAS	Marker-assisted selection
Mb	Million bases
MBS	Mapping-by-sequencing
min	Minute

mL	Millilitre
mM	Millimolar
NaCl	Sodium chloride
NaOH	Sodium hydroxide
ng	Nanogram
NGS	Next-generation sequencing
PCR	Polymerase chain reaction
PR	Pathogenesis-related
PSA	Pooled sample assembly
qPCR	Quantitative polymerase chain reaction
qRT	Quantitative real-time
QTL	Quantitative trait loci
R	Resistant
RAPD	Randomly amplified polymorphic DNA
RFLP	Restriction fragment length polymorphism
RNA	Ribonucleic acid
RNA-Seq	RNA-Sequencing
RPKM	Reads assigned per kilobase of target per million mapped reads
rpm	Revolutions per minute
S	Susceptible
SA	Salicylic acid
SCAR	Sequence characterized amplified region
SDS	Sodium dodecyl sulfate
sec	Second

SNP	Single nucleotide polymorphism
SSR	Simple sequence repeat
STS	Sequence tagged site
Taq polymerase	Polymerase from the bacterial species *Thermus aquaticus*
TF	Transcription factor
TIR-NBS-LRR	Toll/interleukin-1 receptor-nucleotide binding site-leucine-rich repeat
WGS	Whole genome shotgun
wpi	Weeks post-inoculation

1. INTRODUCTION

Members of the genus *Brassica* in the family Brassicaceae are important to the human diet and are also of industrial significance. Ancient civilizations originally cultivated *Brassica* in the Mediterranean and in Asia; recently, world-wide cultivation of *Brassica* crops has been on the rise (Branca and Cartea, 2011). Wild *Brassica* plants were used for medicine, cosmetics and perfumery in the past. To this day, medicinal plants are still an essential source for health care and pharmaceutical industries. The most abundant *Brassica* phytochemicals include indole phytoalexins, phenolics and glucosinolates; all of which play a role in the antioxidant and anticarcinogenic activities in *Brassica* crops (Podsedek, 2007; Jahangir et al., 2009). Modern uses of *Brassica* include use of traditional therapeutic compounds in medicine, as food supplements in the human diet to improve overall health, and in industrial processes such as production of biodiesel fuel. *Brassica* is the most economically significant genus within the Brassicaceae; thus, it is important to enhance security and protection of this staple crop. Clubroot in *Brassica* crops can cause devastating yield loss, so control methods are necessary to mitigate disease. Genetic control of clubroot through introgression of clubroot resistance (CR) genes identified in clubroot resistant cultivars of *Brassica* species is efficient and can reduce the use of fungicides, thereby reducing their negative environmental impact (Saito et al., 2006).

The objective of this research is to map CR genes at high resolution in *Brassica nigra* to facilitate map-based gene cloning. It is therefore essential to develop robust genetic markers tightly linked to the CR genes for use in marker-assisted selection (MAS) in canola and mustard breeding programs. In addition, objectives include utilizing RNA-Seq to mine for SNPs and analyzing the disease reaction of resistant and susceptible *B. nigra* lines following *P. brassicae*

infection by identifying up- and down-regulated differentially expressed genes (DEGs). A graphical representation of the experimental flow during this study is shown in Figure 1.1.

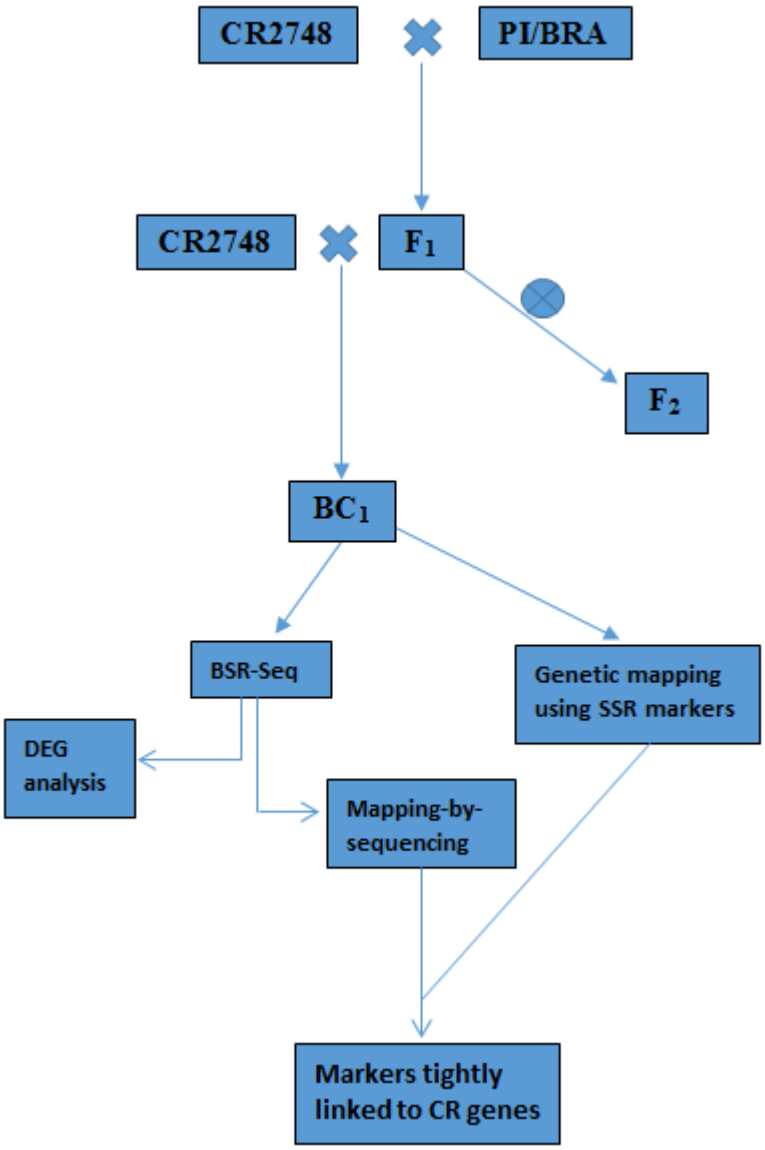

Figure 1.1: An overview of the experimental flow during the study. The research started with developing F_1, F_2 and BC_1 populations for two *B. nigra* cultivars (PI 219576 and BRA192/78), followed by phenotyping. The F_2 populations were analyzed to confirm the segregation ratio of 3R:1S. The BC_1 populations were used to confirm the segregation ratio of 1R:1S and to carry out genetic mapping work. The conventional approach was done by using simple sequence repeat (SSR) markers. The other approach was done via mapping-by-sequencing (MBS) through bulked segregant RNA-Seq (BSR-Seq). Analysis of differentially expressed genes (DEGs) was also conducted with BSR-Seq method.

2. LITERATURE REVIEW

This literature section will review the *Brassica* species within the family Brassicaceae and their economic importance around the world, including Canada. Background on clubroot disease will be discussed, including methods utilized to mitigate the disease. Since the main interest of this thesis is on mapping of clubroot resistance genes in *Brassica nigra*, different mapping approaches and genetic marker based studies are also reviewed.

2.1 *Brassica* species and "The Triangle of U"

Brassica species are significant in the horticultural and agricultural sectors because they are major oilseed and vegetable crops grown worldwide, except the tropics. The genomic relationship between *Brassica* crop species are commonly represented by the "Triangle of U" (U, 1935) (Figure 2.1). This consists of *Brassica nigra* (L.) Koch (genome BB; *n* = 8), *Brassica oleracea* L. (CC; *n* = 9) and *Brassica rapa* L. (AA; *n* = 10), which are diploid species. Also included are *Brassica carinata* A. Braun (BBCC; *n* = 17), *Brassica napus* L. (AACC; *n* = 19) and *Brassica juncea* (AABB; *n* = 18), which are amphidiploid species derived from interspecific hybridization between the corresponding pairs of the diploid species (U, 1935). The *B. nigra* B-genome is 632 Mbp, *B. oleracea* has the largest genome of the diploid *Brassica* at 696 Mbp whereas *B. rapa* has the smallest genome at 529 Mbp. *Brassica juncea* has a genome size of 1068 Mbp, the *B. napus* genome size is 1132 Mbp and *B. carinata* is 1284 Mbp (Johnston et al., 2005). *Brassica rapa* species include Chinese cabbage, turnip, mizuna, field mustard and yellow sarson. Chinese cabbage is an economically important vegetable in Asia; sarson turnip is a minor crop in Europe and in New Zealand; turnip greens and turnip tops are widely used in Portugal and Spain for human nutrition; and turnip rape seed is commonly used for oilseed production in North America. *Brassica oleracea* species include kale, cauliflower, cabbage, broccoli and cabbage

crops (Padilla et al., 2005). *Brassica nigra*, also known as black mustard, is commonly used as a condiment. *Brassica carinata*, Ethiopian mustard, is an emerging crop on the Canadian prairie that could provide sustainable non-food raw material such as biomass or oil. *Brassica juncea* is a major oilseed crop of the Indian sub-continent grown on 6-7 million hectares of land (Chauhan et al., 2011); it is also an important condiment mustard crop and this species can produce low acid food grade oil (Rakow, 2007). *Brassica napus* is an oilseed crop that is important for edible oil production in Canada, Europe and China, and has a high demand (Paritosh et al., 2014). The Canadian canola industry has a target of 26 million tonnes of sustained market demand and production by 2025 (Canola Council of Canada, 2014).

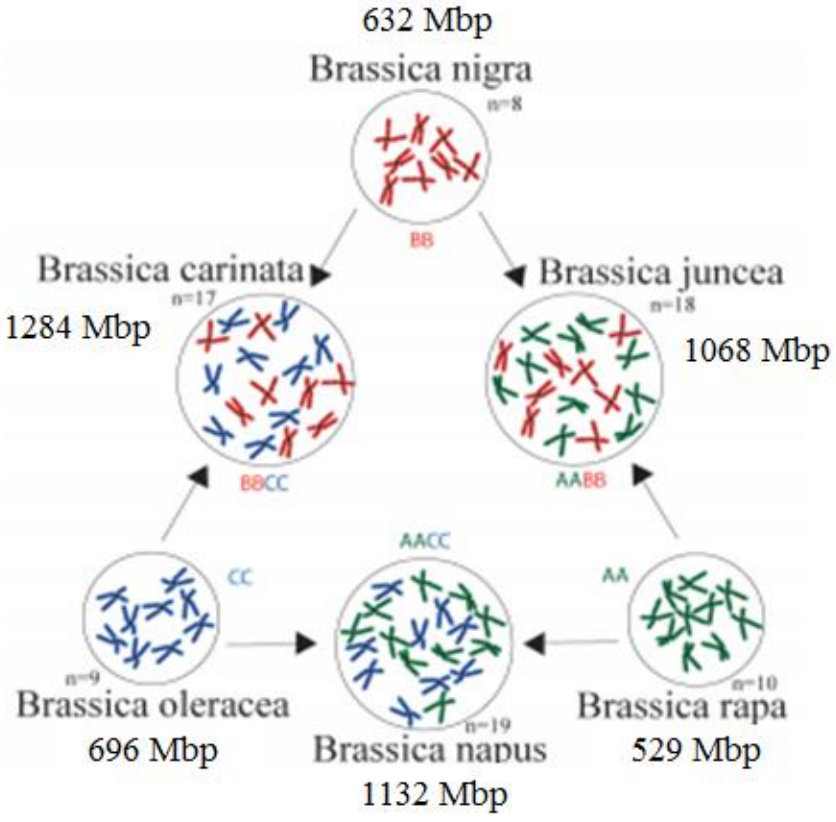

Figure 2.1: The "Triangle of U" diagram showing the genetic relationships between six species of the genus *Brassica* (U, 1935). Chromosomes are represented by different colours in each genome (A, B and C). The diploid species include *Brassica nigra* (genome BB; *n* = 8), *B. oleracea* (CC; *n* = 9) and *B. rapa* (AA; *n* = 10). The amphidiploids include *B. carinata* (BBCC; *n* = 17), *B. juncea* (AABB; *n* = 18) and *B. napus* (AACC; *n* = 19).

Genome similarity is required in order for chromosome pairing and genetic recombination to occur (Leflon et al., 2006). The *Brassica* species are closely related, so inter-specific crosses are believed to have led to the formation of three new species of tetraploid *Brassica.* Introgression of economically important traits from one species into another species is easily facilitated. Modern scientific approaches, especially molecular biology, have substantially improved cultivated *Brassica* accessions. Molecular markers in MAS and breeding have been developed, and genetic transformation has enabled rapid selection of desirable *Brassica* agronomic traits (Branca and Cartea, 2011). The advancement of in vitro techniques including ovary and embryo tissue culture may overcome post-fertilization obstacles in distantly related plant species. In addition, somatic hybridization techniques allow hybridization to occur when pre-fertilization barriers exist, inducing recombination of cytoplasmic and nucleic genomes (reviewed by Glimelius, 1999; Christey, 2004). Economically important traits such as crop yield, pest resistance and abiotic stress tolerance can be rapidly introgressed from wild *Brassica* species to cultivated accessions. Therefore, applications of modern techniques in plant breeding programs can broaden the genetic potential in *Brassica* crops. In particular, genetic maps of *Brassica* species have been produced to facilitate progress in genetics and crop breeding. Genetic maps also allow comparisons of genetic relationships between *Brassica* crops and their related wild species (Branca and Cartea, 2011). *Brassica* species are closely related to the model plant *Arabidopsis thaliana* (Brassicaceae). The average identity of exon sequences at the nucleotide level between *Arabidopsis* and *Brassica* was reported to be around 87% (Cavell et al., 1998). Due to the close identity, the completed genome of *Arabidopsis* provided an essential resource for identifying genes relevant to agronomic and nutritional traits of *Brassica* species (Lan and Paterson, 2000; Li and Quiros, 2003). The A and C genomes of cultivated *Brassica* species are

also completely sequenced and can be used as a valuable reference (reviewed by Sharma et al., 2014).

2.2 Clubroot disease

2.2.1 *Plasmodiophora brassicae*: an emerging pathogen of *Brassica* crops

Clubroot is one of the most serious diseases that affect members of the family Brassicaceae. The soil-borne protist pathogen, *Plasmodiophora brassicae*, is the cause of clubroot, infecting a wide range of *Brassica* crop plants. Originally identified by Russian biologist Mikhail Stepanovich Woronin in 1878, *Plasmodiophora* belongs to the class Phytomyxea which is part of the protist supergroup Rhizaria (Cavalier-Smith and Chao, 2003). There are at least six pathotypes of *P. brassicae* (2, 3, 5, 5X, 6 and 8) identified in Canada, with pathotype 3 being predominant and most virulent on canola in western Canada. Pathotypes 2 and 6 are more common in eastern Canada, affecting cruciferous vegetables and canola crops (Strelkov et al., 2006). Pathotype 5X is a new race of clubroot disease that was recently discovered in a field north of Edmonton in 2013, which can overcome resistance in the canola cultivars available on the Canadian market (Strelkov et al., 2016a). When seedlings are infected by *P. brassicae*, plant death often occurs. However, this is not necessarily the case when infection happens later in plant development. Plants with clubroot infection are stunted compared to healthy plants because the root system forms club-shaped galls (Figure 2.2) that restrict the roots from taking up water and nutrients. Other above-ground symptoms of infected plants include foliar wilting, yellowing and premature senescence (Dixon, 2004).

Figure 2.2: A *Brassica nigra* plant (left) infected with *Plasmodiophora brassicae.* This 5-week old plant extracted from *P. brassicae* contaminated soil and displays the characteristic root swellings (galls) or clubs of the disease. Resistant *B. nigra* control of line PI 219576 is shown on the right.

Clubroot galls contain millions of resting spores that can survive in the soil for up to 20 years, with a half-life of around 4 years (Wallenhammar, 1996). The three stages of the pathogen life cycle consist of resting spore survival in the soil, primary infection of the root hairs and secondary infection of the cortical cells. (Naiki and Dixon, 1987). Resting spores germinate and release an oval-shaped biflagellate zoospore in response to host root exudate secretion. These motile zoospores infect host-plant root hairs, by penetrating the cell wall using mechanical force. Upon the establishment of primary infection in the root hair, primary plasmodia are formed. Plasmodia cleavage produces zoosporangia, each containing 4-16 secondary zoospores, which are released into the soil. The secondary infection stage occurs when the secondary zoospores penetrate to the root cortical cells. Establishment of the secondary infection produces secondary

plasmodia. It is at this time club-shaped galls appear on the roots, coinciding with the development of above-ground disease symptoms. Breakdown of root tissues allow resting spores to be released into the soil. Since zoospore motility is limited in the soil, the pathogen cannot spread quickly and it usually completes one cycle per growing season in western Canada (reviewed by Hwang et al., 2011). An overview of the life cycle of *P. brassicae* is shown in Figure 2.3.

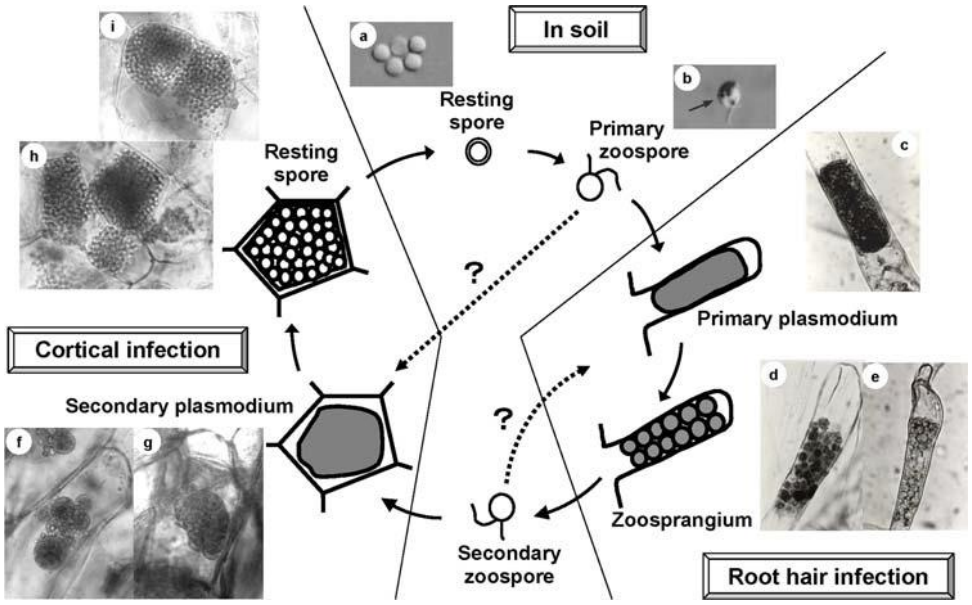

Figure 2.3: The life cycle of *Plasmodiophora brassicae* (Kageyama and Asano, 2009). The pathogen progresses through its life stages as directed in **a-i**. **a:** Resting spore. **b:** Primary zoospore. **c:** Primary plasmodium in root hair. **d:** Zoosporangial cluster in root hair. **e:** Empty zoosporangium. **f, g:** Secondary plasmodia in cortical cells. **h, i:** Resting spores in cortical cells.

Different approaches have been carried out to detect and quantify *P. brassicae* inoculum levels in soil and host roots. A common method is to use bait plants grown in infected soil for 5-6 weeks, which are then extracted from the soil (Figure 2.2) and examined for the presence of root galls (Faggian and Strelkov, 2009). The major obstacle this method fails to overcome is the effect of varying environmental conditions producing different results. An alternative technique is the microscopic examination of root hairs to evaluate root hair infection, but this method is also time-

consuming and requires professional training (reviewed by Hwang et al., 2011). The PCR-based test developed by Cao et al. (2007) for the detection of *P. brassicae* has been used more commonly in recent years. A PCR-based protocol was established where *P. brassicae* primers and reaction conditions are optimized to amplify the DNA only from the pathogen to give a reliable detection (Cao et al., 2007). This process has been commercialized for the routine detection of pathogen in plant and soil materials, and is able to detect 1 x 10^3 resting spores per gram of soil. This method provided a qualitative measure of *P. brassicae*, but not the amount of inoculum present. The introduction of quantitative PCR (qPCR) made it possible to quantify *P. brassicae* inoculum levels (Sundelin et al., 2010; Peng et al., 2015; Deora et al., 2015).

2.2.2 Epidemiology and economic impact of clubroot disease

Clubroot is widespread throughout the world (reviewed by Dixon, 2009). It is believed the Fertile Crescent was the origin of clubroot disease. The presence of *P. brassicae* in China and Japan may have occurred by the movement of ancestral *B. rapa* types (along with the pathogen) from the Fertile Crescent. Clubroot disease has been reported in Europe since the 16th and 17th centuries, and the frequency of reports increased substantially in the 19th century as agriculture intensified and expanded to feed the rising populations during the Industrial Revolution. The spread of this pathogen to the Americas was due to diseased animal fodder transported by colonists (reviewed by Dixon, 2009). Clubroot has been established on canola and mustard in western Canada (Howard et al., 2010). This occurrence is a concern, as western Canada is one of the leading producers of rapeseed (Figure 2.4). The disease has also made an impact on cruciferous vegetables, especially in regions where vegetable production is high such as provinces of British Columbia, Ontario and Quebec (Howard et al., 2010). Clubroot on canola was initially discovered in fields in central Alberta in 2003 (Strelkov et al., 2006). The affected

area has spread with the latest official count being 2,153 clubroot-infested fields confirmed in 2015 (Strelkov et al., 2016b). Infected fields are concentrated in the central part of the province, but some cases have been reported in southern Alberta (Figure 2.5). The Manitoba map for clubroot distribution is shown in Figure 2.6, where municipalities in red have had clubroot symptoms observed in at least one field or have had at least one soil sample with spore levels above 80,000 per gram. The presence of *P. brassicae* inoculum has also been reported in at least one field in the neighbouring province, Saskatchewan (Strelkov et al., 2016b; Cao et al., 2009; Dokken-Bouchard et al., 2010).

In Alberta, estimated yield losses in clubroot infected canola crops have ranged from 30% to 100%. Seed quality was reduced, with 4.7-6.1% decline in oil content and 13-26% decline in 1000-seed weights. Oilseed rape production has also been reduced in China due to this pathogen, with 17% loss of young plants, 15% loss of mature plants, and 10.2% yield loss. In Australia, clubroot can cause at least 50% loss of oilseed rape (reviewed by Hwang et al., 2011). This obligate parasite is present in more than 60 countries where *Brassica* crops are grown and crop yields can be reduced by 10-15% worldwide (reviewed by Dixon, 2009).

As a consequence of clubroot disease, *Brassica* crops face rejection from supermarket buyers due to delayed maturity caused by disease, loss of quality in appearance and size, or may even lose marketability due to unsightly root galls. Fields infested with clubroot result in surviving plants that grow unpredictably and can cause huge variation in sizes. Land management becomes more challenging and costly because clubroot disease complicates crop rotations and transportation of machinery. The asset value of land plummets once it is infected by the pathogen and leads to reduced opportunities for leasing to growers and farmers (reviewed by Dixon, 2009).

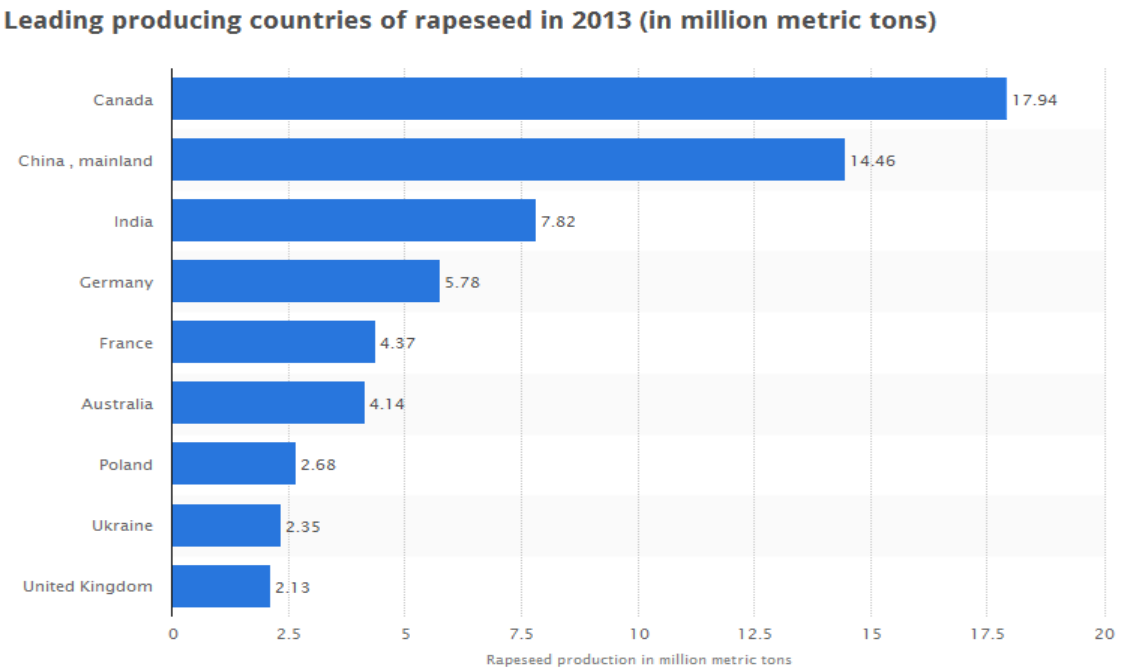

Figure 2.4: The nine leading producers of rapeseed world-wide. Canada is the largest producer of rapeseed, at 17.94 million metric tons in 2013 (Statista, 2013).

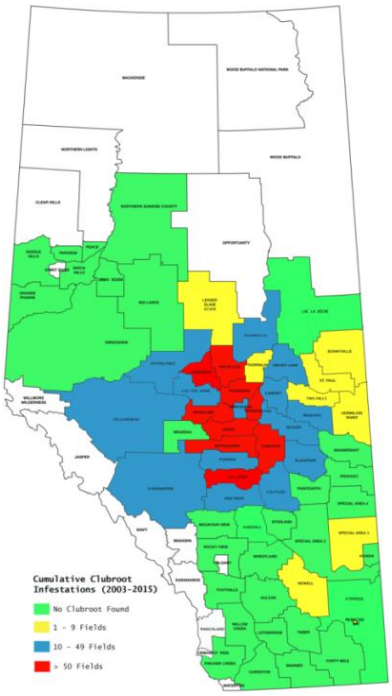

Figure 2.5: The occurrence of clubroot on canola in Alberta as of November 2015 (Strelkov et al., 2016b). Green regions indicate no clubroot found. Clubroot disease was found on canola in yellow (1-9 fields), blue (10-49 fields) and red regions (more than 50 fields).

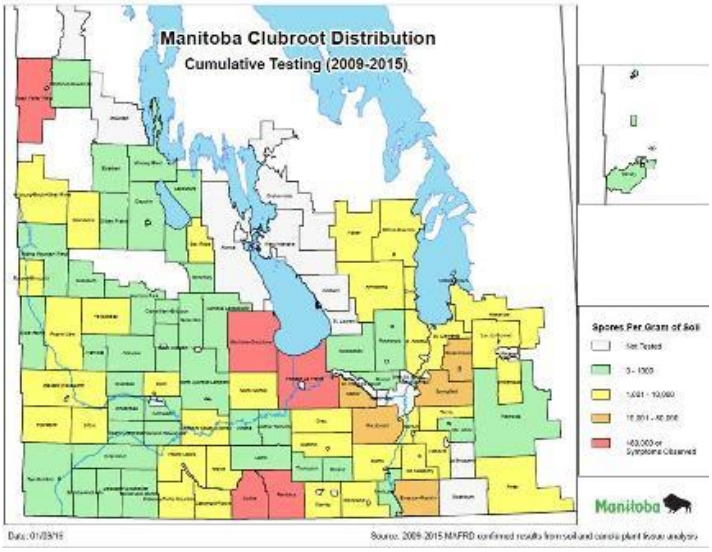

Figure 2.6: Clubroot distribution in Manitoba (MAFRD, 2016). White regions were not tested. Spores per gram of soil found in green (0-1000), yellow (1001-10000), orange (10001-80000) and red regions (more than 80000).

2.2.3 Clubroot control management

To manage clubroot disease effectively, scientific understanding of disease population dynamics and their resistance sources must be maintained. With current *P. brassicae* pathotypes identified, it is important CR be managed by implementing various strategies. This sustained management is necessary because *P. brassicae* populations can shift rapidly in response to the selection pressure by resistant cultivars (Diederichsen et al., 2009). Additionally, resting spores of *P. brassicae* are long-lived in the soil and can spread easily through soil tag, footwear, machinery, grazing animals, transplants, and surface floodwater (reviewed by Donald and Porter, 2009).

Integrated clubroot control involves knowledge of factors including soil type, climatic conditions, and basic site history. The practice of removing diseased plants and crop rotation with non-host species are common methods to avoid clubroot (reviewed by Donald and Porter, 2009). Peng et al. (2015) reported that a greater than 2-year crop rotation reduced resting spores of *P.*

brassicae in soil and the impact of clubroot on canola. In recent years, a molecular technique (qPCR) has been used to rapidly detect and quantify *P. brassicae* to assist with predicting disease and expected yield loss (Sundelin et al., 2010).

Increasing pH levels by raising soil alkalinity is a commonly used practice to manage clubroot. Gossen et al. (2014) demonstrated that a higher pH or alkaline condition in the soil reduces clubroot development. A pH of 7.2 has been known to reduce incidence and severity of disease, but several control failures around this pH level have been reported (reviewed by Donald and Porter, 2009). Mineral nutrition has also been shown to have an impact on disease development; for example, Palm (1963) indicated that potassium stimulated development of gall tissue and that calcium reduced root hair infection by strengthening epidermal cell walls of the host.

Chemical control of *P. brassicae* has been frequently utilized since the growth of the agrochemical industry, but restrictions on the use of chemicals have been implemented due to health and environmental concerns; such chemicals include mercurous chloride (Calomel^{TM}), anbenzimidazoles, and cyazofamid (Ranman^{TM}) (reviewed by Donald and Porter, 2009). Peng et al. (2011a) showed that biocontrol agents such as *Bacillus subtilis* and *Gliocladium catenulatum* can reduce clubroot severity by more than 80% on a highly susceptible canola cultivar 'Fortune RR' inoculated with *P. brassicae*. A similar effect was also observed with the application of the fungicides, fluazinam and cyazofamid.

It was reported that temperature has an effect on infection and subsequent development of clubroot under controlled conditions (Gossen et al., 2012). Sharma et al. (2011) showed that root hair infection was highest and visual symptoms of clubroot observed earliest at 25°C,

intermediate at 20°C and 30°C, and lowest and latest at 15°C and 10°C. This research indicated that cool temperatures result in slower development of clubroot symptoms in *Brassica* crops.

Integrated management methods such as the use of crop rotation, chemical application and mineral nutrition can only partially control the disease. Breeding of resistant cultivars is one of the most effective strategies to reduce crop loss from infection with this pathogen, in combination with other control measures (Saito et al., 2006). Host resistance is a relatively low cost and reliable approach to produce a more robust management strategy. However, a recent study showed that resistance in some resistant cultivars was eroded after 3 to 5 cycles of infection by a single-spore isolate of *P. brassicae* (LeBoldus et al., 2012). For this reason, it is beneficial if there are a wide variety of genes that can resist different virulent races of the pathogen.

2.3 Clubroot resistance in *Brassica* species

2.3.1 Sources of clubroot resistance

Substantial research on the genetics of clubroot resistance (CR) has been done for the economically important *Brassica* species, mainly *B. rapa*, *B. oleracea* and *B. napus*. Clubroot resistance to one or more pathotypes of *P. brassicae* has been reported in all major *Brassica* crops including *B. nigra*, with the exception of *B. juncea* and *B. carinata* (Diederichsen et al., 2009). Both qualitative and quantitative types of resistance have been reported, with most sources of resistance being pathotype specific. Since resistant materials against *P. brassicae* have been identified in the ancestral diploid *Brassica* species, this can expand the genetic base of CR in *B. napus* and generate sources of CR in *B. carinata* and *B. juncea* (reviewed by Hwang et al., 2011).

European fodder turnip (*B. rapa* ssp. *rapifera*) has been used as a source of CR for introgression into Chinese cabbage. Cultivars such as ‘Gelria R’, ‘Siloga’, ‘Debra’ and ‘Milan

White' are used as resistant sources for the development of CR cultivars in Chinese cabbage and *B. napus* (reviewed by Piao et al., 2009). Earlier studies have shown that at least three dominant, pathotype-specific resistance genes are present in *Brassica* A genome or turnip genotypes. Marker trait analysis also confirmed that at least eight CR loci are present in the A-genome of *B. rapa* (Sakamoto et al., 2008).

Clubroot resistance is more common in *B. rapa* compared to *B. oleracea.* Crisp et al. (1989) assessed around 1000 *B. oleracea* accessions and resistant sources were confirmed in some northern and western European kale and cabbage lines. Some of these resistant sources are commonly used in breeding for CR in *B. oleracea*. Pathotype-specific resistance to *P. brassicae* has been reported in some accessions of *B. oleracea* (Voorrips, 1996), with the majority of studies on the C-genome indicating that CR in *B. oleracea* is quantitative and is under polygenic control by one or two major QTLs as well as QTLs with minor effects. However, some reports have shown that CR in *B. oleracea* is qualitative and is under control by dominant or recessive genes (reviewed by Hwang et al., 2011).

Studies of CR in swedes (*B. napus*) indicated that cultivars such as 'Wilhelmsburger' and 'Studsgaard Bangholm' are resistant sources used to breed CR cultivars of *B. napus*. Clubroot resistance is also found in European winter *B. napus* cultivars 'Mendel' and 'Tosca', in which resistance was introgressed from the parental diploid species (Dixon et al., 1972; Frauen, 1999). Most studies revealed oligogenic control of resistance to *P. brassicae*, with proposed models based on three, four and five resistances genes (Gustafsson and Falt, 1986). Another study by Diederichsen and Sacristan (1996) suggested that resistance in resynthesized *B. napus* is under control by at least two dominant and unlinked genes. More than 20 QTLs for CR were identified

and that the genes for resistance in *B. napus* are possibly located in the A-genome from *B. rapa* (Werner et al., 2008; Chiang et al., 1977).

Earlier studies from Canadian researchers reported sources of CR that are effective against pathotypes found in central and eastern parts of the country. For example, the evaluation of 334 *Brassica* lines collected in eastern Canada showed several resistant cabbage lines against pathotypes 2 and 6 (Chiang and Crete, 1972). Another study by Vigier et al. (1989) found that the resistant *B. napus* canola cultivar 'OAC Triton' is resistant against pathotype 2, as it carries a CR gene from *B. rapa.* In recent years, research to observe a broad range of germplasm has been directed towards pathotypes on canola in western Canada (Hasan et al., 2012; Peng et al., 2014). Hasan et al. (2012) assessed 275 accessions from all six *Brassica* species and found several resistant accessions against Canadian pathotypes. Turnip (*B. rapa* var. *rapifera*), black mustard (*B. nigra*) and rutabagas (*B. napus* var. *napobrassica*) were reported to be good sources of resistance to pathotype 3. The turnip accessions also exhibited resistance to pathotypes 2, 5, 6 and 8, with the rutabagas often showing resistance to several pathotypes as well. Highly resistant sources were fewer in *B. oleracea*, with two cultivars, 'Badger Shipper' and 'Bindsachsener', carrying resistance to some pathotypes in Canada (reviewed by Rahman et al., 2014). Peng et al. (2014) also evaluated 955 *Brassica* accessions against all five pathotypes of *P. brassicae* in Canada and found a broad range of CR candidates from the diploid species *B. rapa, B. nigra* and *B. oleracea*. However, no research on inheritance of CR in *B. nigra* was reported.

2.3.2 Genetic mapping of clubroot resistance genes

Mapping of clubroot resistance genes and developing molecular markers linked to CR loci in *B. rapa, B. oleracea,* and *B. napus* have been performed. At least nine CR loci in the A-genome of *B. rapa* have been identified and are distributed to five different chromosomes. The

genes are designated as *Crr1, Crr2, Crr3, Crr4,* and *CRc* which were mapped to chromosomes A08, A01, A03, A06, and A02, respectively. *CRa, CRb* and *CRk* were also mapped on linkage group A03 but are located in a different chromosome region, with the exception of *CRk* and *Crr3* (reviewed by Piao et al., 2009). Recently, Chu et al. (2014) mapped a CR gene in *B. rapa* ssp. *chinensis* (pak choy) using 318 SSR markers in a mapping population consisting of 1,587 backcross plants with resistance to pathotype 3. This gene was designated *Rcr1* and genetically mapped to linkage group A03. Over 20 QTLs have been identified in the C-genome of *B. oleracea,* confirming that clubroot resistance is controlled at several QTLs for this *Brassica* species (Rocherieux et al., 2004; Nagaoka et al., 2010). So far, at least 20 QTLs involving clubroot resistance have been identified in the AC genome of *B. napus*. A major mapped gene, known as *Pb-Bn1*, conferring resistance to two *P. brassicae* isolates has been mapped to chromosome A03; also mapped was a minor QTL for each isolate on linkage groups C2 and C9 (Manzanares-Dauleux et al., 2000). Later, Werner et al. (2008) analyzed a DH population (*B. napus*) produced from a cross between '263/11' and a susceptible cultivar 'Express' using seven *P. brassicae* isolates. There were 19 QTLs involving resistance that were race-specific and were mapped on eight chromosomes: A02, A03, A08, A09, C3, C5, C6 and C9. From these 19 QTLs identified, four were closely linked to each other on chromosome A03, whereas three were linked to chromosome A08. QTLs on linkage groups A03 and C9 have a more significant effect and thus result in broad-spectrum resistance (reviewed by Piao et al., 2009).

2.4 B-genome of *Brassica nigra*

Mapping and identification of clubroot resistance genes have been carried out in three *Brassica* species (*B. oleracea, B. rapa*, and *B. napus*), but not in the *Brassica* crop species containing the B-genome (*B. nigra, B. carinata* and *B. juncea*). In the 1950s, *B. juncea* replaced

B. nigra as the preferred mustard crop in Asia, so not much research in *B. nigra* was done regarding breeding and genetics, unlike the other two diploid species of U's triangle (Chauhan et al., 2011). However, the B-genome is considered to be an essential source of useful genes in *Brassica* breeding, including biotic and abiotic stress tolerance, disease resistance, and oil seed quality (Chevre et al., 1997; Struss et al., 1991). Transferring CR from *B. nigra* to *B. napus* is a possible approach to expand the genetic base in oilseed germplasm (Gerdemann-Knörck et al., 1995). It is known that both of the mustard species, *B. carinata* and *B. juncea* (both containing B-genome), have potential to produce canola oil qualities (Rakow, 2007). Though there has been considerable evidence of recombination between the A and C *Brassica* genomes (Leflon et al., 2006), the B-genome seems exceptional since recombination between the B-genome and its related A/C homologues has been very limited. The B-genome species parted from the A/C lineage about 6 million years ago, but all three diploid *Brassica* genomes revealed significant conservation of gene content and sequence identity; block homeology between A- and B-genomes of *B. juncea* was also shown by Paritosh et al. (2014). However, a number of large-scale chromosomal rearrangements such as inversions set the B-genome apart, causing significant structural difference; this occurrence could be the cause for the lack of recombination between the B-genome and the closely related A/C genomes (Navabi et al., 2013). Many attempts have been made to transfer beneficial traits from *Brassica* B-genome to the corresponding A and C genomes, with little success due to limited or no recombination occurring. Therefore, a better understanding of the genome structure and gene composition of the *Brassica* B-genome could provide insights about its association with other *Brassica* species and increase the utilization of this particular genome to improve *Brassica* crops (Navabi et al., 2013).

2.5 Utilization of molecular markers and genetic mapping approaches

With the advancement of PCR, many DNA marker technologies have been developed to generate high-density genetic maps for the major *Brassica* species as well as analysis of genetic relationships between them. The majority of work on genetic mapping has utilized molecular markers in segregating populations to search for a specific trait of interest, such as disease resistance. Conventional mapping approaches for qualitative traits have made an impact in oilseed *Brassica* breeding by incorporating MAS strategy to facilitate map-based gene cloning.

Marker-assisted applications in *Brassica* began to play a key role with the first development of restriction fragment length polymorphism (RFLP) linkage maps for *B. rapa*, *B. oleracea*, and *B. napus.* The commonly used polymerase chain reaction (PCR) method greatly improved the marker density in genetic maps through the use of markers such as randomly amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP) and inter-simple sequence repeat (ISSR). PCR markers tightly linked to a specific gene can be converted into sequence characterized amplified region (SCAR) or sequence tagged site (STS) markers for commercial oilseed breeders (reviewed by Snowdon and Friedt, 2004). Simple sequence repeat (SSR) markers, also known as microsatellites, have been useful for genetic mapping as they are highly polymorphic, inexpensive and robust in nature (Grist et al., 1993). Mapping of specific genes can be carried out with numerous single-nucleotide polymorphism (SNP) markers. SNP refers to a single base change in a DNA sequence and is the most common type of DNA polymorphism in most organisms. Due to their use in high-throughput analysis, low assay cost, abundance in the genome (every 44-75 bp), locus specificity, co-dominant inheritance and low rate of genotyping errors, SNPs are highly preferred in genetic mapping and have largely replaced other molecular markers (Schlötterer, 2004). Unlike fragment length polymorphisms,

SNPs are unable to be analyzed by traditional gel electrophoresis and detection has been possible with the rise in PCR technologies (reviewed by Snowdon and Friedt, 2004). One particular method for SNP genotyping is by using kompetitive allele specific PCR (KASP). KASP is a fluorescence-based genotyping technique that is allele-specific and is based on fluorescence resonance energy transfer (FRET) for signal production (reviewed by Semagn et al., 2014).

A widely used marker-intensive application for genetic mapping is bulked segregant analysis (BSA) (Michelmore et al., 1991). BSA is a low cost and time-efficient way to identify markers linked to particular traits of interest that does not require a linkage map or the need for a large number of samples in a population. This method can be utilized in extreme mapping, implying that the plants at the extreme end of the phenotype range in a population are pooled, and the differences in genotype can be related to the phenotype of interest. With this approach, QTLs in large populations can be easily identified if the number of available markers is abundant and extensively distributed along the genome (reviewed by Deschamps et al., 2012).

Genetic mapping based on marker analysis has been carried out using high-throughput sequencers. The Sanger method of sequencing is a benchmark for producing high quality sequencing data and has been useful in the completion of whole genome sequencing drafts; in addition, genetic variations such as SNPs can be discovered. Refinements have been made with the development of automated sequencers and fluorescence labelling technology (reviewed by Deschamps et al., 2012). Sanger capillary sequencing systems include the commonly used Applied Biosystems 3xxx series or the GE Healthcare MegaBACE instrument. These systems are based on the same principles from the Sanger method (Sanger et al., 1978), with further enhancement. Other instruments also employ pyrosequencing, reversible terminator chemistry, sequencing-by-ligation and virtual terminator chemistry (Kircher and Kelso, 2010). However,

this conventional mapping method is being replaced by the more efficient mapping-by-sequencing (MBS) approach through next-generation sequencing (NGS) technologies.

The NGS technology relies on massively parallel sequencing that is capable of producing several hundreds of billions of bases per run (Shendure and Ji, 2008). Genotyping-by-sequencing (GBS) enables genome-wide detection and scoring of hundreds of thousands of markers simultaneously, derived from sequence data of the population being genotyped. This GBS procedure starts with creating NGS libraries for sequencing, SNP identification and then genotyping analysis to generate genetic maps for identifying genes of interest from segregating populations. This MBS approach is efficient because no prior sequence data is required and newly identified markers are from the population being genotyped (reviewed by Deschamps et al., 2012). A study by Liu et al. (2012) reported that bulked segregant RNA-Seq (BSR-Seq) is an effective method for gene mapping because polymorphic markers do not have to be previously identified and gene expression data can be produced at no extra cost. This approach can also be mined for DNA sequence polymorphism such as the discovery of SNPs, which are the markers of choice. This is due to their highly polymorphic nature and their capability of genotyping thousands of markers rapidly on automated platforms compared to other PCR-based markers (Raman et al., 2014). A number of major CR genes and more than 20 QTLs have been identified in *Brassica* species and close relatives; however, no robust SNP markers for MAS are available.

2.6 Transcriptomic analysis following infection by *Plasmodiophora brassicae*

Few studies have been done on the molecular mechanisms linked to clubroot resistance genes in *Brassica* species based on transcriptome analysis. The model plant, *Arabidopsis thaliana*, is a useful model host for *P. brassicae* to examine host metabolism, as it has been

reported that gene sequence identity between *Brassica* and *Arabidopsis* ranges from 75-90% (Piao et al., 2009).

Since clubroot is described by the growth of clubs on plant roots that results from cell enlargement and cell proliferation, the plant hormones auxin and cytokinin may play a role during infection (Ludwig-Müller et al., 2009; Piao et al., 2009). Genes involved in indole-3-acetic acid (IAA) biosynthesis were also shown to play a role in clubroot symptom development during later disease stages (Piao et al., 2009). Siemens et al. (2006) observed the differential gene expression between healthy root tissue and infected root tissue of *Arabidopsis* at two different time points. At 10 days post-inoculation (dpi), only genes involved in starch and sulphur metabolism were differentially regulated; however, at 23 dpi, several metabolites such as starch, lipids, and flavonoids were up-regulated.

Host metabolism in *Arabidopsis* was shown to be affected by *P. brassicae* infection based on microarray analysis, with up-regulation of genes involved in carbohydrate metabolism in root samples of the susceptible plant (Siemens et al., 2006); however, the moderately resistant plants indicated down-regulation of genes involved in pathogenesis and cell proliferation in the host (Jubault et al., 2013). Healthy and infected *Arabidopsis* plants were compared and 20 differentially regulated proteins were identified; these proteins were involved in lignin biosynthesis, cytokinin metabolism, glycolysis, intracellular calcium homeostasis and detoxification of reactive oxygen species (Devos et al., 2006). There were down-regulation of genes involved in trehalose and arginine metabolism in moderately resistant *Arabidopsis* plants, compared to the susceptible plants (Jubault et al., 2008).

Flavonoid was shown to induce clubroot symptoms in *Arabidopsis*, while inhibition of oxoglutaric acid-dependent dioxygenases reduced symptom development (Pasold and Ludwig-Müller, 2013). Genes involved in defense-related pathways can be triggered with salicylic acid (SA) or biofungicide application, which can reduce clubroot development in *A. thaliana* and *B. napus* (Lahlali et al., 2013; Lovelock et al., 2013). Additionally, Chen et al. (2016) indicated that up-regulation of PR genes involved in SA signaling pathway was higher in clubroot resistant *B. rapa* lines suggesting that SA signalling played a role in clubroot resistance. Biosynthetic pathways such as plant hormone biosynthesis can differ between host plants depending on the developmental stage, environmental conditions, type of species and individual tissues (Ludwig-Müller, 2009).

Based on the transcriptomic analysis during infection by *P. brassicae* in *Brassica rapa* (Chu et al., 2014), the gene ontology (GO) terms "defence response" and "plasma membrane" made up the majority of up-regulated DEGs in resistant plants; this is because the plasma membrane is a site where most plant-pathogen interaction takes place. The majority of GO terms regarding down-regulation of DEGs in resistant plants were linked to growth and development. This includes cell cycle, cell growth, and morphogenesis. Genes involved in uncontrolled cell division and expansions induced by clubroot were significantly supressed in resistant *B. rapa* lines (Chen et al., 2016). Results from this study indicated that many cellular activities in inoculated resistant plants were greatly up-regulated at 15 dpi due to root infection and pathogen colonization occurring around this time point. Lipid compounds were found to play an important role in clubroot resistance, as lipids can function in detecting infection and possibly in regulating gene transcription (Walley et al., 2013). Jasmonates were also observed to have a role in clubroot resistance, as they have inhibitory impacts on many plant pathogens (Prost et al., 2005). Genes

involved in ethylene metabolic and biosynthetic processes were up-regulated in resistant plants, suggesting that ethylene and jasmonic acid may function together in plant defence, but SA may not have a critical role in clubroot resistance (Chu et al., 2014). Up-regulation of DEGs involved in callose localization and deposition was observed in resistant plants, and a study by Naumann et al. (2013) indicated that callose plays a role in plant cell wall defence against fungal pathogens.

3. MATERIALS AND METHODS

3.1 *Brassica nigra* parental lines and segregating populations

Clubroot resistant accessions identified at Saskatoon Research Centre-Agriculture and Agri-Food Canada (SRC-AAFC) (Peng et al., 2014) were used for this study. Seeds of the resistant lines BRA192/78 and PI 219576 were kindly provided from Viterra, Viterra Inc., Research and Development (201-407 Downey Road, Saskatoon, Saskatchewan S7N 4L8, Canada). The susceptible line CR2748 was obtained from IPK, Leibniz Institute of Plant Genetics and Crop Plant Research (OT Gatersleben, Corrensstrasse 3, D-06466 Seeland, Germany). The two clubroot resistant lines (PI 219576 and BRA192/78) were crossed with the susceptible female line CR2748 to produce F_1 populations. A total of 192 F_1 plants derived from CR2748 × BRA192/78 and CR2748 × PI 219576 were inoculated with pathotype 3 of *P. brassicae* using an established protocol (Peng et al., 2011b). Self-pollination of F_1 generation plants produced F_2 populations. Resistant F_1 plants were backcrossed with susceptible CR2748 plants to produce BC_1 populations. A total of 57 F_2 plants and 234 BC_1 plants with the CR gene from PI 219576 were inoculated with pathotype 3. A total of 90 F_2 plants and 242 BC_1 plants from BRA192/78 were inoculated with pathotype 3. A total of 41 ACDC (a *B. rapa* canola line highly susceptible to clubroot) and 95 CR2748 plants were used in all inoculated trials.

BC_1 populations were made by crossing the F_1 plants derived from PI 219576 and BRA192/78 (male resistant plants) with CR2748 (female susceptible plants). Five well grown buds per female plant were kept for crossing, while the remaining flowers and small buds were removed. When the retained buds had opened, their anthers were detached carefully with forceps. Anthers from newly opened flowers of donor plants were collected and pollen grains were dusted

directly to stigmas of the pistils of the female plants. Pollinated plants were covered with a plastic crossing bag for 5 days.

3.2 Plant growth conditions

Seeds were planted in Sunshine #3 soil-less planting mix (SunGro Horticulture, Vancouver, BC) in tall, narrow plastic containers (5-cm diameter, 20-cm height, Steuwe & Sons, Corvalis, OR). The soil mix was enhanced with 1% (w/v), 16-8-12 (N:P:K) control-released fertilizer. Plants were kept in a greenhouse (22/18°C, day/night) with a 14-h photoperiod (230 µmol/m^2/s at the canopy level). After inoculation with *P. brassicae*, plants were kept in a growth room at 23/20°C and 14-h photoperiod (512 µmol/m^2/s). The reason for utilizing two different facilities for plant growth is to avoid spread of clubroot contamination.

3.3 Inoculum preparation and inoculation

Inoculum was prepared as a resting-spore suspension following the procedures described by Lahlali et al. (2013). Fresh and mature clubroot galls consisting of pathogen resting spores were buried in a tray of Sunshine planting mix at room temperature for around 3-4 weeks to mature the resting spores. The dried galls were then transferred to plastic bags and kept at -20°C until use. Approximately 3-5 grams of clubbed canola roots were soaked in distilled water at room temperature for 1-2 hrs to soften the tissue, and then the clubbed roots were sliced into smaller pieces with scissors. The root materials were soaked in 50-100 mL de-ionized water in a Waring blender at high speed for 2 min. The mixed solution was filtered through 2-3 layers of nylon mesh cloth. About 1 mL of the original suspension was pipetted and added into 99-mL water. After mixing, this diluted suspension was used to estimate spore concentration with a hemacytometer. The inoculum concentration was adjusted to $1x10^7$ spores/mL. For inoculation, 5 mL of resting-spore suspension was pipetted directly on the seed in each container immediately

after planting. Soil was moisturized with water before planting to facilitate infection and water added daily to the bottom of trays.

3.4 Tissue collection and phenotyping of clubroot symptoms

Leaf tissues of BC_1 plants were collected at 2 weeks after inoculation for RNA extraction (stored at -80°C) and collected at 4 weeks after inoculation for DNA extraction (stored at -20°C). Phenotyping was carried out and results evaluated 5 weeks after inoculation using a 0-3 scale (Figure 3.1). A rating of '0' was considered resistant and '1-3' susceptible. A '0' rating indicated no galling; a '1' rating indicated a few small galls; a '2' rating indicated moderate galling; and a '3' rating indicated severe galling (Strelkov et al., 2006). The F_2 populations were analyzed to confirm the expected genetic analysis ratio of 3R:1S and the BC_1 populations (expected 1R:1S ratio) were used for genetic mapping work. The goodness of fit for the segregation was analyzed using a Chi-square (X^2) Test (Sokal and Rohlf, 1995).

Figure 3.1: Symptom grade of clubroot disease caused by *Plasmodiophora brassicae* in *Brassica nigra*. Severity of root infection was assessed on a 0-3 scale, where 0 represents no galling; 1, a few small galls; 2, moderate galling; and 3, severe galling.

3.5 DNA sample preparation

3.5.1 DNA extraction protocol

The following reagents were used to make 110 mL 2X CTAB buffer: 2.2 g CTAB powder (2%), 5.5 mL 2 M Tris-HCl pH 8 (100 mM), 4.4 mL 0.5 M EDTA (20 mM), 30.8 mL 5 M NaCl (1.4 M) and 67.1 mL ddH2O. The following reagents were used in the DNA extraction protocol: chloroform, isopropanol, sterile milli-Q water, RNase A (10 mg/mL) and mercaptoethanol.

The following are the steps to extract DNA from *B. nigra* leaves: 1) Leaf tissues were freeze-dried for 2 days; 2) dried leaf tissues with beads were put in shaker for 5-10 min until fine powder; 3) for each sample in centrifuge tubes, 500 µL of pre-warmed (50°C) CTAB extraction buffer including 2% (v/v) 2-mercaptoethanol was added; 4) tubes were incubated at 65-70°C for at least 1 hr; 5) equal volume (500 µL) of chloroform: isoamyl alcohol (24:1) was added and gently inverted for 10 min; 6) tubes were centrifuged at 10,000 rpm (revolutions per minute) for 10 min and the upper aqueous phase was transferred into new tubes; 7) a volume of 400 µL isopropanol was added to each tube and gently inverted 20 times, and then incubated at room temperature for 30 min; 8) DNA was pelleted by centrifuging tubes at 10,000 rpm for 15 min at 4°C; 9) the supernatant was decanted and the pellet was air-dried in a fume hood at room temperature for 1 hr; 10) the pellet was re-suspended in sterile Milli-Q water (200 µL) and 1 µL of RNase A (10 mg/mL) was added to each tube; 11) NanoDrop ND-2000c spectrophotometer (Thermo Scientific, Wilmington, DE) was utilized to measure DNA concentration (ng/µL) and the 260/280 ratio was analyzed to measure purity of the samples; and 12) DNA concentration was adjusted to 50 ng/µL with sterile Milli-Q water.

3.5.2 PCR conditions

The PCR recipe for SSR marker genotyping contained 0.0375 µL forward primer (5 µM), 0.375 µL reverse primer (5 µM), 0.5 µL 2 mM dNTP, 1 µL 10X standard buffer, 0.125 µL fluoro-T7 primer, 5.8625 µL Milli-Q water, 2 µL 50 ng/µL genomic DNA and 0.1 µL Taq polymerase (New England BioLabs). The PCR conditions were as follows: denaturation at 94°C for 3 min, followed by 22 cycles of 94°C for 30 s, 56°C for 30 s and 72°C for 45 s; then 22 cycles of 94°C for 30 s, 47°C for 30 s, 72°C for 45 s, final extension at 72°C for 10 min and a holding stage at 15°C. The PCR machine used was a 384-well thermocycler (Applied Biosystems, Inc.). Post-PCR steps were as follows: all four PCR reactions were combined, 2 µL was aliquoted from each reaction into a new PCR plate, then diluted with 30 µL of Milli-Q water. From this dilution, 2 µL was added to a final PCR plate with 9 µL of HI-DYE containing 600-LIZ size standard.

Genotyping of SNP markers was based on the KASP method. The PCR mixture contained 4 µL 2X KASP Master Mix (LGC Genomics), 0.14 µL assay mix and 4 µL of 50 ng/µL genomic DNA. The PCR conditions were as follows: pre-PCR holding stage at 30°C for 1 min, denaturation at 95°C for 15 min, followed by 10 cycles of 95°C for 20 s and 61°C for 1 min; then 26 cycles of 95°C for 20 s, 55°C for 1 min and a post-PCR holding stage at 30°C for 1 min. KASP runs were carried out with the StepOne Plus Real-Time PCR system (Applied Biosystems, Inc.).

3.6 Genetic mapping through bulked segregant analysis using microsatellite markers

SSR markers from *B. juncea* were developed by SRC-AAFC (http://aafc-aac.usask.ca/BrassicaMAST). The SSR markers from B-genome linkage groups were used for the rough mapping work since BSR-Seq technology was not available at the time. Bulked segregant analysis was first carried out as follows: R bulk (16 plants) and S bulk (16 plants) from

the BC_1 population of PI 219576; and two parental lines, PI 219576 and CR2748. A pre-run was conducted using extracted DNA from the two bulks and two parental lines, then subjected to PCR using 120 SSR markers that corresponded to the microsatellite loci within the *Brassica juncea* genome (B1-B8 chromosomes).

A fluorescence-based cycle sequencing system, ABI 3130xl genetic analyzer (Applied Biosystems, Inc.), was utilized for SSR marker analysis. There were four fluoro-labelled T7 primers required for analysis on the ABI 3130xl: FAM (blue), PET (red), NED (yellow) and VIC (green). The fluorescently labelled DNA fragments, separated by size, move across a laser beam, which causes the dyes on the fragments to fluoresce.

The Genemapper 5.0 software was used for SSR marker genotyping. Lastly, the genotype scores were compared with the phenotype results to see if there is association with clubroot resistance. A preliminary linkage map was generated using JoinMap 4.1 software.

3.7 Mapping-by-sequencing through BSR-Seq

Leaf tissues collected from BC_1 populations of PI 219576 and BRA192/78 were immediately placed in liquid nitrogen, and then stored in a -80°C freezer until use. Leaf tissues were ground using a mortar and pestle with liquid nitrogen and grouped into resistant and susceptible bulks, containing 30 individuals in each bulk. There were three biological replicates for R and S bulks (R1, R2, R3, S1, S2, S3), and three technical replicates from each leaf tissue collected.

3.7.1 RNA extraction

Total RNA from resistant and susceptible bulks were extracted using the RNeasy Plant Mini Kit (Qiagen; Toronto, CA) with on-column deoxyribonuclease (DNAse) digestion using a Qiagen RNase-Free DNase Set following the manufacturer's instruction. The RNA

concentrations and purity were checked using a NanoDrop ND-2000c spectrophotometer. The Experion RNA StdSens analysis kit (Bio-Rad Laboratories, Inc.) was used for RNA quality analysis with the Experion automated electrophoresis system. RNA quality was analyzed to ensure that the RNA integrity number (RIN) was greater than eight for each sample. RNA samples were stored in -80°C until they were used for cDNA synthesis.

3.7.2 Complementary DNA (cDNA) library construction

The cDNA libraries were prepared following the TruSeq RNA Sample Preparation v2 Guide (Illumina; San Diego, CA). The first step for cDNA library construction was to purify and fragment poly-A containing mRNA from total RNA. First strand cDNA synthesis was performed by reverse transcriptase and random primers. Next, second strand cDNA synthesis was carried out using DNA polymerase I and RNaseH. The cDNA fragment went through an end repair process, addition of a single 'A' base and ligation of adaptors. The purpose of adaptor ligation was to differentiate the samples on the flow cell. These products were purified and then amplified with PCR to create the final cDNA libraries. A NanoDrop ND-2000c spectrophotometer was utilized to check cDNA concentrations and purity. Quality control and qPCR analysis were carried out to validate the cDNA libraries. The Experion DNA 1K Analysis Kit (Bio-Rad Laboratories, Inc.) was used to confirm the size and purity of the cDNA libraries. The final product should be a band at approximately 260 bps. The KAPA Library Quantification Kit v4.11 was used to perform qPCR of cDNA libraries.

3.7.3 Sequencing through the MiSeq platform and data analysis

The cDNA libraries were denatured and diluted to prepare them for paired-end sequencing on the MiSeq system (*Preparing Libraries for Sequencing on the MiSeq*, Illumina 2013). MiSeq Reagent Kit v3 was used. RNA-Seq was carried out on each inoculated resistant

and susceptible bulked samples using the Illumina MiSeq platform in Dr. Markus Hecker's lab at University of Saskatchewan's toxicology building (Saskatoon, Canada). Prior to the MiSeq run, options for adaptor trimming and adapter trimming read 2 were selected in the Illumina experiment manager software, so that the sequencing adaptors can be removed from the resulting raw reads.

High-quality reads were subsequently aligned to the Chinese cabbage (*B. rapa* ssp. *pekinesis*) Chiifu genome (v1.5; http://brassicadb.org/brad) and *B. oleracea* genome v2.1 using SeqMan NGen 11 (DNASTAR, Madison, WI, USA) software with a minimum match percentage of 93%. The RNA-Seq reads were also aligned to a *B. nigra* reference genome v1 that was provided by Dr. I. Parkin at Saskatoon Research Centre, AAFC.

The A-reference genome consists of 10 chromosomes and 40,357 scaffolds; the total lengths of chromosomes and scaffolds are about 257 million bases (Mb) and 27 Mb, equivalent to about 90% and 10% of the reference genome, respectively. The B-reference genome consists of 8 chromosomes and 3,026 scaffolds; the total lengths of chromosomes and scaffolds are about 370 Mb and 111 Mb, equivalent to about 77% and 23% of the reference genome, respectively. The C-reference genome consists of 9 chromosomes and 2,620 scaffolds; the total lengths of chromosomes and scaffolds are about 447 Mb and 22 Mb, equivalent to about 95% and 5% of the reference genome, respectively. To simplify downstream data analysis, only the chromosome sequences were used in the current study.

Two methods were used for short read assembly: 1) short reads from three biological replicates of each R and S bulk were assembled into the reference genome, producing six assembly files – known as single sample assembly (SSA); 2) short reads from a pool of the three R bulks and a pool of the three S bulks were assembled into the reference genome, generating two assembly files – known as pooled sample assembly (PSA). Standard assembling and filtering

parameters were used. Pooled assembly of the three R and three S bulks was performed for SNP identification, whereas single assembly was done for gene expression analysis. SNP discovery for marker development was carried out using SeqMan Pro 11 (DNASTAR, Madison, WI, USA) software. The SNP discovery parameter was set at default, with Q call ≥ 15 and depth ≥ 5. Analysis of gene expression was done by using QSeq (DNASTAR, Madison, WI, USA) software. The heat map was generated, which uses color to display the expression levels of genes in R and S bulks; color intensity indicates the levels of gene activity. By default, hierarchical clustering is displayed in the heat map view. A scatter plot was also produced to give a visual comparison of gene expression levels between R and S bulks. Data was visualized in the scatter plot on a $\log_2$ scale, with the R^2 value between 0.0 and 1.0. The Student's t-test method was used for identifying differentially expressed genes.

3.7.4 Single nucleotide polymorphism (SNP) marker genotyping using KASP

Validation and genotyping of SNP markers were based on the KASP version 4.0 SNP Genotyping Manual, and the data were analyzed with StepOne Software v2.3 (Applied Biosystems, Inc.). Pre-screening was first done on six R and six S BC_1 samples from PI 219576 and BRA192/78 to observe if there was distinctive clustering between the two groups. Polymorphic SNP markers were then tested on BC_1 populations, with R and S parents as controls. Genotype scores were compared with the phenotype results to see if there is association with clubroot resistance. Genetic linkage maps were constructed by JoinMap 4.1 software, with distance measured in centimorgans (cM).

3.8 Annotation of differentially expressed genes (DEGs)

The DEGs for resistant samples were separated into up-regulated and down-regulated groups relative to the gene expression observed in susceptible samples. Gene annotation was obtained from the *Brassica* database (BRAD) (http://www.brassicadb.org).

3.9 Real-time reverse transcription (RT) quantitative PCR (qPCR)

The purpose of this experiment was to confirm the results of RNA-seq data. Primers were designed using PrimerQuest (Primer3 Software, version 2.2.3) and synthesized by Integrated DNA Technologies Inc. (Coralville, IA). The list of primers designed for RT-qPCR is shown in Appendix 1. The planting experiment was divided into inoculated (pathotype 3) and non-inoculated (water) trays. Nine plants were used for each of PI 219576, BRA192/78 and CR2748, in the inoculated and non-inoculated trays, with a total of 27 plants in each tray. The 9 plants for each line were divided into three groups (three plants each) in which the leaf and root tissues were bulked to produce three biological replicates for each line. Leaf and root tissues from the inoculated and non-inoculated trays were collected 2, 3, 4, 5 and 6 weeks post-inoculation (wpi). Non-inoculated plants did not develop obvious clubroot symptoms. RNA extraction was performed using the MasterPure Complete DNA & RNA Purification Kit (Mandel Scientific) following manufacturer's instruction. Verso SYBR Green 1-Step qRT-PCR Kit Plus ROX Vial (Thermo Scientific) was used to quantify RNA in a single step assay. The reaction mix preparation for a 10 µL final volume was as follows: 0.1 µL verso enzyme mix, 5 µL 2X 1-Step qPCR SYBR mix, 0.5 µL RT enhancer, 0.2 µL forward primer (5 µM), 0.2 µL reverse primer (5 µM), 3 µL nuclease-free water and 1 µL template RNA. ROX internal passive reference dye was used to normalize the fluorescent reporter signal generated in qPCR. The 1-step qRT-PCR thermal cycling conditions and melt curve program were set following the manufacturer's

instruction. The melt curve was performed to confirm the specificity of the reaction and absence of primer dimers. The experiment was performed on the StepOne Plus Real-Time PCR system (Applied Biosystems, Inc.) and relative expression data was analyzed using the StepOne software v2.3 (Applied Biosystems, Inc.). The ΔC_T (threshold cycle) values observed from RT-qPCR experiment were compared with $\log_2$ fold changes from RNA-Seq data.

4. RESULTS

4.1 Clubroot assessment of *B. nigra* line PI 219576

4.1.1 Single dominant gene confirmation

All of the PI 219576 plants were resistant to pathotype 3 of *P. brassicae*, showing no clubroot symptoms at 5 weeks after inoculation (Figure 4.1A). The ACDC and CR2748 plants showed complete susceptibility (Figure 4.1B). Complete resistance to pathotype 3 of *P. brassicae* was found in 100 F_1 generation plants derived from the cross of CR2748 with PI 219576 (Figure 4.1C), indicating that PI 219576 was likely a homozygous resistant line and the resistance was controlled by a dominant gene(s). A total of 57 F_2 plants from PI 219576 ($X^2 = 0.05$, $P = 0.82$) showed 3:1 segregation ratio for resistance and susceptibility (Figure 4.1D). Evaluation of 234 BC_1 plants from PI 219576 ($X^2 = 0.62$, $P = 0.43$) showed a 1:1 segregation ratio (Figure 4.1E), indicating that CR in PI 219576 is controlled by a single dominant gene. The Chi-Square statistics for the 3:1 and 1:1 ratios are summarized in Table 4.1. The CR gene in PI 219576 was designated as *Rcr6a*.

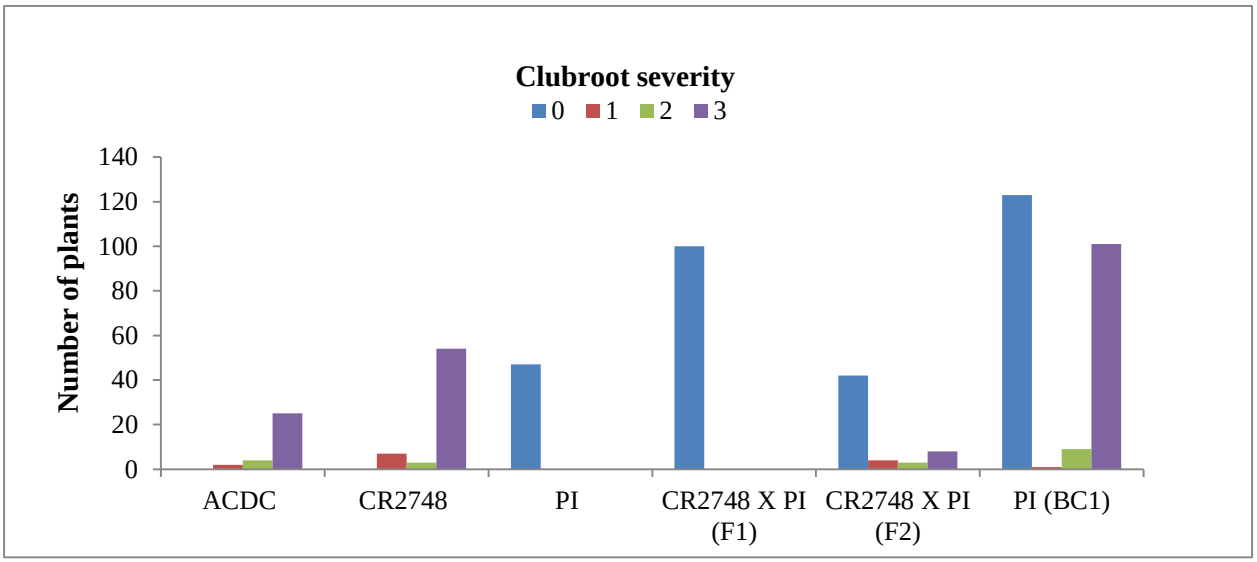

Figure 4.1: Segregation of clubroot resistance for *B. nigra* parents (PI 219576 and CR2748), F_1, F_2 and BC_1 populations. Plants were inoculated with pathotype 3 of *P. brassicae*. Resistant parent showed complete resistance. Susceptible parent (CR2748) and susceptible control (ACDC of *B. rapa*) showed clubroot symptoms. F_1 population showed complete resistance. F_2 population showed segregation ratio of 3R:1S. BC_1 population showed segregation ratio of 1R:1S.

Table 4.1: Chi-Square tests confirm anticipated BC_1 and F_2 segregation ratios in the PI 219576 population of *B. nigra*. The segregation ratio was 1R:1S for BC_1 and 3R:1S for F_2. Plants were inoculated with pathotype 3 of *P. brassicae*.

	Phenotype					
Population	Total	Resistant (R)	Susceptible (S)	Expected ratio (R:S)	χ^2	*P-value*
BC_1	234	123	111	1:1	0.62	0.43
F_2	57	42	15	3:1	0.05	0.82

4.1.2 SSR marker screening and development of a preliminary linkage map

In the SSR marker screening analysis, a pre-run was first conducted to select polymorphic markers that can be used to test on the BC_1 population from PI 219576. A total of 120 SSR markers from B-genome linkage groups B1 to B8 were initially tested on DNA samples from R parent PI 219576, S parent CR2748, R bulk and S bulk (pre-run). Sixty-six markers out of a total of 120 exhibited polymorphism in the pre-screening run; these markers were then tested on a small population of 32 BC_1 plants from PI 219576, with 42 SSR markers displaying polymorphism (Appendix 2). Selected polymorphic markers were then tested on the entire BC_1 population from PI 219576. Genemapper 5.0 software was used to analyze the fragment bands of R and S samples so that the bands from R parents were scored as '+' and the bands from S parents were scored as '−'. A scoring of '0' was assigned to ambiguous data. DNA fragments from the BC_1 population could be separated into 2 bands (1 and 2) by some markers. Band 1 was from PI 219576 carrying the CR gene and band 2 from CR2748 (Figure 4.2).

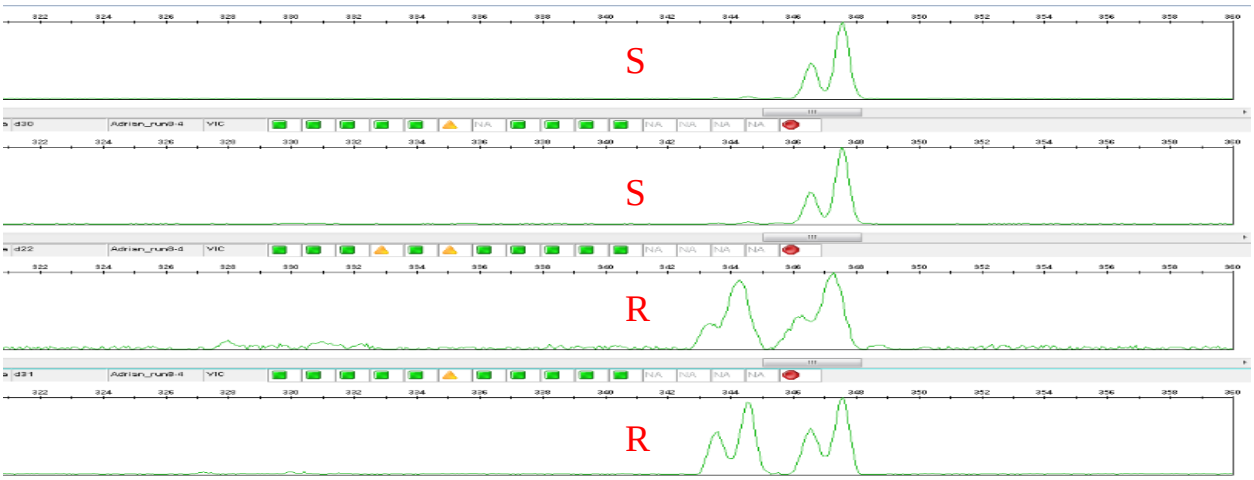

Figure 4.2: Genotype plot from Genemapper 5.0 software showing two typical examples for each of the R and S samples of BC_1 population from PI 219576. The R samples were heterozygous for resistance containing two peaks, whereas the S samples were homozygous recessive containing only one peak. DNA fragments were separated into 2 bands (1 and 2) by some markers. Band 1 was from PI 219576 carrying the CR gene and band 2 from CR2748.

A genetic linkage map was constructed as shown in Figure 4.3. Only 4 markers (sJ8211a, sB3868a, sJ4216a and sJ7148a) displayed association with resistance. The markers sJ8211a, sB3868a and sJ4216a were mapped on B3. However, the last marker sJ7148a was not mapped in the *B. juncea* mapping population (http://aafc-aac.usask.ca/BrassicaMAST). Based on this research, sJ7148a is likely a marker on B3 since it was linked to the three markers on B3. The CR gene *Rcr6a* was roughly mapped to a wide range of 89.1 cM flanked by the markers sJ4216a and sJ7148a. There was a high rate of recombination with a total of 93 recombinants identified via comparison of flanking markers and phenotype data over 234 BC_1 plants, with 61 falling between sJ4216a and *Rcr6a* and 32 between *Rcr6a* and sJ7148a. This conventional approach for genetic mapping was inefficient and there were not sufficient markers identified to narrow down the location of the CR gene. For this reason, further genetic mapping through the MBS approach was performed.

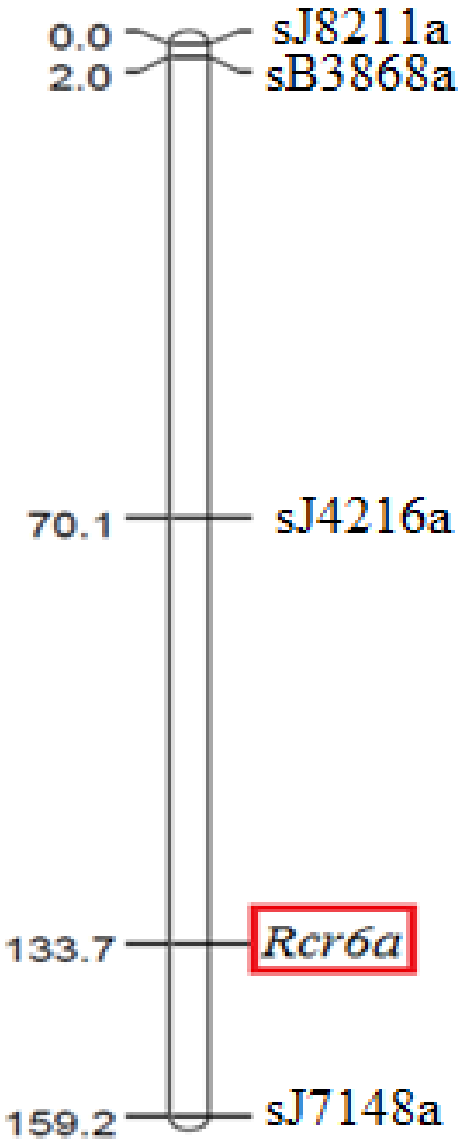

Figure 4.3: Preliminary linkage map on chromosome B3 of *B. nigra* developed in a segregating BC_1 population (234 plants) with the CR gene *Rcr6a* from PI 219576. *Rcr6a* was roughly mapped in between SSR markers sJ4216a and sJ7148a. Plants were inoculated with pathotype 3 of *P. brassicae.* The genetic distance (cM) is shown on the left.

4.1.3 Mapping of *Rcr6a* with SNP markers via BSR-Seq using reference genomes of *B. rapa* and *B. oleracea*

Our group has developed a new mapping-by-sequencing (MBS) method based on the distribution of polymorphic variants (F. Yu, unpublished data). To further define genomic location of *Rcr6* gene(s), polymorphic variants that could be utilized as genetic markers were identified in the mapping populations using BSR-Seq data. The read length of the MiSeq data was 76 bases, and the total amount of data was 8.6 Gb for PI 219575. The short reads from RNA-Seq were aligned to *B. rapa* and *B. oleracea* reference genomes, respectively, since the genome for *B. nigra* has not been published.

4.1.3.1 RNA-Seq and sequence alignment

Pooled sample assembly of the three R bulks and three S bulks from PI 219576 was carried out to identify variants (SNPs and InDels). A total of 32.0 million (M) sequences, 2062.2 Mb in length, with 8-fold coverage of the reference A-genome were assembled into *B. rapa* chromosomes from the pool of three R bulks, and 39.5 M sequences, 2523.4 Mb in length, with 10-fold coverage were assembled from the pool of three S bulks (Table 4.2). More sequences were aligned into the longer chromosomes A03 and A09 and fewer sequences were aligned into the shorter chromosomes A04 and A10 (Table 4.2).

A total of 33.0 M sequences, 2097.6 Mb in length, with 5-fold coverage of the reference C-genome were assembled into *B. oleracea* chromosomes from the pool of three R bulks, and 39.0 M sequences, 2490.4 Mb in length, with 6-fold coverage were assembled from the pool of three S bulks (Table 4.3). More sequences were aligned into chromosomes C3 and C4, but C9 is

a longer chromosome than C4; fewer sequences were aligned into chromosomes C6 and C7, but C8 is a shorter chromosome than C7 (Table 4.3).

Table 4.2: Short reads from the resistant (R) and susceptible (S) bulks of PI 219576 were assembled into chromosomes of the *B. rapa* reference genome in pooled sample assembly (PSA).

Chromosome number	Chromosome size (bases x 10^6)	Number of sequences (x 10^6)		Accumulated length of sequences (bases x 10^6)	
		R	S	R	S
A01	26.8	3.0	3.7	191.4	233.8
A02	27.0	2.8	3.4	176.4	215.3
A03	31.8	4.4	5.4	284.8	342.2
A04	19.3	2.6	3.3	170.4	210.7
A05	25.3	3.0	3.6	192.5	230.7
A06	25.2	3.4	4.4	219.2	285.0
A07	25.9	3.3	4.0	211.6	255.5
A08	20.8	3.0	3.6	192.9	234.6
A09	38.9	4.5	5.6	292.2	356.2
A10	16.4	2.0	2.5	130.8	159.4
Total	257.4	32.0	39.5	2062.2	2523.4

Table 4.3: Short reads from the resistant (R) and susceptible (S) bulks of PI 219576 were assembled into chromosomes of the *B. oleracea* reference genome in pooled sample assembly (PSA).

Chromosome number	Chromosome size (bases x 10^6)	Number of sequences (x 10^6)		Accumulated length of sequences (bases x 10^6)	
		R	S	R	S
C1	43.8	3.2	3.8	202.0	241.6
C2	52.9	3.1	3.7	194.3	233.1
C3	65.0	5.5	6.5	351.3	415.2
C4	53.8	4.3	5.0	271.0	316.2
C5	46.9	3.8	4.4	240.0	281.4
C6	39.8	2.7	3.3	176.2	211.0
C7	48.4	2.9	3.4	184.6	220.2
C8	41.8	3.7	4.4	235.2	280.6
C9	54.7	3.8	4.5	243.0	291.1
Total	447.1	33.0	39.0	2097.6	2490.4

The B-genome reference was also verified. A total of 41.1 M sequences, 2841.3 Mb in length, with 8-fold coverage of the reference B-genome were assembled into *B. nigra* chromosomes from the pool of three R bulks, and 47.9 M sequences, 3332.9 Mb in length, with 9-fold coverage were assembled from the pool of three S bulks (Table 4.4). More sequences were aligned into chromosomes B3 and B6, than in the longer chromosomes B2 and B8; fewer sequences were aligned into chromosomes B4 and B7 than in the shortest chromosome B6 (Table 4.4).

Table 4.4: Short reads from the resistant (R) and susceptible (S) bulks of PI 219576 were assembled into chromosomes of the *B. nigra* reference genome in pooled sample assembly (PSA).

Chromosome number	Chromosome size (bases x 10^6)	Number of sequences (x 10^6)		Accumulated length of sequences (bases x 10^6)	
		R	S	R	S
B1	42.3	5.0	5.7	345.6	395.1
B2	52.7	5.3	6.2	362.6	426.7
B3	46.9	5.6	6.4	385.9	449.1
B4	43.4	4.8	5.6	334.6	393.3
B5	51.8	5.0	6.0	347.5	415.9
B6	36.7	5.9	6.9	407.7	481.9
B7	41.7	4.1	4.8	281.6	331.0
B8	54.2	5.4	6.3	375.8	439.9
Total	369.7	41.1	47.9	2841.3	3332.9

4.1.3.2 Identification of variants

The numbers of variants in PI 219576 was identified using the PSA method aligned with A-, C- and B-genome references. There were a total of 120.2 thousand (K) polymorphic (poly) variants when aligned with A-genome, 303.2 K poly variants when aligned with C-genome, and 256.0 K poly variants when aligned with B-genome. Chromosome A08 had the highest percentage of variants in the A-genome at 30.6% (Figure 4.4A), C8 had the highest percentage of variants in the C-genome at 30.2% (Figure 4.4B), and B3 had the highest percentage of variants in the B-genome at 48.8% (Figure 4.4C).

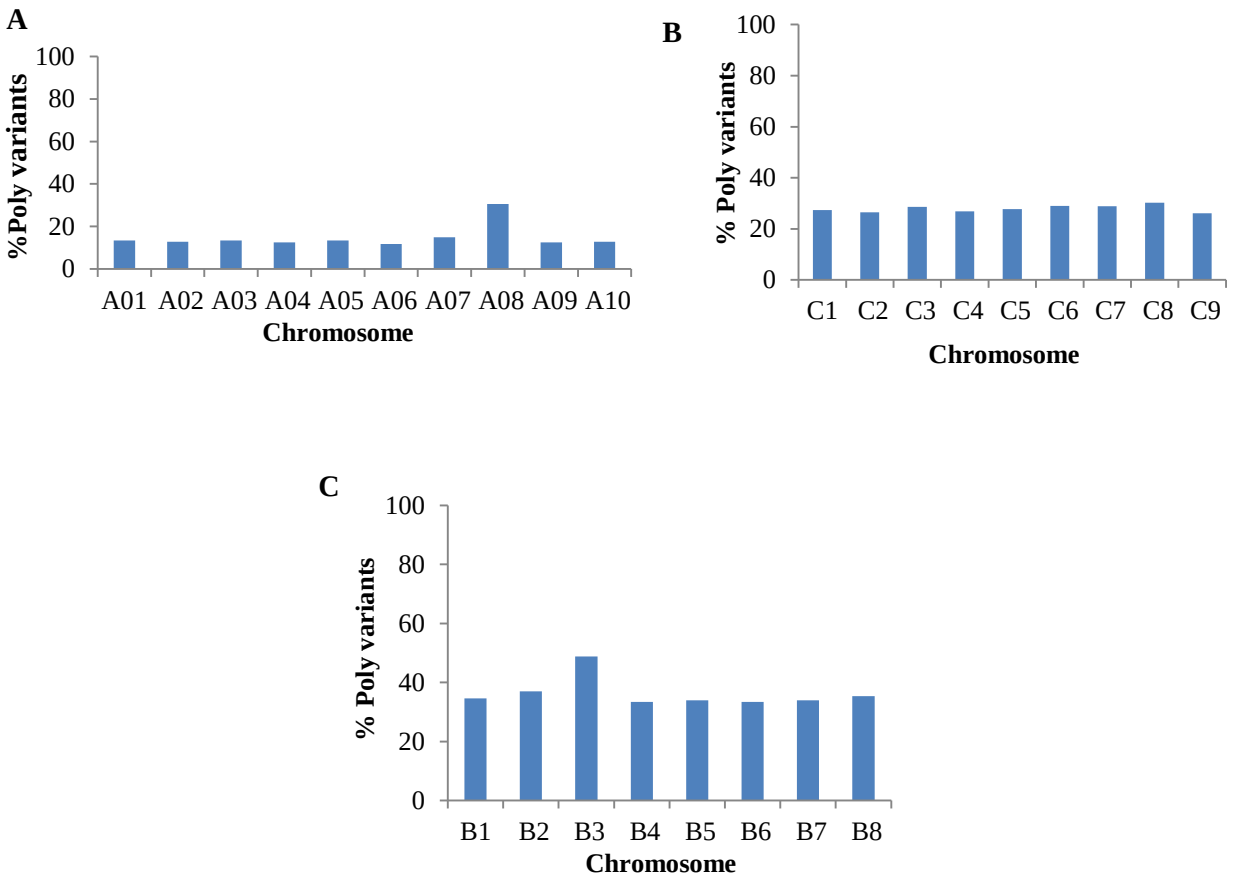

Figure 4.4: Percentage of poly variants (SNPs and InDels) from PI 219576 using various *Brassica* reference genomes. **A)** Percentage of poly variants by PSA using A-reference genome of *B. rapa*. **B)** Percentage of poly variants by PSA using C-reference genome of *B. oleracea*. **C)** Percentage of poly variants by PSA using B-reference genome of *B. nigra*.

The poly variant (SNPs and InDels) distribution (%) was analyzed using the A- and C-reference genomes. The B-genome reference is not yet complete, so the distribution results did not determine which chromosome contained the highest percentage of poly variants along the physical distance. For PI 219576, chromosome A08 contained the highest percentage of poly variants located in the physical range from approximately 10 Mb to 20 Mb (Figure 4.5); and C8 contained the highest percentage of poly variants located in the physical range from approximately 18 Mb to 24 Mb (Figure 4.6). Variants identified in the R and S samples could be the same (monomorphic, mono) or different (polymorphic, poly) (Figure 4.7).

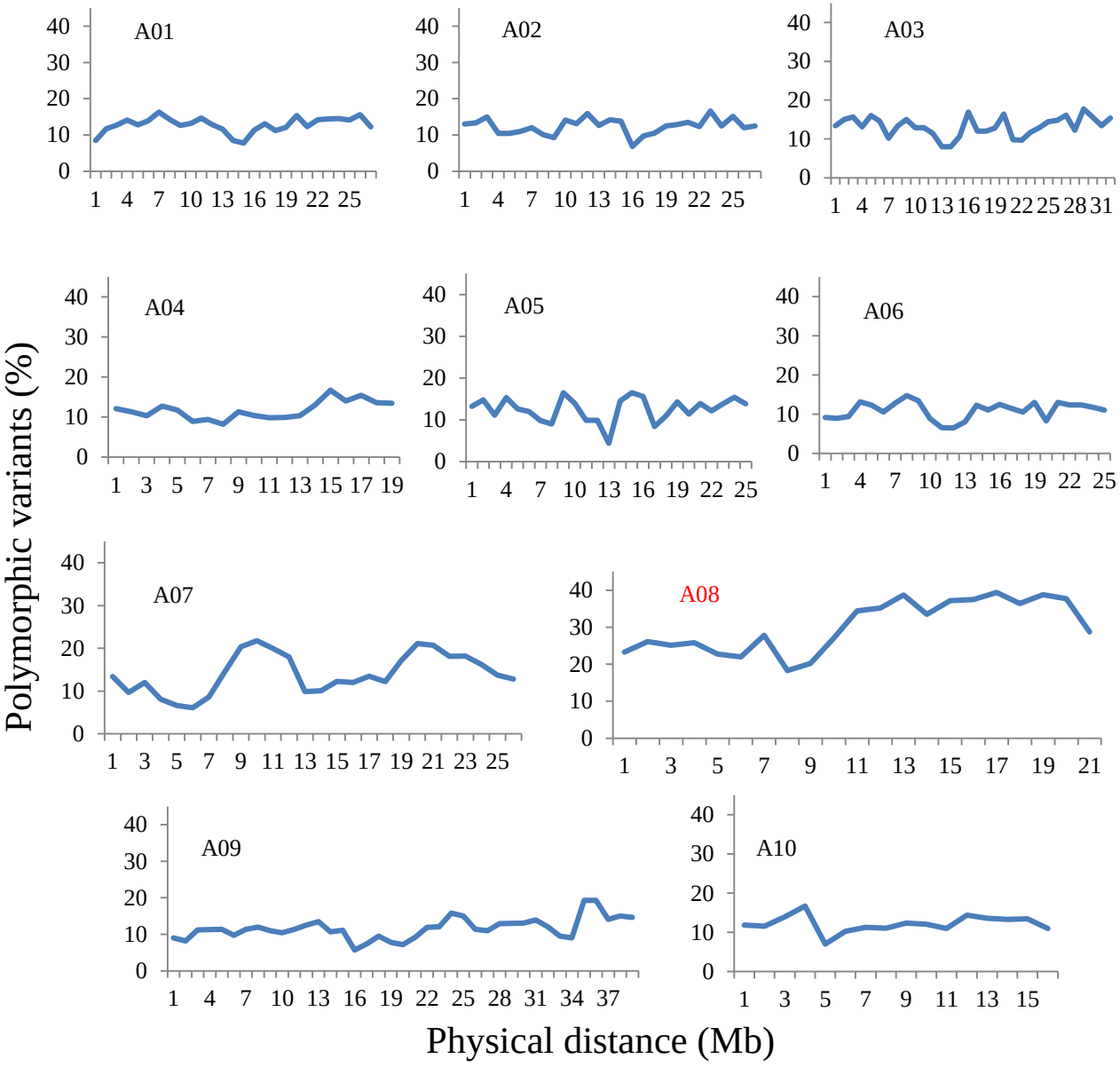

Figure 4.5: Percentage of poly variants (y-axis) distributed along the physical distance (x-axis) in PI 219576 using the A-reference genome of *B. rapa*. Chromosome A08 contained the highest percentage of poly variants at a wider physical range compared to other chromosomes.

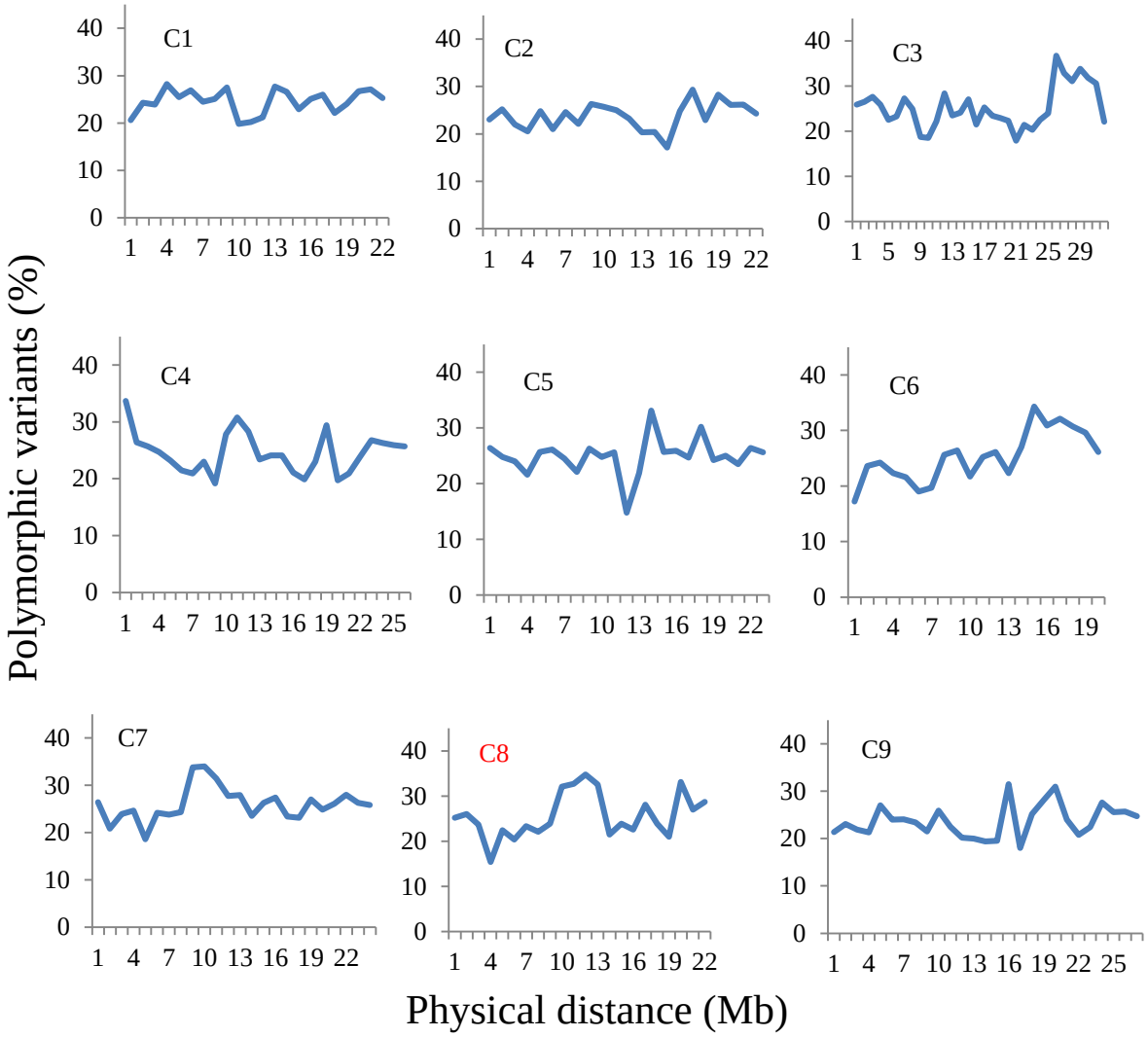

Figure 4.6: Percentage of poly variants (y-axis) distributed along the physical distance (x-axis) in PI 219576 using the C-reference genome of *B. oleracea*. Chromosome C8 contained the highest percentage of poly variants at a wider physical range compared to other chromosomes.

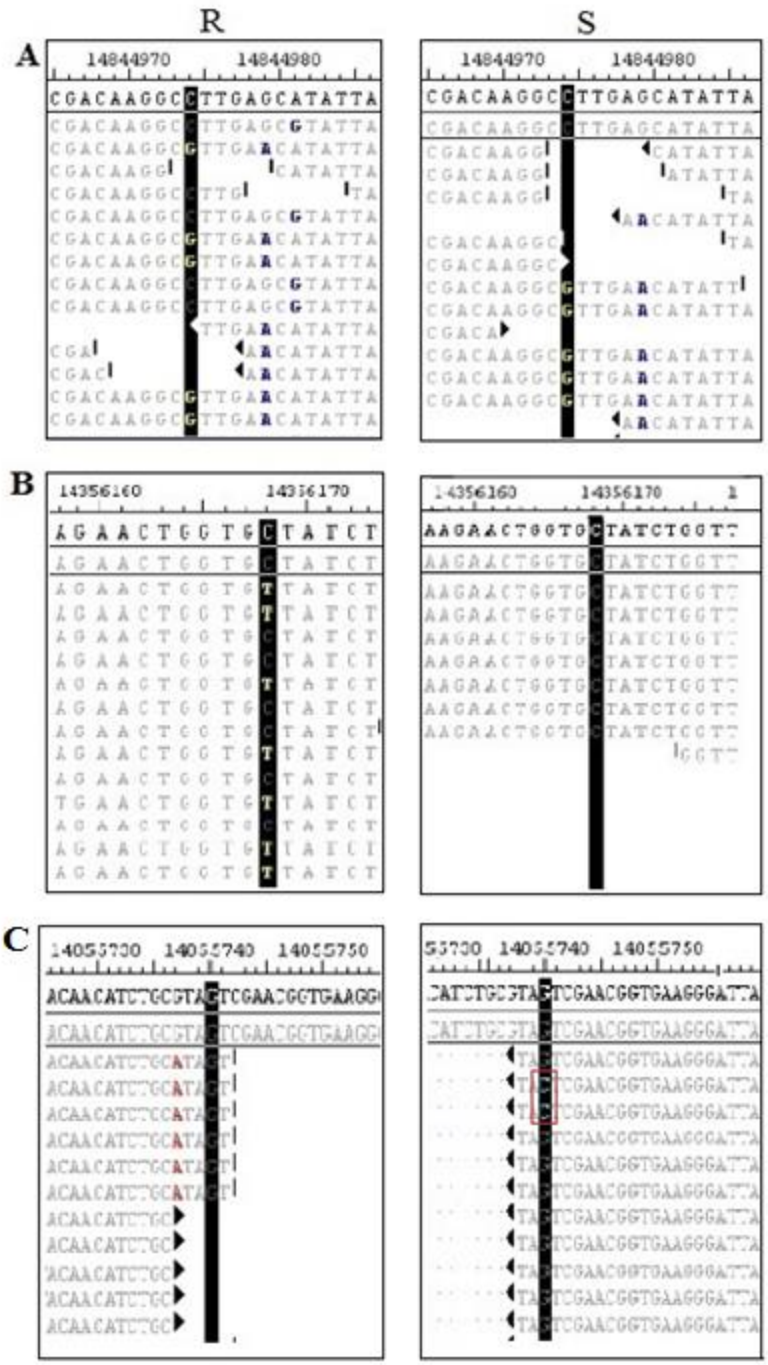

Figure 4.7: Examples of polymorphic and monomorphic variants in PI 219576 using the *B. rapa* reference sequence (A08) displayed by SeqMan Pro (DNASTAR, Madison, WI, USA). SNPs are highlighted in black. **A)** Monomorphic variant at location 14844974; SNP "G" in both R and S bulks. **B)** Polymorphic variant at the location 14356168; SNP "T" in R, but not in S. **C)** Polymorphic variant at the location 14055740; SNP "C" in S, but not in R. The numbers in the first row of each snapshot are the physical location on chromosome A08 in the *B. rapa* reference genome. The DNA sequences shown in the second rows are the reference genome sequences. The sequences below the reference genome are the short reads from either R or S bulks assembled into the reference genome.

4.1.3.3 Validation and genotyping of SNP markers

KASP results for the progeny and parental lines of BC_1 were analyzed in the allelic discrimination plot and samples were scored as '+' for resistance and '−' for susceptibility. The R parental line (PI 219576) was homozygous resistant (allele1/allele1) that separated away individually and the S parental line (CR2748) was homozygous susceptible (allele2/allele2) that clustered within the susceptible population (allele2/allele2). The resistant population was heterozygous resistant (allele1/allele2) which formed a cluster positioned between the susceptible cluster and the R parental line (Figure 4.8).

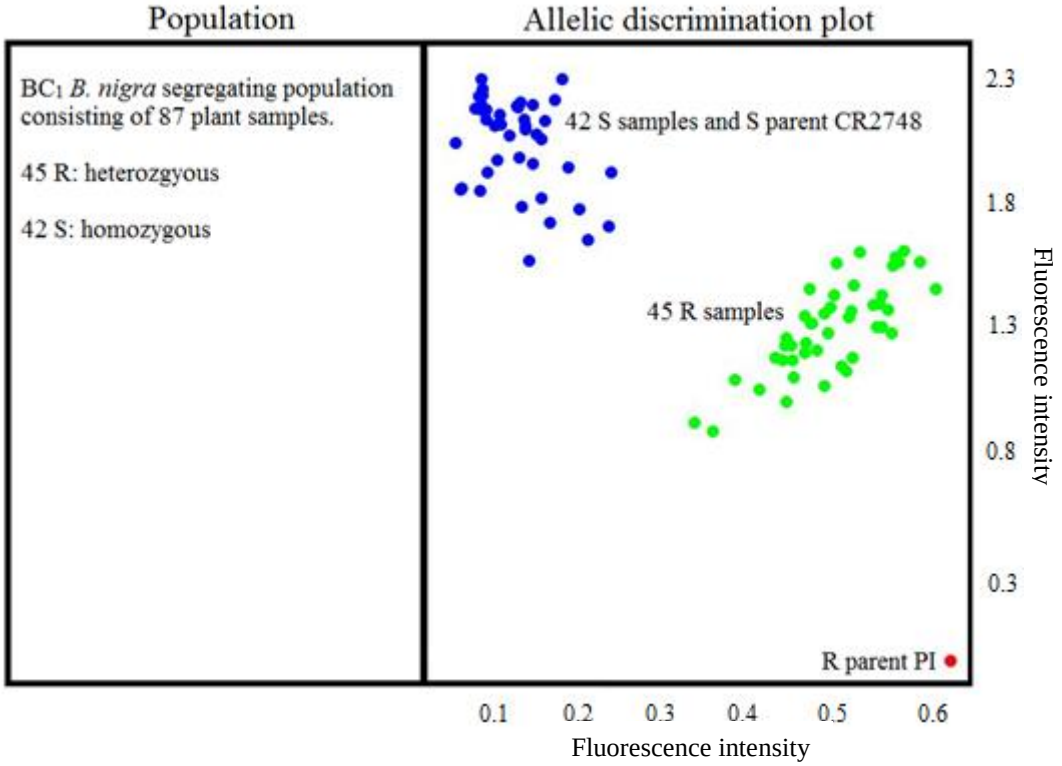

Figure 4.8: Validation and genotyping of SNP markers by KASP method. The R parent (homozygous resistant) segregated individually. The S individuals (homozygous recessive) from BC_1 population shifted to the upper left region in the form of a cluster. The S parent (homozygous recessive) was located within the S cluster. The R individuals (heterozygous resistant) from the BC_1 population were positioned between the S cluster and R parent. The data was taken from the PI 219576 population.

4.1.3.4 Genetic map from PI 219576 cultivar inoculated with pathotype 3

A CR gene designated as *Rcr6a* was mapped on chromosome B3 of PI 219576, in a region homologous to one on *B. rapa* chromosome A08. A total of 39 SNP markers derived from B3 and A08 linkage groups were validated and shown in the genetic map constructed (Figure 4.9). *Rcr6a* was mapped to a range of 3.0 cM, with a physical distance of approximately 0.4 Mb, flanked by the markers SNP_A08_47 and SNP_A08_17. The flanked segment is homologous to the region between 14.36 Mb and 14.84 Mb of A08 (*B. rapa* reference genome sequence, Chromosome v1.5), with 4 genes (*Bra010551, Bra010552, Bra010588* and *Bra010589*) identified as encoding TIR-NBS-LRR class of proteins. *Bra010551* and *Bra010552* were located between 14.34 Mb and 14.36 Mb. *Bra010588* and *Bra010589* were located between 14.57 Mb and 14.58 Mb. A total of six recombinants were identified via comparison of flanking markers and phenotype data over the 234 BC_1 plants, with 5 falling between SNP_A08_47 and *Rcr6a* and 1 between *Rcr6a* and SNP_A08_17. Genetic maps were constructed separately with linked SNP markers from A08 (Figure 4.10A) and B3 (Figure 4.10B) chromosomes. In the A08 linkage map, *Rcr6a* was flanked by SNP markers A08_47 and A08_17. In the B3 linkage map, *Rcr6a* was flanked by SNP markers B3_44 and B3_51.

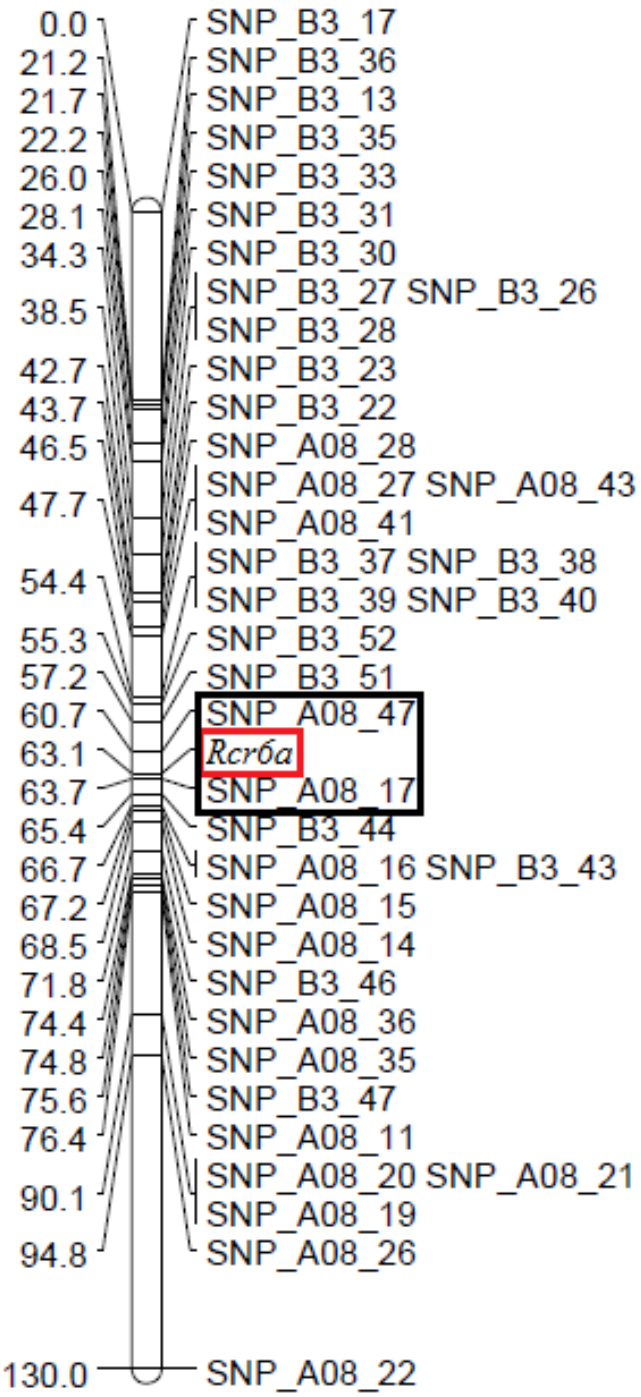

Figure 4.9: Linkage map of B3 chromosome of *B. nigra* from the PI 219576 population showing the location of *Rcr6a* flanked in between the SNP markers A08_47 and A08_17. Both A08 and B3 SNP markers were combined together. Plants were inoculated with pathotype 3 of *P. brassicae.* Genetic distance is given in centiMorgans (cM).

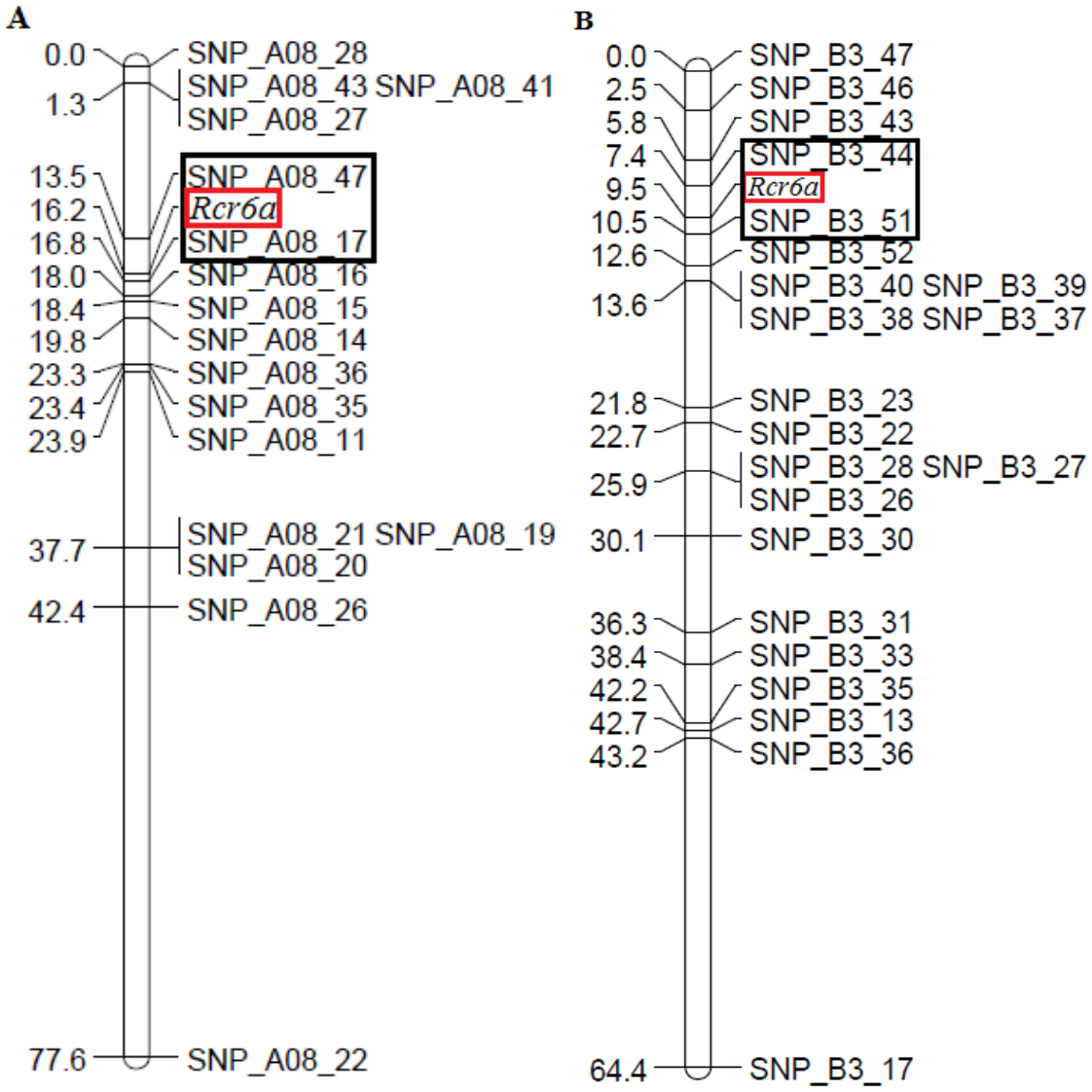

Figure 4.10: Genetic maps of the PI 219576 population of *B. nigra* constructed separately with linked SNP markers from A08 and B3 chromosomes. Plants were inoculated with pathotype 3 of *P. brassicae.* Genetic distance is given in centiMorgans (cM). **A)** A08 linkage map of *Rcr6a* flanked in between SNP markers A08_47 and A08_17. **B)** B3 linkage map of *Rcr6a* flanked in between SNP markers B3_44 and B3_51.

4.1.3.5 Analysis of variants in the target region

Rcr6a from PI 219576 of *B. nigra* was mapped on chromosome B3 in an interval of 3.0 cM homologous to a region on *B. rapa* chromosome A08, located between 14.36 Mb and 14.84 Mb. In the current study, the target regions were analyzed using the PSA method. There were 9 genes annotated in the *Rcr6a* interval based on the A-reference genome v1.5. All the annotated genes showed expression and variants could be identified. Information on BLASTX (best hit) to *A. thaliana* and gene annotation was obtained from http://brassicadb.org/brad/index.php (Appendix 3). Four genes in the *Rcr6a* interval (*Bra010551, Bra010552, Bra010588* and *Bra010589*) encode TNL (TIR-NBS-LRR)-class disease resistance proteins. The numbers of poly variants between the R and S bulks among the TNL genes in the *Rcr6a* region was assessed because the poly variants signify differences in the DNA sequences between R and S bulks. Analysis using the PSA method indicated that the TNL gene (*Bra010552*) carried the highest number of poly variants in the *Rcr6a* interval, suggesting that *Bra010552* is likely the candidate gene (Table 4.5).

Table 4.5: Number of polymorphic variants identified in each TNL gene located in the *Rcr6a* interval in PSA analysis.

Gene	No. of variants within the *Rcr6a* interval
Bra010551	5
Bra010552	25
Bra010588	6
Bra010589	13

4.1.3.6 Transcriptome analysis based on RNA-Seq

The inoculated BC_1 population from PI 219576 was bulked (R and S) and examined for global transcriptomes using RNA-Seq. The reads from R and S bulked RNA-seq samples were assembled into the *B. rapa* reference genome v1.5. Single assembly of the three R bulks and three S bulks was done for gene expression analysis. The QSeq summary report indicated the amount of unique reads and repeated reads assigned (Table 4.6); approximately 66 million unique reads and 71 thousand repeated reads were assigned from a total of six pooled bulks.

Table 4.6: QSeq summary report. RNA-Seq reads from inoculated R and S bulks of PI 219576 leaf samples (BC_1). A-template reference was used.

	Unique Reads Assigned	Repeated Reads Assigned
PI (R1 bulk)	8851341	10520
PI (R2 bulk)	12645829	12844
PI (R3 bulk)	8413683	7390
PI (S1 bulk)	11780139	14847
PI (S2 bulk)	15034905	16834
PI (S3 bulk)	8959439	8984

A total of 11 DEGs were identified in PI 219576 at 95% confidence, with eight genes up-regulated and 3 down-regulated in the R samples relative to the S samples. The genes in the mapped *Rcr6a* region between 14.36 Mb and 14.84 Mb of chromosome A08 of *B. rapa* were expressed. However, none of the TIR-NBS-LRR genes were differentially expressed and many of them showed no major difference in expression levels between inoculated R and S. A heat map of the 11 DEGs for PI 219576 is shown in Figure 4.11 and the list of annotated DEGs for PI 219576 is provided in Appendix 4.

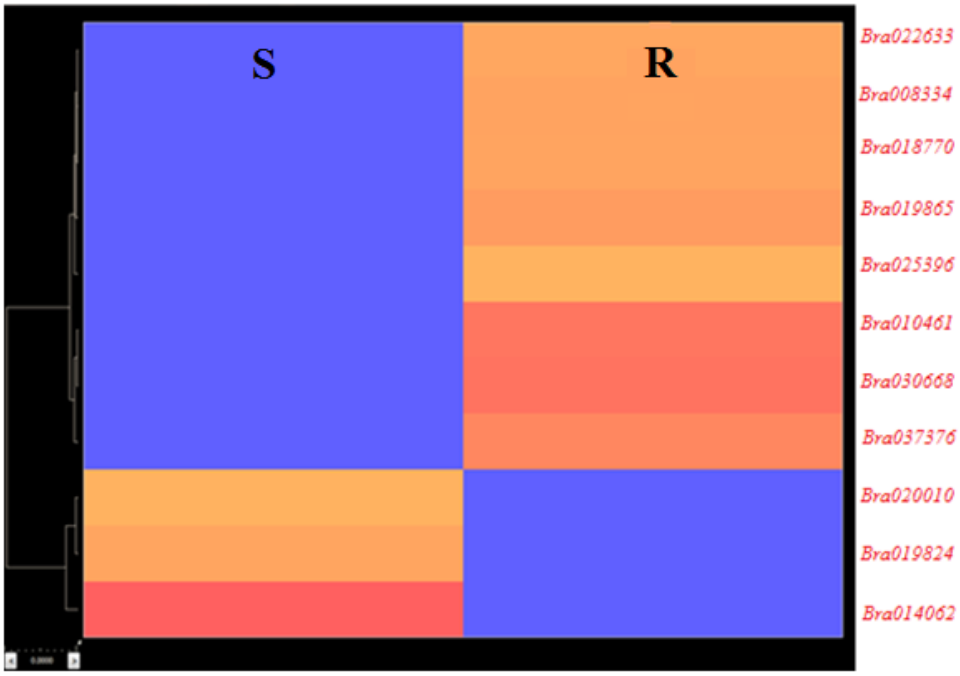

Figure 4.11: Heat map of 11 DEGs in the PI 219576 resistant and susceptible bulks. Up-regulation of genes is shown in orange and down-regulation of genes in blue. Eight genes were up-regulated and three genes were down-regulated in the R samples relative to S samples. The gene accessions are shown on the right.

RT-qPCR analysis of four selected genes from PI 219576 over the same R and S bulk leaf samples validated the RPKM (reads assigned per kilobase of target per million mapped reads) calculations and statistical analysis of transcript data (Figure 4.12A). The patterns for RNA-Seq are quite consistent but not for RT-qPCR. The RT-qPCR data confirmed the up-regulation of an S-locus lectin protein kinase family protein gene (*Bra019865*), glycine-rich protein gene (*Bra018770*) and two proteins with unknown function genes (*Bra037376* and *Bra030688*) in R samples from PI 219576 detected via RNA-Seq. Root tissues were also used for RT-qPCR analysis (Figure 4.12B).

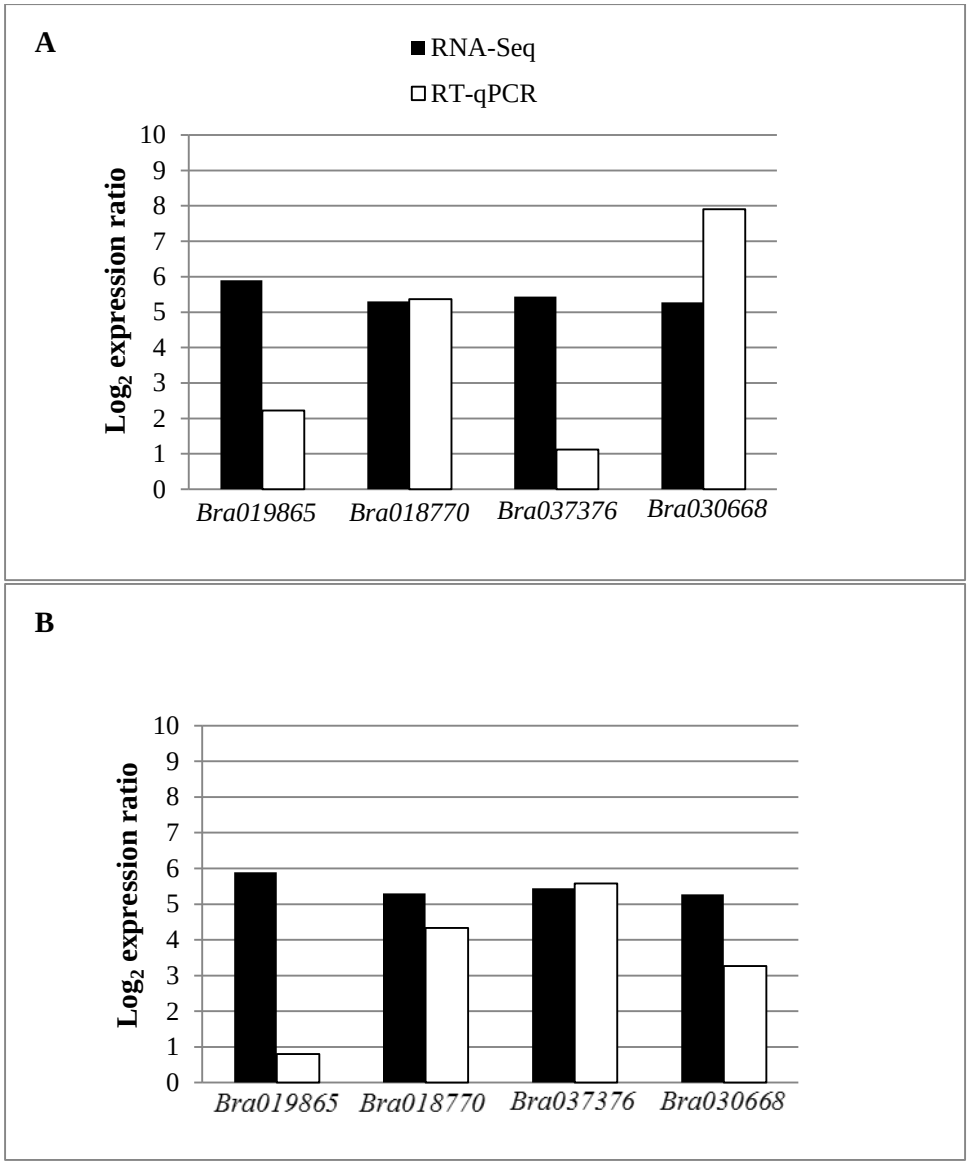

Figure 4.12: Validation of RPKM-calculated expression ratios for selected DEGs from PI 219576 using RT-qPCR. RPKM values from RNA-Seq are denoted in black, and RT-qPCR results in white. Plants were inoculated with pathotype 3 of *P. brassicae*. Leaf and root tissues were collected 2 wpi. **A)** Leaf tissues. **B)** Root tissues. *Bra019865* = S-locus lectin protein kinase family protein; *Bra018770* = glycine-rich protein; *Bra037376* = unknown function protein; *Bra030668* = unknown function protein.

4.1.3.7 Annotation of up- and down-regulated DEGs

The up- and down-regulated DEGs in the R samples relative to S samples obtained from RNA-Seq data were functionally annotated using the *Brassica* database. A plant defence gene (*Bra025396*), F-box family protein gene (*Bra022633*), glycine-rich protein gene (*Bra018770*) and S-locus lectin protein kinase gene (*Bra019865*) were up-regulated. A myeloblastosis (MYB) family transcription factor gene (*Bra019824*) and emp24 family protein gene (*Bra020010*) were down-regulated. Complete annotations of up- and down-regulated DEGs for PI 219576 are shown in Appendix 4.

4.2 Clubroot assessment of *B. nigra* line BRA 192/78

4.2.1 Single dominant gene confirmation

All of the BRA192/78 plants were resistant to pathotype 3 of *P. brassicae*, showing no clubroot symptom at 5 weeks after inoculation (Figure 4.13A). All of the ACDC and CR2748 plants were susceptible (Figure 4.13B). Complete resistance to pathotype 3 of *P. brassicae* was found in 92 F_1 plants derived from the cross of CR2748 with BRA192/78 (Figure 4.13C), indicating that BRA192/78 was likely a homozygous resistant line and the resistance was controlled by dominant gene(s). A total of 90 F_2 plants from BRA192/78 ($X^2 = 0.73$, $P = 0.39$) showed 3:1 segregation ratio for resistance and susceptibility (Figure 4.13D). Evaluation of 242 BC_1 plants from BRA192/78 showed 1:1 segregation ratio ($X^2 = 1.65$, $P = 0.20$) (Figure 4.13E), indicating that CR in BRA192/78 is controlled by a single dominant gene. The Chi-Square statistics for the 3:1 and 1:1 ratios are summarized in Table 4.7. The CR gene in BRA192/78 was designated as *Rcr6b*.

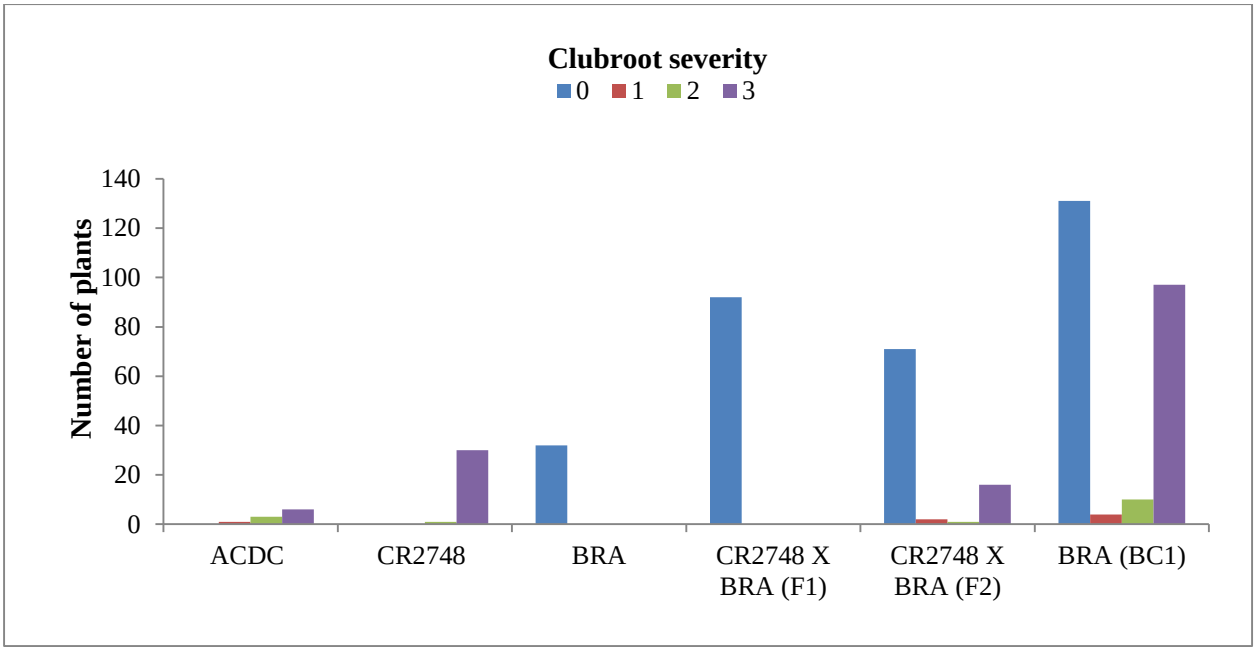

Figure 4.13: Segregation of clubroot resistance for parents (BRA192/78 and CR2748), F_1, F_2 and BC_1 populations of *B. nigra*. Plants were inoculated with pathotype 3 of *P. brassicae*. Resistant parent showed complete resistance. Susceptible parent (CR2748 of *B. nigra*) and susceptible control (ACDC of *B. rapa*) showed clubroot symptoms. F_1 population showed complete resistance. F_2 population showed segregation ratio of 3R:1S. BC_1 population showed segregation ratio of 1R:1S.

Table 4.7: Chi-Square tests confirmed BC_1 and F_2 segregation ratios in the BRA192/78 population of *B. nigra*. Pathotype 3 of *P. brassicae* was used as inoculum.

	Phenotype					
Population	Total	Resistant (R)	Susceptible (S)	Expected ratio (R:S)	χ^2	*P-value*
BC_1	242	131	111	1:1	1.65	0.20
F_2	90	71	19	3:1	0.73	0.39

4.2.2 Mapping of *Rcr6b* with SNP markers via BSR-Seq using the reference genomes of *B. rapa* and *B. oleracea*

A MBS approach via BSR-Seq was carried out for BRA192/78 because SNP markers linked to *Rcr6a* from PI 219576 were not completely associated with *Rcr6b*. The read length of the MiSeq data was 76 bases, and the total amount of data was 8.9 Gb for BRA192/78. The short reads from RNA-Seq were aligned to *B. rapa* and *B. oleracea* reference genomes because the *B. nigra* genome has not been published.

4.2.2.1 RNA-Seq and sequence alignment

Pooled sample assembly of the three R bulks and three S bulks from BRA 192/78 was carried out to identify variants. A total of 25.3 M sequences, 1650.2 Mb in length, with 6-fold coverage of the reference A-genome were assembled into *B. rapa* chromosomes from the pool of three R bulks, and 41.2 M sequences, 2637.7 Mb in length, with 10-fold coverage were assembled from the pool of three S bulks (Table 4.8). More sequences were aligned into the longer chromosomes A03 and A09 and fewer sequences were aligned into the shorter chromosomes A04 and A10 (Table 4.8).

A total of 25.4 M sequences, 1655.3 Mb in length, with 4-fold coverage of the reference C-genome were assembled into *B. oleracea* chromosomes from the pool of three R bulks, and 41.0 M sequences, 2657.3 Mb in length, with 6-fold coverage were assembled from the pool of three S bulks (Table 4.9). More sequences were aligned into chromosomes C3 and C4, but C9 is a longer chromosome than C4; fewer sequences were aligned into C6 and C7, but C8 is a shorter chromosome than C7 (Table 4.9).

Table 4.8: Short reads from BRA192/78 assembled into chromosomes of the *B. rapa* reference genome in pooled sample assembly (PSA) of the resistant (R) and susceptible (S) bulks.

Chromosome number	Chromosome size (bases x 10^6)	Number of sequences (x 10^6)		Accumulated length of sequences (bases x 10^6)	
		R	S	R	S
A01	26.8	2.4	3.9	154.0	250.8
A02	27.0	2.1	3.5	136.9	222.0
A03	31.8	3.5	5.7	225.2	371.7
A04	19.3	2.2	3.3	140.9	215.0
A05	25.3	2.3	3.7	153.9	239.4
A06	25.2	2.5	4.1	166.2	271.1
A07	25.9	2.6	4.1	167.8	264.0
A08	20.8	2.5	3.9	163.3	256.5
A09	38.9	3.6	5.8	237.5	376.4
A10	16.4	1.6	2.6	104.5	170.8
Total	257.4	25.3	41.2	1650.2	2637.7

Table 4.9: Short reads from BRA192/78 assembled into chromosomes of the *B. oleracea* reference genome in pooled sample assembly (PSA) of the resistant (R) and susceptible (S) bulks.

Chromosome number	Chromosome size (bases x 10^6)	Number of sequences (x 10^6)		Accumulated length of sequences (bases x 10^6)	
		R	S	R	S
C1	43.8	2.5	4.1	160.5	262.9
C2	52.9	2.3	3.7	147.0	236.7
C3	65.0	4.1	6.8	271.1	441.9
C4	53.7	3.5	5.3	228.1	343.2
C5	46.9	2.9	4.6	188.0	298.2
C6	39.8	2.0	3.4	133.9	219.8
C7	48.4	2.2	3.6	141.6	236.2
C8	41.8	3.0	4.8	194.2	312.5
C9	54.7	2.9	4.7	190.9	305.9
Total	447.0	25.4	41.0	1655.3	2657.3

A total of 33.2 M sequences, 2281.9 Mb in length, with 6-fold coverage of the reference B-genome were assembled into *B. nigra* chromosomes from the pool of three R bulks, and 53.3 M sequences, 3704.6 Mb in length, with 10-fold coverage were assembled from the pool of three S bulks (Table 4.10). More sequences were aligned into the chromosomes B6 and B8, but B2 is a longer chromosome than B6; fewer sequences were aligned into chromosomes B4 and B7, but B6 is the shortest chromosome (Table 4.10).

Table 4.10: Short reads from BRA192/78 assembled into chromosomes of the *B. nigra* reference genome in pooled sample assembly (PSA) of the resistant (R) and susceptible (S) bulks.

Chromosome number	Chromosome size (bases x 10^6)	Number of sequences (x 10^6)		Accumulated length of sequences (bases x 10^6)	
		R	S	R	S
B1	42.2	4.1	6.4	280.9	437.8
B2	52.7	4.1	6.7	282.9	463.9
B3	46.8	4.4	7.1	304.5	493.3
B4	43.4	3.8	6.1	267.6	430.1
B5	51.8	4.1	6.7	280.9	462.6
B6	36.7	4.8	7.7	335.9	536.1
B7	41.6	3.2	5.4	223.4	376.6
B8	54.2	4.4	7.2	305.8	504.2
Total	369.4	33.2	53.3	2281.9	3704.6

4.2.2.2 Identification of variants

The numbers of variants in BRA 192/78 was identified using the PSA method aligned with the A-, C- and B-reference genomes. There were a total of 241.2 K poly variants when aligned with the A-genome, 274.2 K poly variants when aligned with the C-genome, and 268.0 K poly variants when aligned with the B-genome. Chromosome A08 had the highest percentage of variants at 28.6% in the A-genome (Figure 4.14A), C8 had the highest percentage of variants at 43% in the C-genome (Figure 4.14B), and B3 had the highest percentage of variants at 50.5% in the B-genome (Figure 4.14C).

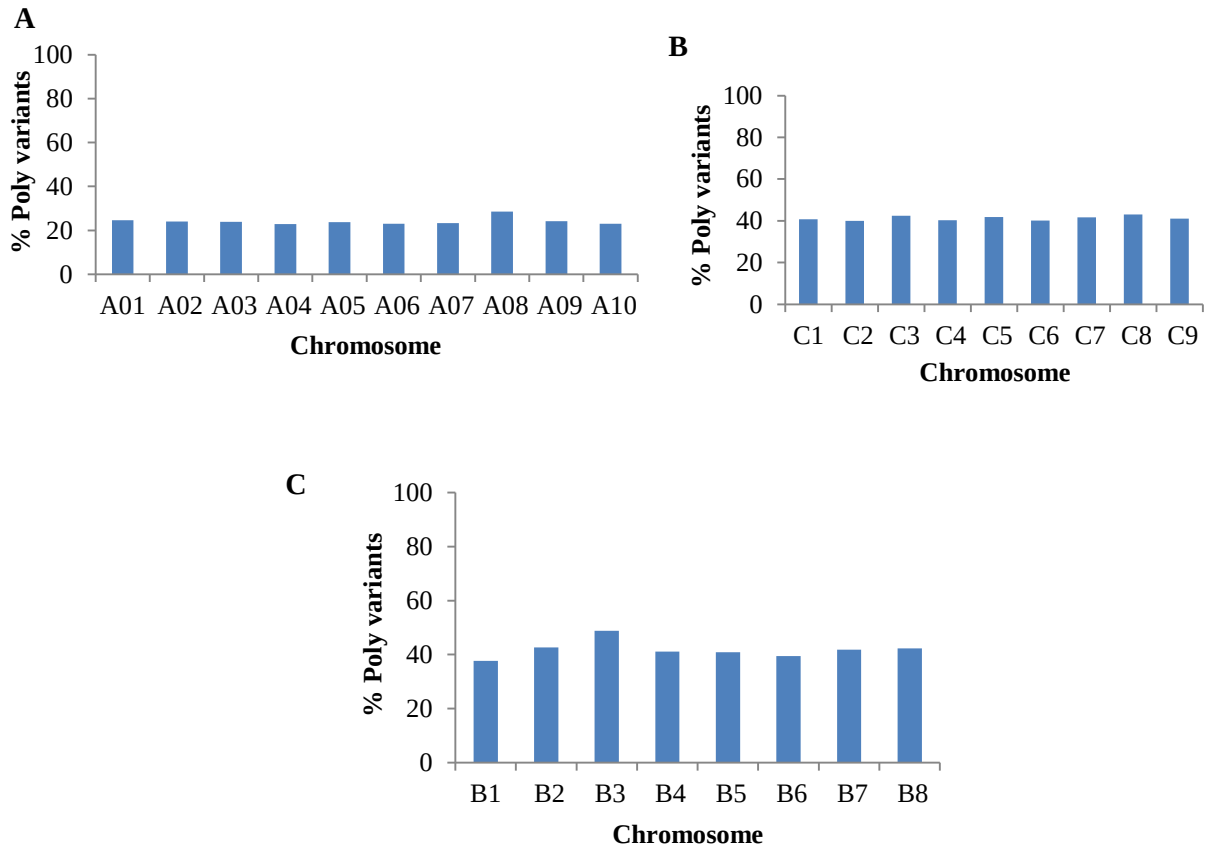

Figure 4.14: Percentage of poly variants (SNPs and InDels) from BRA192/78 using various *Brassica* reference genomes. **A)** Percentage of poly variants by PSA using the A reference genome of *B. rapa*. **B)** Percentage of poly variants by PSA using the C reference genome of *B.oleracea*. **C)** Percentage of poly variants by PSA using the B reference genome of *B. nigra*.

The poly variant distribution (%) was analyzed using the A and C reference genomes. The pattern for BRA192/78 was similar to the previous PI 219576 results, with A08 containing the highest percentage of poly variants located in the physical range from approximately 12 Mb to 20 Mb (Figure 4.15); and C8 containing the highest percentage of poly variants located in the physical range from approximately 18 Mb to 24 Mb (Figure 4.16). The B reference sequence was not yet complete, so the distribution results did not provide a clear observation to determine which chromosome contained the highest percentage of poly variants along the physical distance.

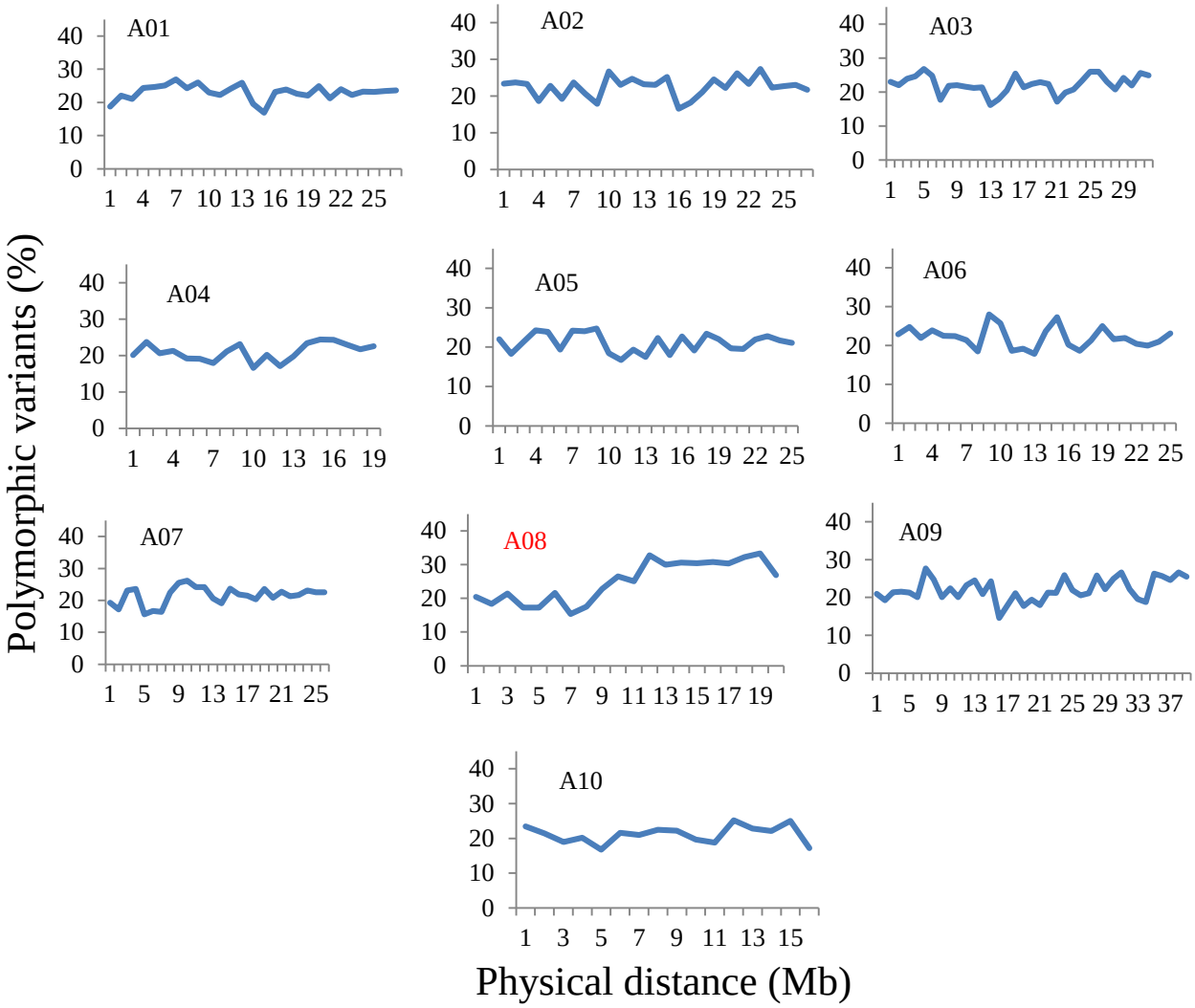

Figure 4.15: Percentage of poly variants (y-axis) distributed along the physical distance (x-axis) in BRA192/78 using the A-reference genome of *B. rapa*. Chromosome A08 contained the highest percentage of poly variants at a wider physical range compared to other chromosomes.

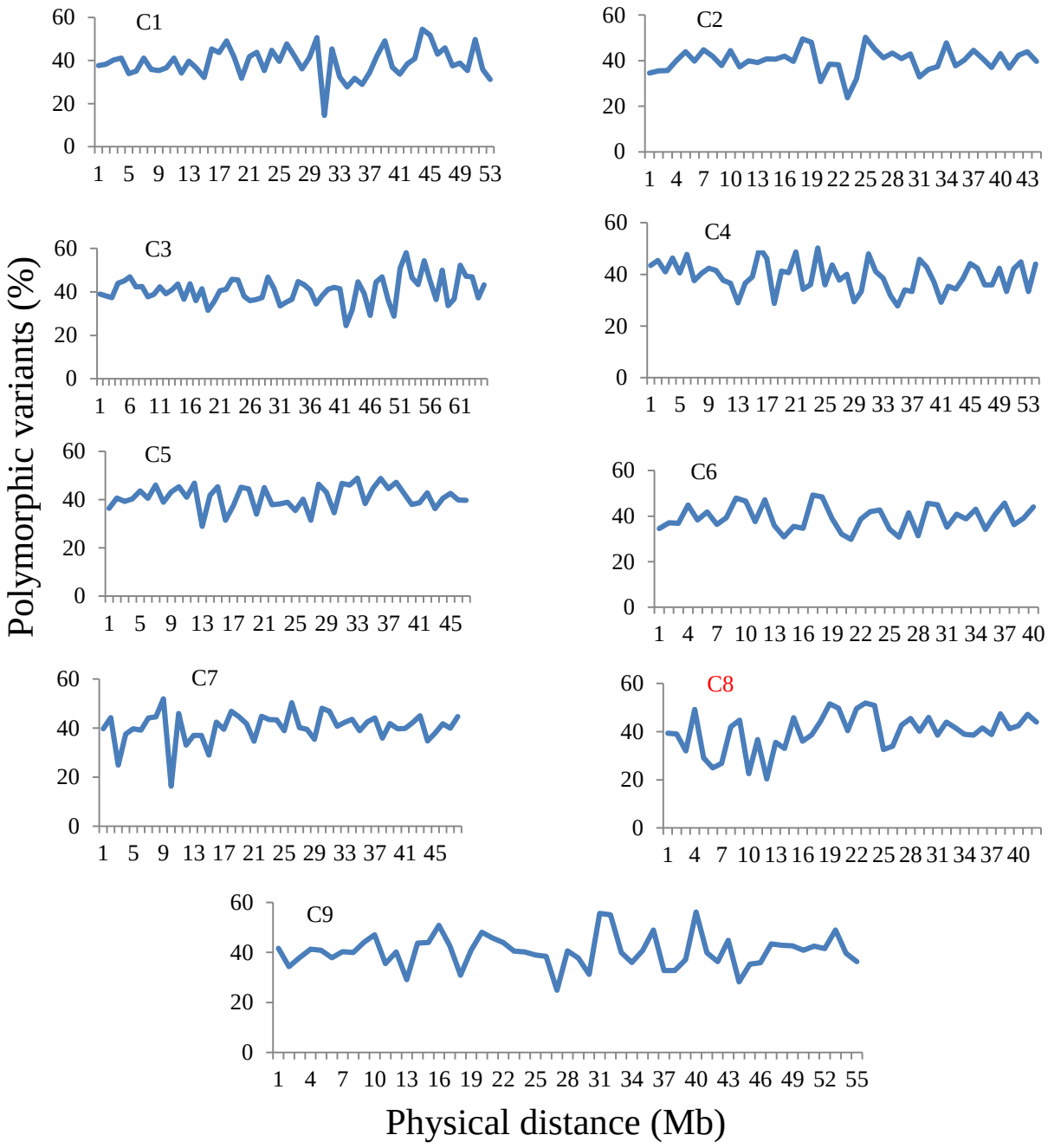

Figure 4.16: Percentage of poly variants (y-axis) distributed along the physical distance (x-axis) in BRA192/78 using the C-reference genome of *B. oleracea*. Chromosome C8 contained the highest percentage of poly variants at a wider physical range compared to other chromosomes.

4.2.2.3 Genetic map from BRA192/78 cultivar inoculated with pathotype 3

The BC_1 population from BRA192/78 was tested with SNP markers linked to *Rcr6a*, which showed that the CR gene in BRA192/78 was likely in the *Rcr6a* region, and was designated as *Rcr6b.* A total of 19 SNP markers derived from B3 and A08 linkage groups were validated and shown in the genetic map constructed (Figure 4.17). *Rcr6b* was mapped to a range of 0.5 cM in the B3 chromosome, with a physical distance of approximately 0.3 Mb. *Rcr6b* was flanked by the markers SNP_B3_65 and SNP_B3_67 at the upper region; and SNP_B3_62, SNP_B3_63 and SNP_B3_64 at the lower region. The SNP_A08_17 co-segregated with *Rcr6b*. The flanked segment is homologous to the region between 12.76 Mb and 14.84 Mb of A08 (*B. rapa* reference genome sequence, Chromosome v1.5), with two genes (*Bra010551* and *Bra010552*) between 14.34 Mb and 14.36 Mb identified as encoding TIR-NBS-LRR class of proteins. Another gene (*Bra034556*) located near the flanked segment at 12.05 Mb was identified as TIR-NBS-LRR. There were no recombinants identified via comparison of flanking markers and phenotype data over the 242 BC_1 plants. Genetic maps were constructed separately with linked SNP markers from A08 (Figure 4.18A) and B3 (Figure 4.18B) chromosomes. In the A08 linkage map, *Rcr6b* was flanked by SNP markers A08_47 and A08_56, with A08_17 co-segregating with the CR gene. In the B3 linkage map, *Rcr6b* was flanked by SNP markers B3_62, B3_63 and B3_64 at the upper region; and SNP markers B3_65 and B3_67 at the lower region.

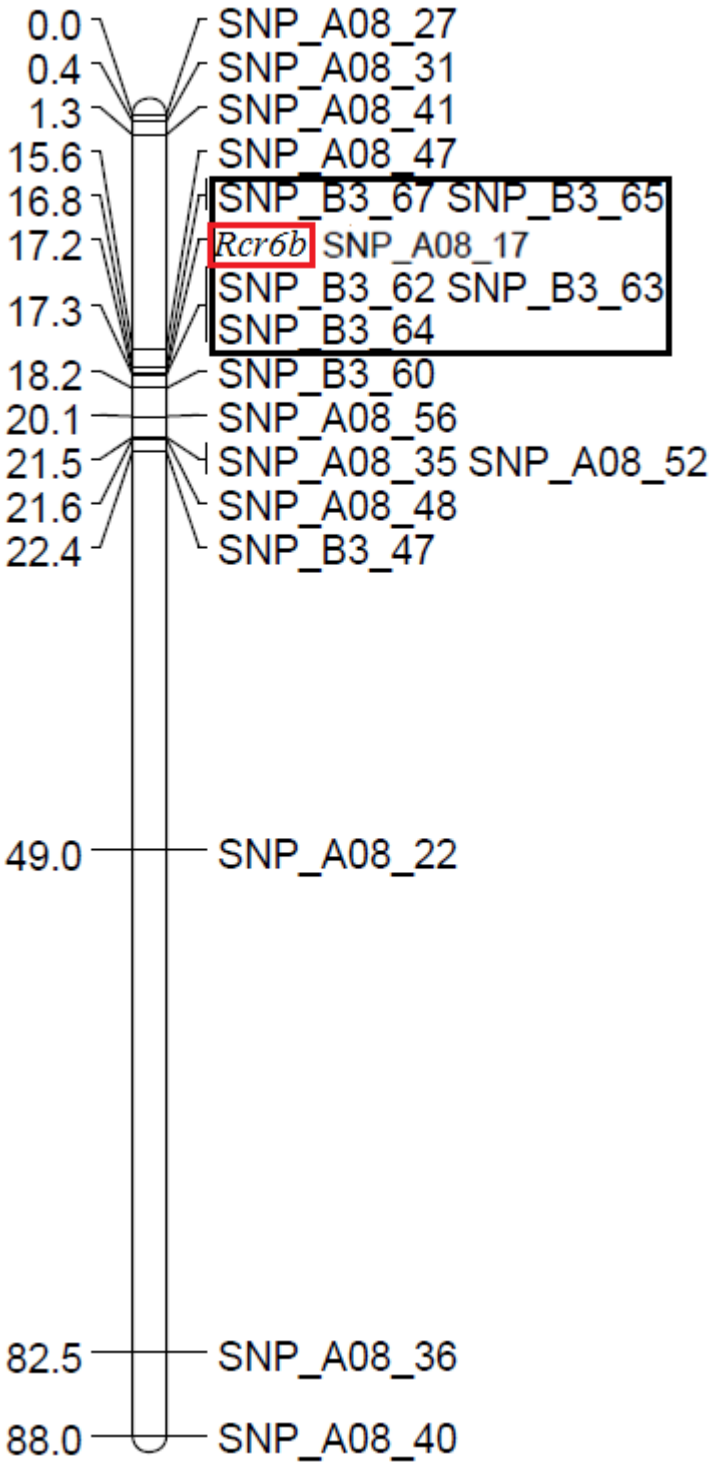

Figure 4.17: Linkage map of B3 chromosome from the BRA192/78 population of *B. nigra* showing the location of *Rcr6b* flanked in between SNP markers B3_65 and B3_67 at the upper region; and SNP markers B3_62, B3_63 and B3_64 at the lower region. SNP_A08_17 co-segregated with *Rcr6b*. Plants were inoculated with pathotype 3 of *P. brassicae*. Genetic distance is given in centiMorgans (cM).

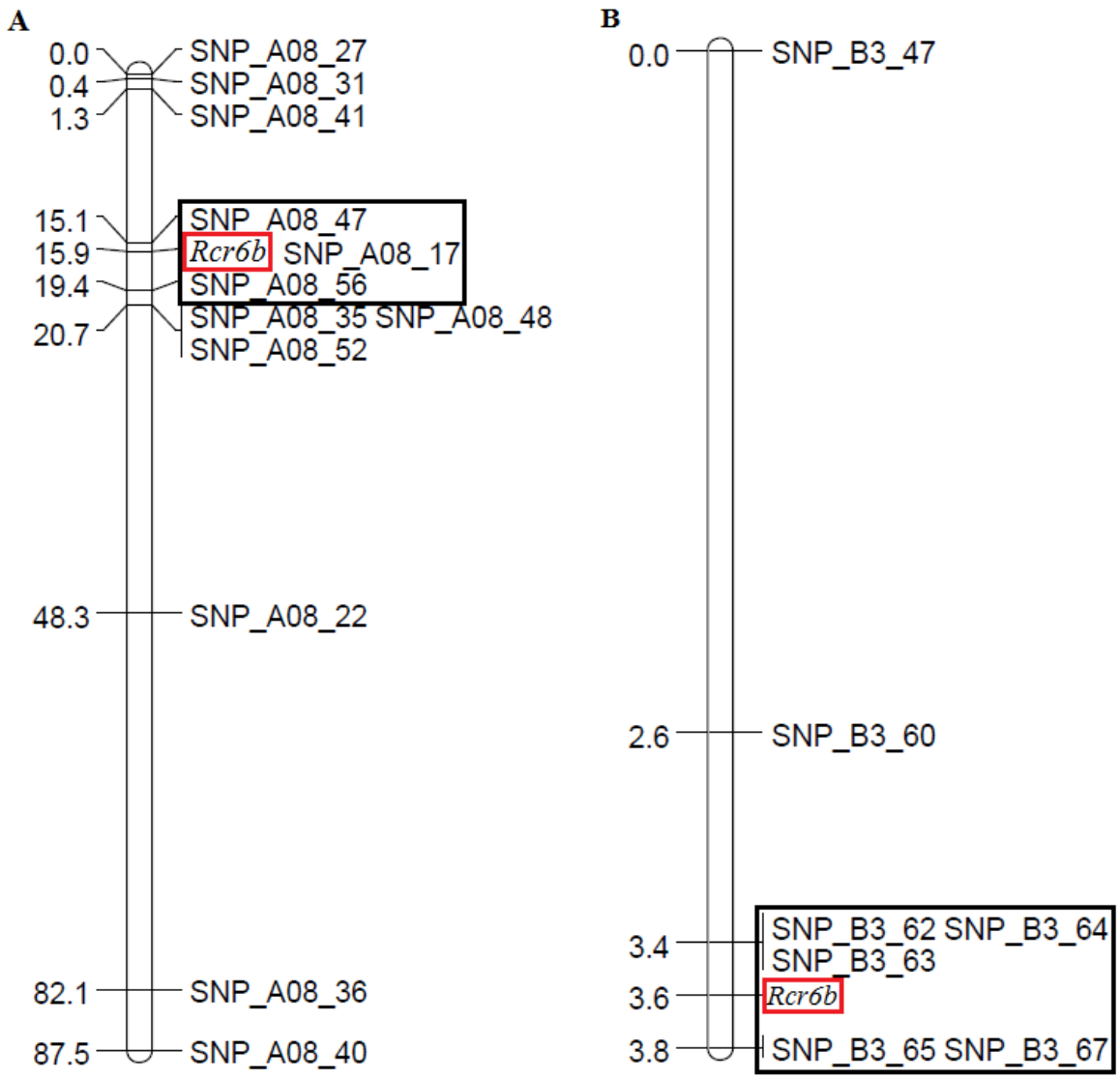

Figure 4.18: Genetic maps of the BRA192/78 population of *B. nigra* constructed separately with linked SNP markers from A08 and B3 chromosomes of *B. rapa* and *B. nigra*, respectively. Plants were inoculated with pathotype 3 of *P. brassicae*. Genetic distance is given in centiMorgans (cM). **A)** A08 linkage map with *Rcr6b* flanked in between SNP markers A08_47 and A08_56; A08_17 co-segregated with *Rcr6b*. **B)** B3 linkage map with *Rcr6b* flanked in between SNP markers B3_62, B3_63 and B3_64 at the upper region; and SNP markers B3_65 and B3_67 at the lower region.

4.2.2.4 Analysis of variants in the target region

The gene *Rcr6b* from BRA 192/78 was mapped in an interval of 0.5 cM homologous to a region on chromosome A08 of *B. rapa*, located between 12.76 Mb and 14.84 Mb. In the current study, the target regions were analyzed using the PSA method. There were 19 genes annotated in the *Rcr6b* interval, based on the A-reference genome v1.5. All of the annotated genes showed expression and the variants could be identified. Information on BLASTX (best hit) to *A. thaliana* and gene annotation was obtained from http://brassicadb.org/brad/index.php (Appendix 3). Two genes in the *Rcr6b* interval (*Bra010551* and *Bra010552*) encode TNL-class disease resistance proteins. The numbers of poly variants between the R and S bulks among the TNL genes in the *Rcr6b* region was assessed because the poly variants signify differences in the DNA sequences between the R and S bulks. Analysis using the PSA method indicated that the TNL gene *Bra010552* carried the highest number of poly variants in the *Rcr6b* interval, suggesting that *Bra010552* is likely the candidate gene (Table 4.11).

Table 4.11: Number of polymorphic variants identified in each TNL gene located in the *Rcr6b* interval in PSA analysis.

Gene	No. of variants within the *Rcr6b* interval
Bra010551	4
Bra010552	22

4.2.2.5 Transcriptome analysis based on RNA-Seq

The inoculated BC_1 population from BRA192/78 was bulked (R and S) and examined for global transcriptomes using RNA-Seq. The reads from R and S bulked RNA-seq samples were assembled into the *B. rapa* reference genome v1.5. Single assembly of the three R bulks and three S bulks was performed for gene expression analysis. The QSeq summary report indicated the amount of unique reads and repeated reads assigned (Table 4.12). Approximately 63 million unique reads and 67 thousand repeated reads were assigned from a total of six pooled bulks.

Table 4.12: QSeq summary report. RNA-Seq reads from inoculated R and S bulks of BRA192/78 leaf samples (BC_1) of *B. nigra* . A-template reference was used.

	Unique Reads Assigned	**Repeated Reads Assigned**
BRA (R1 bulk)	8121384	8369
BRA (R2 bulk)	8951807	9787
BRA (R3 bulk)	6958810	8505
BRA (S1 bulk)	8121372	8696
BRA (S2 bulk)	20568684	20911
BRA (S3 bulk)	10276906	11026

A total of 382 DEGs were identified in BRA192/78 at 95% confidence, with 81 genes up-regulated and 301 down-regulated in the R samples relative to S samples. The genes in the mapped *Rcr6b* region between 12.76 Mb and 14.84 Mb of A08 were expressed. Only one of the TIR-NBS-LRR genes (*Bra010551*) was differentially expressed and many of them showed no major difference in expression levels between inoculated R and S. *Bra010551* exhibited an insignificant change in the inoculated R treatment relative to the inoculated S treatment. A scatter plot of the 382 DEGs for BRA192/78 is shown in Figure 4.19 and the list of annotated DEGs for BRA192/78 is provided in Appendix 4.

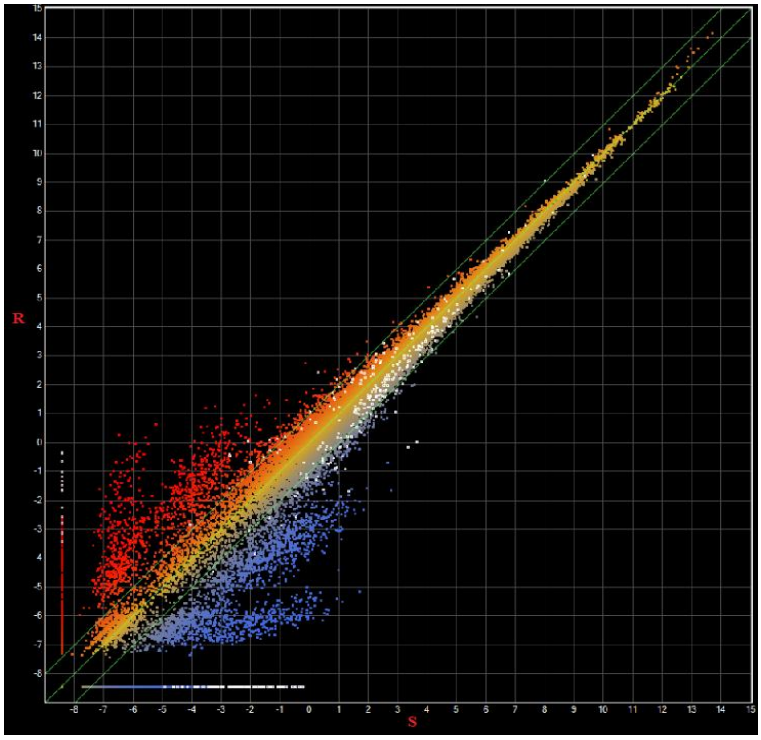

Figure 4.19: Scatter plot of 382 DEGs in the BRA192/78 resistant (y-axis) and susceptible (x-axis) bulks of *B. nigra*. There were 81 genes up-regulated and 301 genes down-regulated in the R samples relative to S samples.

A RT-qPCR analysis of seven selected genes from BRA192/78 over the same R and S bulk leaf samples showed a trend consistent with that of RPKM calculations and statistical analysis of transcript data (Figure 4.20A). The RT-qPCR data confirmed the down-regulation of a PR transcriptional factor (*Bra023748*), extensin-like repeat protein (*Bra014807*), F-box family protein (*Bra017261*), MYB domain protein (*Bra023486*), WOX5 transcription factor (*Bra034855*), plant G-box-binding factor (*Bra008670*) and WRKY44 transcription factor (*Bra005210*) genes in R samples from BRA192/78 as indicated in RNA-Seq analysis. Root tissues were also used for RT-qPCR analysis (Figure 4.20B).

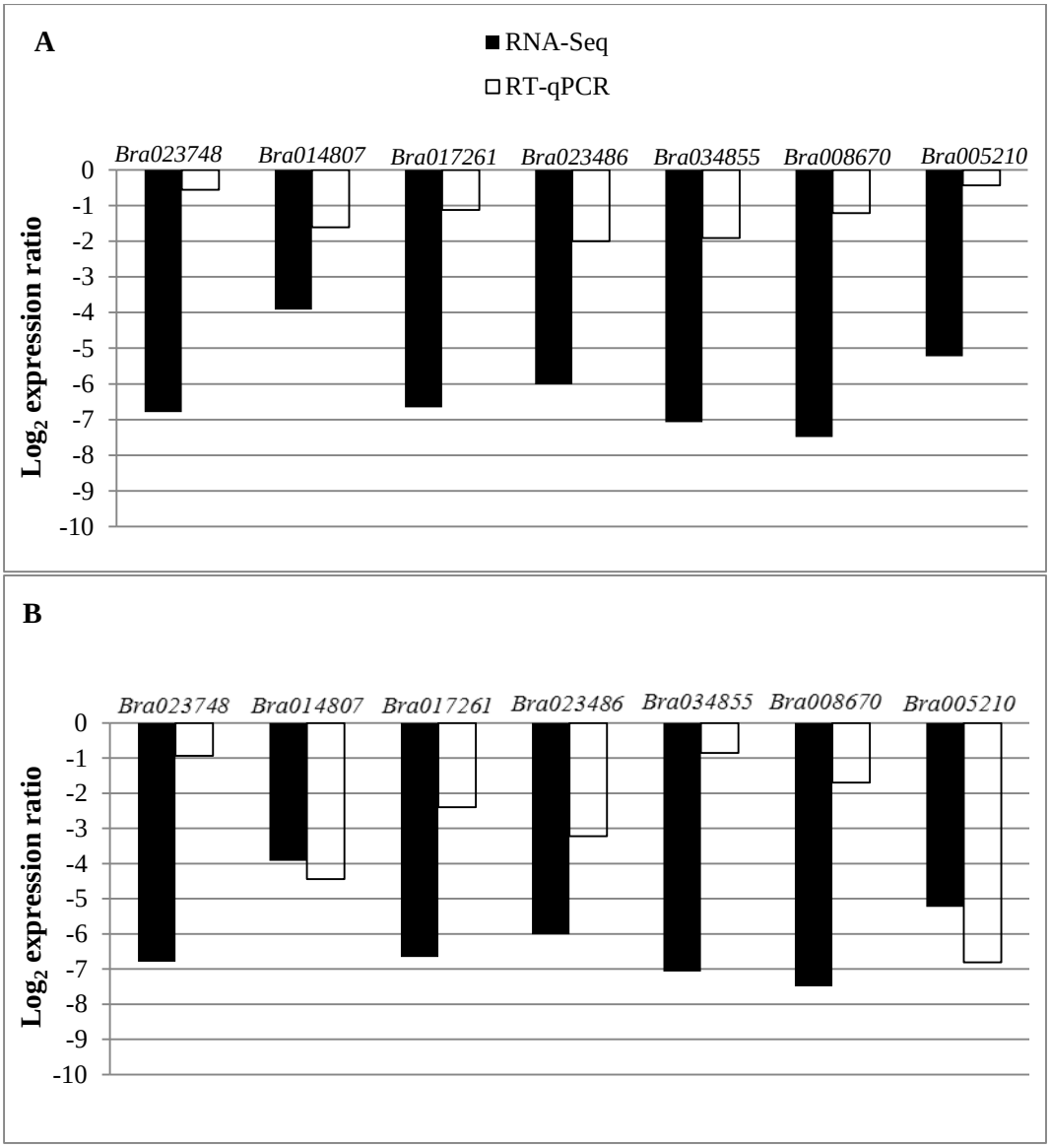

Figure 4.20: Validation of RPKM-calculated expression ratios for selected DEGs from BRA192/78 using RT-qPCR. RPKM values from RNA-Seq are denoted in black, and RT-qPCR results in white. Plants were inoculated with pathotype 3. Leaf and root tissues were collected 2 wpi. **A)** Leaf tissues. **B)** Root tissues. *Bra023748* = PR transcriptional factor; *Bra014807* = extensin-like repeat protein; *Bra017261* = F-box family protein; *Bra023486* = MYB domain protein; *Bra034855* = WOX5 transcription factor; *Bra008670* = plant G-box-binding factor; *Bra005210* = WRKY44 transcription factor.

4.2.2.6 Annotation of up- and down-regulated DEGs

A defence response gene (*Bra016325*) was down-regulated. Several DEGs associated with plant disease resistance proteins (*Bra010542*, *Bra010551*, *Bra026922*, and *Bra002117*) and PR transcriptional factors (*Bra023748*, *Bra023404*, *Bra033138*, and *Bra017612*) were down-regulated. A gene involved in cell development (*Bra032981*) and an IAA-inducible gene (*Bra026166*) were down-regulated. The down-regulated auxin-responsive protein genes were *Bra009535*, *Bra025770*, *Bra025857*, *Bra012231*, *Bra003665*, *Bra002327*, and *Bra008836*. Down-regulated oxidoreductase family protein genes included *Bra004558* and *Bra037749*. Both defence-related ethylene-responsive protein (*Bra035761*) and plant lipid transfer protein (*Bra000775*) genes were up-regulated. A fatty acid desaturase family protein gene (*Bra030638*) was up-regulated. Complete annotations of up- and down-regulated DEGs for BRA192/78 are shown in Appendix 4.

5. DISCUSSION

The main objective of this research was to map clubroot resistance genes in two resistant lines (PI 219576 and BRA192/78) of *B. nigra*. Two CR alleles, *Rcr6a* and *Rcr6b*, were mapped with closely-linked SNP markers. Another objective was to analyze differential gene expression in the host when inoculated with *P. brassicae* to better understand the molecular mechanisms of clubroot resistance. Successful research outcomes in genetic mapping of CR genes, problems and challenges during the study and future perspectives of this research are discussed in this section.

5.1 Importance of mapping CR genes in *Brassica nigra*

So far, no studies have been conducted to map CR genes in the *Brassica* B-genome species – a gene pool that can be used for breeding *B. juncea* and *B. carinata* cultivars. Although the A-, B- and C-genome diploid species were found to carry CR to different *P. brassicae* pathotypes, clubroot disease resistance has only been reported in one amphidiploid species, *Brassica napus*, but not in *B. juncea* and *B. carinata*, which contain the B-genome (Diederichsen et al., 2009).

Resistant *B. nigra* can be used as a valuable genetic resource for introgression of CR into *B. juncea* and *B. carinata* by interspecific hybridization or re-synthesis of these species. In addition, novel traits such as disease resistance and tolerance to abiotic stress can be transferred into *B. napus* (Navabi et al., 2013). However, it is important to consider that there is limited probability of introgression of resistance from the B-genome chromosome of *B. nigra* to the A- or C-genome chromosome of *B. napus*. Navabi et al. (2010) verified that B-genome chromosomes can be transferred into *B. napus* using a backcrossing approach, but there were still major challenges to developing genotypes with introgressed B-genome material. The main reason was because the B-genome chromosomes appeared to be inherited as either whole chromosomes

or chromosomes with terminal deletions. Another study demonstrated the presence of B-genome chromatin in *Brassica napus* x *B. juncea* interspecific progeny (Schelfhout et al., 2006). Due to differentiation of the *Brassica* genomes caused by genome rearrangements and gene deletions, there is low homeology between the B chromosomes with the A/C chromosomes (Attia et al., 1987). Inversions and significant structural divergence have also been proposed to have an effect on plant evolution and could have initiated the isolation and speciation of the B-genome (Navabi et al., 2013). Yet at the genetic and sequence level, there is distinct conservation across the three *Brassica* genomes, such as the comparable level of karyotype and microstructure divergence (Panjabi et al., 2008). Synteny analysis between the A- and B-genomes was performed in *B. juncea* through RNA-Seq (Paritosh et al., 2014). Homeology blocks of linkage group B3 with A03, A06 and A09 were found. In the current study, synteny between B3 and A08 was identified. We successfully fine mapped B3 CR genes *Rcr6a* and *Rcr6b* of *B. nigra* into a homeologous region of chromosome A08 using the *B. rapa* reference genome. SNP markers tightly linked to CR genes were developed, facilitating introgression of the CR genes into canola for use in MAS.

5.2 Mapping of clubroot resistance in two *Brassica nigra* lines by BSR-Seq and development of genetic markers

No resistance gene against clubroot disease has been mapped in *B. nigra*. From this study, the CR genes *Rcr6a* from PI 219576 and *Rcr6b* from BRA192/78 were confirmed to be single dominant genes when the BC_1, F_1 and F_2 populations were used for genetic analysis against pathotype 3 of *P. brassicae*. It is unlikely that multiple genes control the resistance because the 1:1 segregation ratio in BC_1 populations and 3:1 segregation ratio in the F_2 populations confirmed the single gene model for clubroot resistance. Complete resistance in the F_1 population indicated that the resistant parental lines were homozygous dominant. Heterozygosity for resistance in PI

219576 and BRA192/78 would have produced a classic Mendelian segregation ratio (1:1) in the F_1, which was not the case.

Pathotype 3 is the predominant pathotype of *P. brassicae* on canola in Alberta (Strelkov et al., 2006), so it was selected to test the resistant PI 219576 and BRA192/78 populations of *B. nigra* for mapping CR genes. The majority of resistant (PI 219576 and BRA192/78) and susceptible (CR2748) parental plants in the population were either highly resistant or highly susceptible to this pathotype, enabling phenotyping of the host reactions.

RNA-Seq produces a large amount of short DNA sequence reads (Yu et al., 2016) from random places in the transcriptome. Transcriptome analysis and SNP discovery based on RNA-Seq have been conducted in various species, including *B. napus* and *B. rapa* (Bancroft et al., 2011; Paritosh et al., 2013). This study characterized variants in two *B. nigra* populations that carried *Rcr6a* and *Rcr6b*, and identified the most likely candidate gene for clubroot resistance based on DNA variants and development of closely-linked SNP markers to the gene for use in MAS. SNP variants were found to be the most common DNA sequence variation in the current study; similar results were also found in rice, containing an average of 99,955 putative SNPs and 14,617 putative InDels (Takano et al., 2014).

Abundant SNPs and InDels were identified in the *B. nigra* genome using the PSA approach. The SSA method involves the assessment of three biological replicates of resistance and susceptible bulks and provides estimates of variance used to assess the statistical significance of variation of SNPs and InDels; this method is useful for identifying DEGs through RNA-Seq. In this study, a PSA approach was used to merge the data across biological replicates so

statistical analysis was not involved. The PSA method can provide greater depth of sequencing data, providing more accurate results in variant discovery (Yu et al., 2016).

From the RNA-Seq alignment data of PI 219576, there was coverage of 8-fold of the A-reference genome in R samples and 10-fold in S samples. Similarly, there was coverage of 8-fold of the B-reference genome in R samples and 9-fold in S samples. However, there was a lower coverage of 5-fold of the C-reference genome in R samples and 6-fold in S samples. From the RNA-Seq alignment data of BRA 192/78, there was coverage of 6-fold of the A-reference genome in R samples and 10-fold in S samples. Equally, there was coverage of 6-fold of the B-reference genome in R samples and 10-fold in S samples. However, there was a lower coverage of 4-fold of the C-reference genome in R samples and 6-fold in S samples. These results could have contributed to the greater abundance of variants when A- and B- reference genomes were used in the current study. The genome size of *B. oleracea* is largest, then followed by *B. nigra* and *B. rapa*; so coverage of the C-genome was lower.

We genetically mapped *Rcr6a* and *Rcr6b* using the *B. rapa* reference genome. Chromosome A08 had a markedly higher percentage of poly variants in the physical range approximately from 10 Mb to 20 Mb than the rest of the chromosomes in the A-genome. However, the difference in poly variant distribution (%) was not very apparent in chromosome C8 compared to other chromosomes. Chromosomes A08 and B3 showed a significantly higher percentage of polymorphic variants compared to the other chromosomes in both PI 219576 and BRA192/78, whereas C8 exhibited a slightly higher percentage of polymorphic variants. The greater frequency of poly variants could suggest that there was a higher level of difference between the resistant and susceptible bulks in *Rcr6a, Rcr6b* and their adjacent regions compared to other areas of the genome. These results suggest that there are possibly homeologous regions

in A08, B3 and C8 chromosomes. SNP markers were utilized from A08 and B3, but not from C8 because of its slightly higher percentage of polymorphic variants and minor difference in poly variant distribution (%) along the physical distance compared to other chromosomes. Conversely, there were significantly higher percentages of polymorphic variants in A08 and B3 chromosomes.

In addition to the phenotype data, genetic linkage mapping confirmed that one major genetic locus was associated with CR. Closely linked SNP markers to CR genes were developed in the current study using the *B. rapa* and *B. nigra* reference genome information. *Brassica nigra* genome has only been recently sequenced, but has not yet been published. A total of 58 SNP markers were confirmed to be associated with *Rcr6a* and *Rcr6b*. Any plant sample that did not show all '+' for the resistant phenotype or all '–' for susceptible phenotype was considered a recombinant. All of the recombinants were confirmed by re-analyzing the SNP markers using the KASP method. The CR genes were located on the B3 linkage group, in a region homologous to one on *B. rapa* chromosome A08. The markers SNP_A08_47 and SNP_A08_17 flanked the *Rcr6a* gene from PI 219576 to a range of 3.0 cM. In the A08 linkage map from PI 219576, SNP_A08_47 and SNP_A08_17 markers flanked *Rcr6a* in their corresponding upper and lower flanking positions, as shown in the combined genetic map. In the B3 linkage map from PI 219576, the SNP_B3_44 and SNP_B3_51 markers flanked *Rcr6a* in their respective upper and lower flanking positions when compared to the combined genetic map (positions of B3 SNP markers were inverted). The SNP markers B3_62, B3_63, B3_64, B3_65 and B3_67 flanked *Rcr6b* from BRA192/78 to a range of 0.5 cM. Since there were no recombinants between the B3 flanking markers and *Rcr6b*, the markers should have co-segregated with the CR gene; however, there were several undetermined ('0') genotyping results that could have influenced marker

positions. The gap between SNP_A08_56 and *Rcr6b* might be due to the presence of distorted markers, which can affect recombination frequency between markers and skew the genetic distance between markers and the gene of interest (Raman et al., 2014). In addition, the majority of SNP markers were designed from the RNA-Seq data of PI 219576 which could also affect the genetic distance. Other reasons for the wide gap between markers and genes of interest include phenotyping and genotyping errors. Resolution of CR genes can be improved by increasing the mapping population size as well as designing primers closer to *Rcr6a* and *Rcr6b*. SNP_A08_17 co-segregated with *Rcr6b*, indicating that there were no recombinants. The A08 linkage map from BRA192/78 was consistent with the combined genetic map, showing that SNP_A08_47 and SNP_A08_56 markers flanked *Rcr6b*, with SNP_A08_17 co-segregating with the CR gene. The B3 linkage map from BRA192/78 confirmed that the SNP markers B3_65 and B3_67 were located at the lower flanking region of *Rcr6b*, whereas B3_62, B3_63 and B3_64 were located at the upper flanking region. This result was consistent with the combined genetic map, with the positions of B3 SNP markers inverted. The *Rcr6a* and *Rcr6b* genes were confirmed to be linked on the B3 chromosome and were possibly in the same region based on the corresponding positions of the upper flanking marker SNP_A08_47 and the tightly linked marker SNP_A08_17. Based on the relationship between common markers and the location of these CR loci, *Rcr6a* and *Rcr6b* may be identical, allelic or closely linked. To confirm if these two CR genes are in the same position, it will be necessary to clone these genes and compare them at the nucleotide level. Robust SNP markers to CR genes have been developed, which would facilitate the map-based cloning of *Rcr6a* and *Rcr6b*. The markers linked to and nested in *Rcr6a* and *Rcr6b* showed polymorphism between the resistant and susceptible parents, so these genetic markers will be beneficial for MAS in resistance breeding.

The mechanisms for clubroot resistance are not well understood. Within the region defined by SNP_A08_17 and SNP_A08_47 from PI 219576, four TIR-NBS-LRR genes have been located; and within the region defined by SNP_A08_47 and SNP_A08_56 from BRA 192/78, two TIR-NBS-LRR genes have been located (http://brassicadb.org/brad). It is likely that the TNL gene *Bra010552* is a candidate for *Rcr6a* and *Rcr6b.* It is also possible that a combination of the TNL genes are necessary for the expression of *Rcr6a* and *Rcr6b* resistance. Identifying the TNL gene(s) that corresponds with *Rcr6a* and *Rcr6b* will be addressed after the genes have been cloned.

The largest class of resistance genes cloned to date is characterized by a family of proteins consisting of NBS and LRR domains (Chisholm et al., 2006). Nucleotide binding motifs have sequence matches with the NB regions of apoptosis regulators such as *CED4* from *Caenorhabditis elegans* and *Apaf-1* from humans. The LRR is usually 20-30 amino acids in length and these motifs have been identified in proteins from viruses to eukaryotes, participating in activities such as disease resistance. In addition, the NBS-LRR class of R genes can be divided into coiled-coil (CC)-NBS-LRR and TIR-NBS-LRR according to their N-terminal domain (Chisholm et al., 2006). In *Arabidopsis*, more than 150 proteins are expected to be NBS-LRR proteins that determine resistance to bacterial, viral, fungal and oomycete pathogens (Dangl and Jones, 2001). Previous studies have shown that the R genes *CRa* and *Crr1* isolated from *B. rapa* also encode TIR-NBS-LRR proteins (Ueno et al., 2012; Hatakeyama et al., 2013). Thus, it is possible that *Rcr6a*/*Rcr6b* is one or a cluster of TIR-NBS-LRR genes in the mapped region.

The mapping data generated from BSR-Seq was consistent with previous mapping results obtained using SSR markers that corresponded to the microsatellite loci within the *B. juncea* genome (B1-B8). A preliminary linkage map was developed with the pathotype 3 resistance

locus from PI 219576, flanked by the SSR markers sJ4216a and sJ7148a. Three of the four linked SSR markers were selected from chromosome B3 of *B. juncea*, which corresponds to the B3 linkage group in *B. nigra*, verifying that RNA-Seq could provide reliable SNPs for both genome-wide linkage analysis and fine mapping of a specific region. The conventional approach did not provide a reliable genetic map with precise location of CR genes, so MBS through BSR-Seq was carried out. The use of SNP markers also provides the practicality of genotyping thousands of markers rapidly compared to SSR markers (Raman et al., 2014).

Breeding of CR lines using DNA markers in *B. rapa* has been carried out extensively. This approach was based on the notion that CR is controlled by a single dominant locus (Yoshikawa, 1993). Due to the variation in the pathogen, one CR gene is generally not sufficient to protect *Brassica* crops from virulent populations of *P. brassicae* (Hatakeyama et al., 2004). Therefore, breeding of more resistant cultivars would require the addition of more than two CR genes in a single cultivar. Pyramiding CR genes using MAS is an efficient method for this difficult breeding procedure. The SNP markers developed for both *Rcr6a* and *Rcr6b* in *B. nigra* are codominant and their polymorphisms are easily detectable with KASP genotyping. Hence, they will be an effective tool in marker-assisted pyramiding of CR genes in *B. nigra* and other *Brassica* crops.

5.3 Advantages of MBS through BSR-Seq and potential problems

Besides genetic mapping of CR genes, RNA-Seq can provide a robust approach in quantifying gene expression. A BSA-based mapping strategy (BSR-Seq) that depends on RNA-Seq data for SNP marker development has been carried out. Whole genome shotgun (WGS) sequencing is not cost efficient especially for mapping in larger genomes, thus BSR-Seq is an inexpensive and efficient method (Liu et al., 2012). Other advantages include low background

signal and no requirement of a hybridization step, which can result in non-specific binding. However, due to fragmentation throughout the RNA-Seq experiment by either RNA hydrolysis or cDNA shearing, read coverage over genes may be biased against 3' and 5' extremities and can underestimate expression levels of short genes. NGS is an emerging method for SNP genotyping but this method can be time-consuming in terms of informatics needs. Not only did BSR-Seq provide the map positions of two CR genes but also data on differential gene expression. This strategy produced an abundance of polymorphic SNP markers tightly linked to CR genes which can be used for fine mapping and for facilitating gene cloning.

The BSR-Seq approach was used in the current study to successfully map two genes, *Rcr6a* and *Rcr6b*. The size of mapping interval depends on the number of individuals included in the resistant and susceptible bulks, the sequencing depth and the density of polymorphisms in the mapping population. For each factor, more was better. Increasing the number of individuals in each R and S bulks would allow mapping CR genes to a smaller interval. Though the clubroot susceptible plants used to map CR genes were fully recessive and easily differentiated from clubroot resistant plants, classification of R and S individuals was not always straightforward. A phenotypic rating of '0' for resistance and '1' for susceptibility could have been imprecise because some plants may start to develop clubroot symptoms later than 5-week post-inoculation, when the inoculated plants were rated. However, inclusion of a few susceptible individuals in the resistant pool is less likely to significantly affect mapping accuracy. To avoid false-positive SNPs from the RNA-Seq data, we did not select any SNPs that inaccurately displayed a high probability of linkage to the CR gene outside of chromosomes A08 and B3.

The RNA-Seq data provided information on the effects of clubroot disease on global patterns of gene expression. Leaf tissues were used in the study because they were suitable for

mapping work as well as providing gene expression data from a systemic response to *P. brassicae*. For gene expression data, however, it would be best to extract RNA from a tissue in which the mutant phenotype is noticeable; in this case, the root tissues would be ideal. Root tissues can provide gene expression data from a direct response to *P. brassicae*, but collecting roots for each plant is time-consuming and may potentially damage the roots during the collection process. This could interfere with defence mechanisms and disrupt plant growth. Whether to collect leaf or root tissues does not seriously influence the genetic mapping aspect of BSR-Seq, because SNPs in all genes located close to the CR gene and those SNPs that are expressed in these samples can be used as markers for genetic mapping (Liu et al., 2012).

Of the 382 differentially expressed genes from BRA192/78, 301 were down-regulated and 81 were up-regulated in the resistant bulk relative to the susceptible bulk. Of the 11 differentially expressed genes from PI 219576, 8 were up-regulated and 3 were down-regulated in the R bulk relative to the S bulk. There was a large difference between the numbers of DEGs identified between PI 219576 and BRA192/78. The difference might have been due to the timing for leaf tissue collection at 2-week post-inoculation. For PI 219576, plant defence response to *P. brassicae* could have occurred much earlier or later; but for BRA192/78, plant defence response may have occurred exactly at the 2-week post-inoculation period. The up- and down-regulated fold changes were not an accurate measurement because the total raw count could have been very low, but the corresponding fold changes were much higher. For this reason, interpretation of differential gene expression was analyzed carefully. Replicated RNA-Seq data can provide better accuracy in identifying DEGs, but the lack of replication would not affect mapping results (Liu et al., 2012).

BSR-Seq relies on access to an available reference genome and this approach is affected by the quality of the reference genome and the degree of structural variation within the species of interest (Liu et al., 2012). The fully sequenced *B. rapa* reference genome was used for differential gene expression analysis because alignment of RNA-Seq reads with *B. rapa* reference generated greater numbers of DEGs for both PI 219576 and BRA192/78, in comparison to alignment with *B. nigra* reference. Conversely, both *B. nigra* and *B. rapa* reference genomes were used for SNP discovery and abundant SNPs were identified. Mis-assemblies in the reference genome and copy number variation (CNV) in *B. nigra* and the reference genome may hinder mapping success. Even though the quality of *B. nigra* reference genome was not completely refined, two CR genes were successfully mapped using BSR-Seq.

5.4 Transcriptome analysis and annotation of DEGs

Transcriptome analysis was carried out to better understand the molecular mechanisms of disease resistance; thus, DEGs were annotated to determine the function of genes associated with CR. Previous research has been done on characterizing several molecular components involved in disease development, particularly the role of cytokinins on *A. thaliana* and *B. juncea* (Evans and Scholes, 1995; Luo et al., 2013). In this study, 11 DEGs from PI 219576 and 382 DEGs from BRA192/78 were identified between inoculated resistant and susceptible plants. Gene annotation was performed using the *Brassica* database and 69 DEGs retrieved no hit in the annotation search, indicating that a number of unknown genes were expressed in inoculated resistant plants. Several cellular activities in inoculated resistant plants were shown to be highly up-regulated at 2-week post-inoculation, which is consistent with previous studies suggesting that root infection and pathogen colonization occur at 15 dpi (Chu et al., 2014). Differential gene expression

analysis indicated that infection may have activated or deactivated specific genes in the host that may influence clubroot development on resistant and susceptible plants.

A number of CR mechanisms reported previously were supported in this current study. A gene involved in plant lipid transfer (*Bra000775*) was significantly up-regulated, suggesting that lipids play an important role in inoculated resistant plants. This was consistent with a previous study indicating that lipids play a critical role in detecting infection and could regulate gene transcription (Walley et al., 2013). Infection by *P. brassicae* may cause resistant plants to be able to mobilize lipid biosynthesis and metabolism (Chu et al., 2014). An ethylene-responsive protein gene (*Bra035761*) was up-regulated, which has been shown to prevent clubroot development in *A. thaliana* and may be involved in induced resistance mediated by biofungicides (Knaust and Ludwig-Muller, 2013; Lahlali et al., 2013). A previous report suggested that jasmonic acid and ethylene may function together in plant defence (Kazan and Manners, 2012); however, no altered expression of genes associated with JA biosynthesis was identified in this current study. A plant cell wall-associated kinase gene (*Bra008381*) was found to be up-regulated, signifying that reinforcement of plant cell wall may be essential in disease resistance. This may cause a change in the structure and composition of the cell wall, which could prevent secondary infection in epidermal and cortical cells (Chu et al., 2014). Up-regulated DEGs involved in callose localization and deposition have been shown to function in resistance against fungal pathogen penetration of plant cell walls (Naumann et al., 2013).

Several DEGs associated with TIR-NBS-LRR class of disease resistance proteins were identified such as *Bra010551* and *Bra002117*. This result indicated that *Rcr6a* and *Rcr6b* may encode a TIR-NBS-LRR protein. A known defense signaling pathway related to this class of resistant proteins involves the biosynthesis of SA and PR proteins, along with a hypersensitive

reaction (HR) (Joshi and Nayak, 2011). SA biosynthesis was known to play a role in host resistance against biotrophic pathogens (Thaler et al., 2012). However, in this current study, there was no altered expression of genes associated with SA signalling and biosynthesis. Lack of significant change in transcription of genes involved in SA biosynthesis has been reported previously for clubroot resistance (Lahlali et al., 2013). Several down-regulated DEGs (*Bra023748, Bra023404, Bra033138* and *Bra017612*) were identified as PR transcriptional factors, yet there is no evidence to link any of these genes to SA- or HR-related defence responses. The down-regulation of PR transcriptional factor genes may play a role in clubroot resistance because certain metabolic processes are suppressed in the host that are necessary for pathogenesis. It is uncertain whether genes involved in HR or SA signaling pathways contribute to clubroot resistance.

Past studies have shown that auxin levels increased in roots during secondary infection by *P. brassicae* (Evans and Scholes, 1995). All of the auxin-responsive genes (*Bra025770, Bra025857, Bra012231, Bra003665, Bra002327, Bra008836* and *Bra009535*) identified in this study were down-regulated in resistant plants relative to susceptible plants, suggesting that pathogen-activated auxin synthesis may have been suppressed. Chu et al. (2014) found that a gene encoding the SAUR (Small Auxin-Up RNA) family of proteins was highly induced in the resistant plants relative to susceptible plants, and these proteins act as a negative regulator of auxin biosynthesis and transport (Kant et al., 2009). Additionally, the GH3 family protein has been linked with increased basal immunity by disrupting pathogen-activated auxin synthesis (Fu et al., 2011). The current study also revealed the down-regulation of an IAA-inducible gene (*Bra026166*) in resistant plants, which was previously reported to play a role in clubroot development (Piao et al., 2009).

The down-regulated DEGs related to cell development and morphogenesis identified in this study included *Bra032981* and *Bra020254*. Down-regulation of these genes related to cell growth may be involved in clubroot resistance facilitated by *Rcr6a* and *Rcr6b*. The development of clubroot symptom is due to hypertrophy, which can disrupt normal physiological events (Kageyama and Asano, 2009). A study on *A. thaliana* showed that cell enlargement and proliferation were inhibited in a partially resistant line to *P. brassicae* (Jubault et al., 2013). Chu et al. (2014) revealed that genes related to "uni-dimensional cell growth" were up-regulated in inoculated susceptible plants. It is likely that down-regulation of these DEGs function with up-regulation of defense-related genes for CR in plants with *Rcr6a* and *Rcr6b.* It is not known if down-regulation of DEGs related to growth and development is linked to suppression of auxin synthesis (Chu et al., 2014).

Many of the DEGs analyzed were related to transcription factors (TFs). For example, a WRKY DNA-binding protein gene (*Bra011861*) was significantly up-regulated. WRKY-family TFs were reported to be involved in disease resistance through effector-triggered immunity (Ishihama and Yoshioka, 2012). A MYB-family transcription factor gene (*Bra009228*) was up-regulated, and this gene may be involved in secondary cell-wall biosynthesis (Wang and Dixon, 2012) which can result in the strengthening of host cell walls. Several other TFs were also identified such as the down-regulation of heat-stress transcription factor gene (*Bra007739*), but their involvement in CR is not well understood.

6. CONCLUSIONS AND FUTURE RESEARCH

Single dominant CR alleles *Rcr6a* and *Rcr6b* against pathotype 3 of *P. brassicae* were mapped to the B3 chromosome of *B. nigra* using closely linked SNP markers. Flanking markers were tightly linked to *Rcr6a* in PI 219576 and *Rcr6b* in BRA192/78. *Rcr6a* was flanked by SNP_A08_47 and SNP_A08_17 markers to a range of 3.0 cM, with a physical distance of approximately 0.4 Mb. *Rcr6b* was flanked by SNP_B3_65, SNP_B3_67, SNP_B3_62, SNP_B3_63 and SNP_B3_64 markers to a range of 0.5 cM, with a physical distance of approximately 0.3 Mb. Development of more robust SNP markers and increasing the population size can fine map the CR genes. It is unclear whether *Rcr6a* and *Rcr6b* are identical, allelic or closely linked; hence both genes will be cloned and functionally characterized. Future testing with the newly discovered pathotype 5X on a large population size will be used to assess if CR is controlled by a single dominant gene linked on the B3 chromosome. Fine mapping will be used to determine if the CR genes against pathotypes 3 and 5X are located in the same region.

Genetic resistance is the foundation for management of clubroot on *Brassica* species and pyramiding of CR genes can be beneficial for long-term clubroot management in western Canada where it is a serious threat to canola production. In this study, genetic mapping identified two CR genes (*Rcr6a* and *Rcr6b*) in *B. nigra*. Tightly linked SNP markers were developed for use in MAS in canola and mustard breeding programs. Differential gene expression was analyzed to provide insight into the molecular mechanisms of clubroot resistance associated with *Rcr6a* and *Rcr6b*. Gene expression data can be utilized for designing a breeding strategy based on the different actions of CR genes to improve the durability and effectiveness of clubroot resistance. Several DEGs identified from RNA-Seq data were annotated as TIR-NBS-LRR class of disease resistance proteins (*Bra010551* and *Bra002117*) and pathogenesis-related transcriptional factors

(*Bra023748, Bra023404, Bra033138*, and *Bra017612*). Several auxin-responsive genes (*Bra009535*, *Bra025770*, *Bra025857*, *Bra012231*, *Bra003665*, *Bra002327*, and *Bra008836*) and a gene associated with cell growth/development (*Bra032981*) were down-regulated in resistant plants. Proteins involved in biosynthetic and signalling pathways of JA/ET have been identified previously for resistance to clubroot. An ethylene-responsive protein gene (*Bra035761*) was up-regulated, but no genes involved in JA biosynthesis showed altered expressions during infection processes in the current study. Future research will examine the roles of plant hormones and other signalling molecules in clubroot development, and identify candidate transcription factors that are involved in regulating clubroot resistant responses conferred by *Rcr6a*, *Rcr6b* and other CR genes.

7. REFERENCES

Attia, T., Busso, C., & Röbbelen, G. (1987). Digenomic triploids for an assessment of chromosome relationships in the cultivated diploid *Brassica* species. *Genome*, *29*(2), 326-330.

Bancroft, I., Morgan, C., Fraser, F., Higgins, J., Wells, R., Clissold, L., Baker, D., Long, Y., Meng, J., Wang, X., Liu, S., & Trick, M. (2011). Dissecting the genome of the polyploid crop oilseed rape by transcriptome sequencing. *Nature Biotechnology*, *29*(8), 762-766.

Branca, F., & Cartea, E. (2011). *Brassica*. *Wild Crop Relatives: Genomic and Breeding Resources*, 17-36.

Brassica Marker Access and Search Tool. (2016). Retrieved from http://aafc-aac.usask.ca/BrassicaMAST.

Canola Council of Canada. (2014). Retrieved from http://www.canolacouncil.org/uploads/canola_growing_great_2015.pdf.

Cao, T., Manolii, V. P., Strelkov, S. E., Hwang, S. F., & Howard, R. J. (2009). Virulence and spread of *Plasmodiophora brassicae* [clubroot] in Alberta, Canada. *Canadian Journal of Plant Pathology*, *31*(3), 321-329.

Cao, T., Tewari, J., & Strelkov, S. E. (2007). Molecular detection of *Plasmodiophora brassicae*, causal agent of clubroot of crucifers, in plant and soil. *Plant Disease*, *91*(1), 80-87.

Cavalier-Smith, T., & Chao, E. E. Y. (2003). Phylogeny and classification of phylum Cercozoa (Protozoa). *Protist*, *154*(3), 341-358.

Cavell, A. C., Lydiate, D. J., Parkin, I. A. P., Dean, C., & Trick, M. (1998). Collinearity between a 30-centimorgan segment of *Arabidopsis thaliana* chromosome 4 and duplicated regions within the *Brassica napus* genome. *Genome*, *41*(1), 62-69.

Chauhan, J. S., Singh, K. H., Singh, V. V., & Kumar, S. (2011). Hundred years of rapeseed-mustard breeding in India: accomplishments and future strategies. *The Indian Journal of Agricultural Sciences*, *81*(12), 1093-1109.

Chen, J., Pang, W., Chen, B., Zhang, C., & Piao, Z. (2016). Transcriptome analysis of *Brassica rapa* near-isogenic lines carrying clubroot resistant and susceptible alleles in response to *Plasmodiophora brassicae* during early infection. *Frontiers in Plant Science*, *6*, 1183-1196.

Chevre, A. M., Eber, F., Barret, P., Dupuy, P., & Brace, J. (1997). Identification of the different *Brassica nigra* chromosomes from both sets of *B. oleracea-B. nigra* and *B. napus-B. nigra* addition lines with a special emphasis on chromosome transmission and self-incompatibility. *Theoretical and Applied Genetics*, *94*(5), 603-611.

Chiang, M. S., & Crete, R. (1972). Screening crucifers for germplasm resistance to clubroot *Plasmodiophora brassicae*. *Canadian Plant Disease Survey*, *52*(2), 45-50.

Chiang, M. S., Chiang, B. Y., & Grant, W. F. (1977). Transfer of resistance to race 2 of *Plasmodiophora brassicae* from *Brassica napus* to cabbage (*B. oleracea* var. capitata). I. Interspecific hybridization between *B. napus* and *B. oleracea* var. capitata. *Euphytica*, *26*(2), 319-336.

Chisholm, S. T., Coaker, G., Day, B., & Staskawicz, B. J. (2006). Host-microbe interactions: shaping the evolution of the plant immune response. *Cell*, *124*(4), 803-814.

Christey, M. C. (2004). *Brassica* protoplast culture and somatic hybridization. *Brassica*, 119-148.

Chu, M., Song, T., Falk, K. C., Zhang, X., Liu, X., Chang, A., Lahlali, R., McGregor, L., Gossen, B. D., Yu, F., & Peng, G. (2014). Fine mapping of *Rcr1* and analyses of its effect on transcriptome patterns during infection by *Plasmodiophora brassicae*. *BMC Genomics*, *15*(1), 1166-1185.

Chu, M., Yu, F., Falk, K. C., Liu, X., Zhang, X., Chang, A., & Peng, G. (2013). Identification of the clubroot resistance gene *Rpb1* and introgression of the resistance into canola breeding lines using a marker-assisted approach. *Acta Horticulturae*, *1005*, 599-605.

Crisp, P., Crute, I. R., Sutherland, R. A., Angell, S. M., Bloor, K., Burgess, H., & Gordon, P. L. (1989). The exploitation of genetic resources of *Brassica oleracea* in breeding for resistance to clubroot (*Plasmodiophora brassicae*). *Euphytica*, *42*(3), 215-226.

Dangl, J. L., & Jones, J. D. (2001). Plant pathogens and integrated defence responses to infection. *Nature*, *411*(6839), 826-833.

Deora, A., Gossen, B. D., Amirsadeghi, S., & McDonald, M. R. (2015). A multiplex qPCR assay for detection and quantification of *Plasmodiophora brassicae* in soil. *Plant Disease*, *99*(7), 1002-1009.

Deschamps, S., Llaca, V., & May, G. D. (2012). Genotyping-by-sequencing in plants. *Biology*, *1*(3), 460-483.

Devos, S., Laukens, K., Deckers, P., Van Der Straeten, D., Beeckman, T., Inzé, D., Van Onckelen, H., Witters, E., & Prinsen, E. (2006). A hormone and proteome approach to picturing the initial metabolic events during *Plasmodiophora brassicae* infection on *Arabidopsis*. *Molecular Plant-Microbe Interactions*, *19*(12), 1431-1443.

Diederichsen, E., Frauen, M., Linders, E. G., Hatakeyama, K., & Hirai, M. (2009). Status and perspectives of clubroot resistance breeding in crucifer crops. *Journal of Plant Growth Regulation*, *28*(3), 265-281.

Diederichsen, E., & Sacristan, M. D. (1996). Disease response of resynthesized *Brassica napus* L. lines carrying different combinations of resistance to *Plasmodiophora brassicae* Wor. *Plant Breeding*, *115*(1), 5-10.

Diederichsen, E., Wagenblatt, B., Schallehn, V., Deppe, U., & Sacristán, M. (1994). Transfer of clubroot resistance from resynthesised *Brassica napus* into oilseed rape-identification of race-specific interactions with *Plasmodiophora brassicae*. *ISHS Brassica Symposium-IX Crucifer Genetics Workshop*, *407*, 423-430.

Dixon, G. R. (2004). The biology of *Plasmodiophora brassicae* Wor.-A review of recent advances. *IV International Symposium on Brassicas and XIV Crucifer Genetics Workshop*, *706*, 271-282.

Dixon, G. R. (2009). The occurrence and economic impact of *Plasmodiophora brassicae* and clubroot disease. *Journal of Plant Growth Regulation*, *28*(3), 194-202.

Dixon, G. R., Doodson, J. K., Beeney, B. W., Davies, H., & Moxon, R. H. (1972). Studies of resistance in swede seed stocks to clubroot (*Plasmodiophora brassicae*). *Plant Varieties Seeds*, *12*, 456-463.

Dokken-Bouchard, F. L., Bouchard, A. J., Ippolito, J., Peng, G., Strelkov, S., Kirkham, C. L., & Kutcher, H. R. (2010). Detection of *Plasmodiophora brassicae* in Saskatchewan, 2008. *Canadian Plant Disease Survey*, *90*, 126.

Donald, C., & Porter, I. (2009). Integrated control of clubroot. *Journal of Plant Growth Regulation*, *28*(3), 289-303.

Evans, J. L., & Scholes, J. D. (1995). How does clubroot alter the regulation of carbon metabolism in its host? *Aspects of Applied Biology*, *42*, 125-132.

Faggian, R., & Strelkov, S. E. (2009). Detection and measurement of *Plasmodiophora brassicae*. *Journal of Plant Growth Regulation*, *28*(3), 282-288.

Frauen, M. (1999). A new clubroot resistant variety in winter oilseed rape. *Proceedings of the 10th International Rapeseed Congress, Canberra, Australia*.

Fu, J., Liu, H., Li, Y., Yu, H., Li, X., Xiao, J., & Wang, S. (2011). Manipulating broad-spectrum disease resistance by suppressing pathogen-induced auxin accumulation in rice. *Plant Physiology*, *155*(1), 589-602.

Gerdemann-Knörck, M., Nielen, S., Tzscheetzsch, C., Iglisch, J., & Schieder, O. (1995). Transfer of disease resistance within the genus *Brassica* through asymmetric somatic hybridization. *Euphytica*, *85*(1), 247-253.

Glimelius, K. (1999). Somatic hybridization. *Biology of Brassica Coenospecies*, 107-148.

Gossen, B. D., Adhikari, K. K. C., & McDonald, M. R. (2012). Effects of temperature on infection and subsequent development of clubroot under controlled conditions. *Plant Pathology*, *61*(3), 593-599.

Gossen, B. D., Deora, A. D., Peng, G., Hwang, S. F., & McDonald, M. R. (2014). Effect of environmental parameters on clubroot development and the risk of pathogen spread. *Canadian Journal of Plant Pathology*, *36*(S1), 37-48.

Grist, S. A., Firgaira, F. A., & Morley, A. A. (1993). Dinucleotide repeat polymorphisms isolated by the polymerase chain reaction. *BioTechniques*, *15*(2), 304-309.

Gustafsson, M., & Falt, A. S. (1986). Genetic studies on resistance to clubroot in *Brassica napus*. *Annals of Applied Biology*, *108*(2), 409-415.

Hasan, M. J., Strelkov, S. E., Howard, R. J., & Rahman, H. (2012). Screening of *Brassica* germplasm for resistance to *Plasmodiophora brassicae* pathotypes prevalent in Canada for broadening diversity in clubroot resistance. *Canadian Journal of Plant Science*, *92*(3), 501-515.

Hatakeyama, K., Fujimura, M., Ishida, M., & Suzuki, T. (2004). New classification method for *Plasmodiophora brassicae* field isolates in Japan based on resistance of F_1 cultivars of Chinese cabbage (*Brassica rapa* L.) to clubroot. *Breeding Science*, *54*(2), 197-201.

Hatakeyama, K., Suwabe, K., Tomita, R. N., Kato, T., Nunome, T., Fukuoka, H., & Matsumoto, S. (2013). Identification and characterization of *Crr1a*, a gene for resistance to clubroot disease (*Plasmodiophora brassicae* Woronin) in *Brassica rapa* L. *PloS One*, *8*(1), e54745-e54754.

Howard, R. J., Strelkov, S. E., & Harding, M. W. (2010). Clubroot of cruciferous crops–new perspectives on an old disease. *Canadian Journal of Plant Pathology*, *32*(1), 43-57.

Hwang, S. F., Strelkov, S. E., Feng, J. I. E., Gossen, B. D., & Howard, R. J. (2011). *Plasmodiophora brassicae*: a review of an emerging pathogen of the Canadian canola (*Brassica napus*) crop. *Molecular Plant Pathology*, *13*(2), 105-113.

Ishihama, N., & Yoshioka, H. (2012). Post-translational regulation of WRKY transcription factors in plant immunity. *Current Opinion in Plant Biology*, *15*(4), 431-437.

Jahangir, M., Kim, H. K., Choi, Y. H., & Verpoorte, R. (2009). Health-affecting compounds in Brassicaceae. *Comprehensive Reviews in Food Science and Food Safety*, *8*(2), 31-43.

Johnston, J. S., Pepper, A. E., Hall, A. E., Chen, Z. J., Hodnett, G., Drabek, J., Lopez, R., & Price, H. J. (2005). Evolution of genome size in Brassicaceae. *Annals of Botany*, *95*(1), 229-235.

Joshi, R. K., & Nayak, S. (2011). Functional characterization and signal transduction ability of nucleotide-binding site-leucine-rich repeat resistance genes in plants. *Genetic Molecular Research*, *10*(4), 2637-2652.

Jubault, M., Hamon, C., Gravot, A., Lariagon, C., Delourme, R., Bouchereau, A., & Manzanares-Dauleux, M. J. (2008). Differential regulation of root arginine catabolism and polyamine metabolism in clubroot-susceptible and partially resistant *Arabidopsis* genotypes. *Plant Physiology*, *146*(4), 2008-2019.

Jubault, M., Lariagon, C., Taconnat, L., Renou, J. P., Gravot, A., Delourme, R., & Manzanares-Dauleux, M. J. (2013). Partial resistance to clubroot in *Arabidopsis* is based on changes in the host primary metabolism and targeted cell division and expansion capacity. *Functional & Integrative Genomics*, *13*(2), 191-205.

Kageyama, K., & Asano, T. (2009). Life cycle of *Plasmodiophora brassicae*. *Journal of Plant Growth Regulation*, *28*(3), 203-211.

Kant, S., Bi, Y. M., Zhu, T., & Rothstein, S. J. (2009). SAUR39, a small auxin-up RNA gene, acts as a negative regulator of auxin synthesis and transport in rice. *Plant Physiology*, *151*(2), 691-701.

Kazan, K., & Manners, J. M. (2012). JAZ repressors and the orchestration of phytohormone crosstalk. *Trends in Plant Science*, *17*(1), 22-31.

Kircher, M., & Kelso, J. (2010). High-throughput DNA sequencing–concepts and limitations. *Bioessays*, *32*(6), 524-536.

Knaust, A., & Ludwig-Müller, J. (2013). The ethylene signaling pathway is needed to restrict root gall growth in *Arabidopsis* after infection with the obligate biotrophic protist *Plasmodiophora brassicae*. *Journal of Plant Growth Regulation*, *32*(1), 9-21.

Lahlali, R., Peng, G., Gossen, B. D., McGregor, L., Yu, F. Q., Hynes, R. K., Hwang, S. F., McDonald M. R., & Boyetchko, S. M. (2013). Evidence that the biofungicide Serenade (*Bacillus subtilis*) suppresses clubroot on canola via antibiosis and induced host resistance. *Phytopathology*, *103*(3), 245-254.

Lan, T. H., & Paterson, A. H. (2000). Comparative mapping of quantitative trait loci sculpting the curd of *Brassica oleracea*. *Genetics*, *155*(4), 1927-1954.

LeBoldus, J. M., Manolii, V. P., Turkington, T. K., & Strelkov, S. E. (2012). Adaptation to *Brassica* host genotypes by a single-spore isolate and population of *Plasmodiophora brassicae* (clubroot). *Plant Disease*, *96*(6), 833-838.

Leflon, M., Eber, F., Letanneur, J. C., Chelysheva, L., Coriton, O., Huteau, V., Ryder, C.D., Barker, G., Jenczewski, E., & Chevre, A. M. (2006). Pairing and recombination at meiosis of *Brassica rapa* (AA) × *Brassica napus* (AACC) hybrids. *Theoretical and Applied Genetics*, *113*(8), 1467-1480.

Li, G., & Quiros, C. F. (2003). In planta side-chain glucosinolate modification in *Arabidopsis* by introduction of dioxygenase *Brassica* homolog BoGSL-ALK. *Theoretical and Applied Genetics*, *106*(6), 1116-1121.

Liu, S., Yeh, C. T., Tang, H. M., Nettleton, D., & Schnable, P. S. (2012). Gene mapping via bulked segregant RNA-Seq (BSR-Seq). *PLoS One*, *7*(5), e36406-e36413.

Lovelock, D. A., Donald, C. E., Conlan, X. A., & Cahill, D. M. (2013). Salicylic acid suppression of clubroot in broccoli (*Brassicae oleracea* var. *italica*) caused by the obligate biotroph *Plasmodiophora brassicae*. *Australasian Plant Pathology*, *42*(2), 141-153.

Ludwig-Müller, J. (2009). Glucosinolates and the clubroot disease: defense compounds or auxin precursors? *Phytochemistry Reviews*, *8*(1), 135-148.

Ludwig-Müller, J., Prinsen, E., Rolfe, S. A., & Scholes, J. D. (2009). Metabolism and plant hormone action during clubroot disease. *Journal of Plant Growth Regulation*, *28*(3), 229-244.

Luo, Y., Yin, Y., Liu, Y., Pei, X., Wang, Z., Fan, Y., Wang, X., & Leng, R. (2013). Identification of differentially expressed genes in *Brassica juncea* var. *tumida* Tsen following infection by *Plasmodiophora brassicae*. *European Journal of Plant Pathology*, *137*(1), 43-53.

Manitoba Agriculture, Food and Rural Development. (2016). Retrieved from https://www.gov.mb.ca/agriculture/crops/plant-diseases/clubroot-distribution-in-manitoba.html.

Manzanares-Dauleux, M. J., Delourme, R., Baron, F., & Thomas, G. (2000). Mapping of one major gene and of QTLs involved in resistance to clubroot in *Brassica napus*. *Theoretical and Applied Genetics*, *101*(5), 885-891.

Michelmore, R. W., Paran, I., & Kesseli, R. V. (1991). Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. *Proceedings of the National Academy of Sciences*, *88*(21), 9828-9832.

Nagaharu, U. (1935). Genome analysis in *Brassica* with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. *Journal of Japanese Botany*, *7*, 389-452.

Nagaoka, T., Doullah, M. A. U., Matsumoto, S., Kawasaki, S., Ishikawa, T., Hori, H., & Okazaki, K. (2010). Identification of QTLs that control clubroot resistance in *Brassica oleracea* and comparative analysis of clubroot resistance genes between *B. rapa* and *B. oleracea*. *Theoretical and Applied Genetics*, *120*(7), 1335-1346.

Naiki, T., & Dixon, G. R. (1987). The effects of chemicals on developmental stages of *Plasmodiophora brassicae* (clubroot). *Plant Pathology*, *36*(3), 316-327.

Naumann, M., Somerville, S. C., & Voigt, C. A. (2013). Differences in early callose deposition during adapted and non-adapted powdery mildew infection of resistant *Arabidopsis* lines. *Plant Signaling & Behavior*, *8*(6), e24408-e24411.

Navabi, Z. K., Huebert, T., Sharpe, A. G., O'Neill, C. M., Bancroft, I., & Parkin, I. A. (2013). Conserved microstructure of the *Brassica* B Genome of *Brassica nigra* in relation to homologous regions of *Arabidopsis thaliana*, *B. rapa* and *B. oleracea*. *BMC Genomics*, *14*(1), 250-264.

Navabi, Z. K., Parkin, I. A. P., Pires, J. C., Xiong, Z., Thiagarajah, M. R., Good, A. G., & Rahman, M. H. (2010). Introgression of B-genome chromosomes in a doubled haploid population of *Brassica napus* × *B. carinata*. *Genome*, *53*(8), 619-629.

Padilla, G., Cartea, M. E., Rodríguez, V. M., & Ordás, A. (2005). Genetic diversity in a germplasm collection of *Brassica rapa* subsp *rapa* L. from northwestern Spain. *Euphytica*, *145*(1-2), 171-180.

Palm, E. T. (1963). Effect of mineral nutrition on the invasion and response of turnip tissue to *Plasmodiophora brassicae* Woronin. *Contribution from Boyce Thompson Institute*, *22*, 91-112.

Panjabi, P., Jagannath, A., Bisht, N. C., Padmaja, K. L., Sharma, S., Gupta, V., Pradhan, A. K., & Pental, D. (2008). Comparative mapping of *Brassica juncea* and *Arabidopsis thaliana* using Intron Polymorphism (IP) markers: homoeologous relationships, diversification and evolution of the A, B and C *Brassica* genomes. *BMC Genomics*, *9*(1), 113-131.

Paritosh, K., Gupta, V., Yadava, S. K., Singh, P., Pradhan, A. K., & Pental, D. (2014). RNA-seq based SNPs for mapping in *Brassica juncea* (AABB): synteny analysis between the two constituent genomes A (from *B. rapa*) and B (from *B. nigra*) shows highly divergent gene block arrangement and unique block fragmentation patterns. *BMC Genomics*, *15*(1), 396-409.

Paritosh, K., Yadava, S. K., Gupta, V., Panjabi-Massand, P., Sodhi, Y. S., Pradhan, A. K., & Pental, D. (2013). RNA-seq based SNPs in some agronomically important oleiferous lines of *Brassica rapa* and their use for genome-wide linkage mapping and specific-region fine mapping. *BMC Genomics*, *14*(1), 463-475.

Päsold, S., & Ludwig-Müller, J. (2013). Reduction of clubroot (*Plasmodiophora brassicae*) formation in *Arabidopsis thaliana* after treatment with prohexadione-calcium, an inhibitor of oxoglutaric acid-dependent dioxygenases. *Plant Pathology*, *62*(6), 1357-1365.

Peng, G., McGregor, L., Lahlali, R., Gossen, B. D., Hwang, S. F., Adhikari, K. K., Strelkov, S. E., & McDonald, M. R. (2011a). Potential biological control of clubroot on canola and crucifer vegetable crops. *Plant Pathology*, *60*(3), 566-574.

Peng G., Falk, K. C., James, B., & Grugel, R. K. (2011b). Evaluation of *Brassica* germplasm for resistance to clubroot of canola. *Canadian Journal of Plant Pathology, 33*(2), 261.

Peng, G., Falk, K. C., Gugel, R. K., Franke, C., Yu, F., James, B., Strelkov S. E., Hwang S. F., & McGregor, L. (2014). Sources of resistance to *Plasmodiophora brassicae* (clubroot) pathotypes virulent on canola. *Canadian Journal of Plant Pathology, 36*(1), 89-99.

Peng, G., Pageau, D., Strelkov, S. E., Gossen, B. D., Hwang, S. F., & Lahlali, R. (2015). A >2-year crop rotation reduces resting spores of *Plasmodiophora brassicae* in soil and the impact of clubroot on canola. *European Journal of Agronomy, 70*, 78-84.

Piao, Z., Ramchiary, N., & Lim, Y. P. (2009). Genetics of clubroot resistance in *Brassica* species. *Journal of Plant Growth Regulation, 28*(3), 252-264.

Podsędek, A. (2007). Natural antioxidants and antioxidant capacity of *Brassica* vegetables: A review. *LWT-Food Science and Technology, 40*(1), 1-11.

Prost, I., Dhondt, S., Rothe, G., Vicente, J., Rodriguez, M. J., Kift, N., Carbonne, F., Griffiths, G., Esquerre-Tugaye, M. T., Rosahl, S., & Castresana, C. (2005). Evaluation of the antimicrobial activities of plant oxylipins supports their involvement in defense against pathogens. *Plant Physiology, 139*(4), 1902-1913.

Rahman, H., Peng, G., Yu, F., Falk, K. C., Kulkarni, M., & Selvaraj, G. (2014). Genetics and breeding for clubroot resistance in Canadian spring canola (*Brassica napus* L.). *Canadian Journal of Plant Pathology, 36*, 122-134.

Rakow, G. (2007). Rapeseed genetics and breeding research for sustainable oilseed production. *Proceedings of the 12th International Rapeseed Congress, 1*, 207-209.

Raman, H., Dalton-Morgan, J., Diffey, S., Raman, R., Alamery, S., Edwards, D., & Batley, J. (2014). SNP markers-based map construction and genome-wide linkage analysis in *Brassica napus*. *Plant Biotechnology Journal, 12*(7), 851-860.

Rocherieux, J., Glory, P., Giboulot, A., Boury, S., Barbeyron, G., Thomas, G., & Manzanares-Dauleux, M. J. (2004). Isolate-specific and broad-spectrum QTLs are involved in the control of clubroot in *Brassica oleracea*. *Theoretical and Applied Genetics, 108*(8), 1555-1563.

Saito, M., Kubo, N., Matsumoto, S., Suwabe, K., Tsukada, M., & Hirai, M. (2006). Fine mapping of the clubroot resistance gene, *Crr3*, in *Brassica rapa*. *Theoretical and Applied Genetics, 114*(1), 81-91.

Sakamoto, K., Saito, A., Hayashida, N., Taguchi, G., & Matsumoto, E. (2008). Mapping of isolate-specific QTLs for clubroot resistance in Chinese cabbage (*Brassica rapa* L. ssp. *pekinensis*). *Theoretical and Applied Genetics, 117*(5), 759-767.

Sanger, F., Coulson, A. R., Friedmann, T., Air, G. M., Barrell, B. G., Brown, N. L., Fiddes J. C., Hutchison C. A., Slocombe P. M., & Smith, M. (1978). The nucleotide sequence of bacteriophage φX174. *Journal of Molecular Biology*, *125*(2), 225-246.

Schelfhout, C. J., Snowdon, R., Cowling, W. A., & Wroth, J. M. (2006). Tracing B-genome chromatin in *Brassica napus* × *B. juncea* interspecific progeny. *Genome*, *49*(11), 1490-1497.

Schlötterer, C. (2004). The evolution of molecular markers-just a matter of fashion? *Nature Reviews Genetics*, *5*(1), 63-69.

Semagn, K., Babu, R., Hearne, S., & Olsen, M. (2014). Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): overview of the technology and its application in crop improvement. *Molecular Breeding*, *33*(1), 1-14.

Sharma, A., Li, X., & Lim, Y. P. (2014). Comparative genomics of Brassicaceae crops. *Breeding Science*, *64*(1), 3-13.

Sharma, K., Gossen, B. D., & McDonald, M. R. (2011). Effect of temperature on primary infection by *Plasmodiophora brassicae* and initiation of clubroot symptoms. *Plant Pathology*, *60*(5), 830-838.

Shendure, J., & Ji, H. (2008). Next-generation DNA sequencing. *Nature Biotechnology*, *26*(10), 1135-1145.

Siemens, J., Keller, I., Sarx, J., Kunz, S., Schuller, A., Nagel, W., Schmulling, T., Parniske, M., & Ludwig-Müller, J. (2006). Transcriptome analysis of *Arabidopsis* clubroots indicate a key role for cytokinins in disease development. *Molecular Plant-Microbe Interactions*, *19*(5), 480-494.

Snowdon, R. J., & Friedt, W. (2004). Molecular markers in *Brassica* oilseed breeding: current status and future possibilities. *Plant Breeding*, *123*(1), 1-8.

Sokal, R. R., & Rohlf, F. J. (1995). The principles and practice of statistics in biological research. *New York: Edition*, *3*.

Statista. (2013). Retrieved from http://www.statista.com.

Strelkov, S. E., Hwang, S. F., Manolii, V. P., Cao, T., & Feindel, D. (2016a). Emergence of new virulence phenotypes of *Plasmodiophora brassicae* on canola (*Brassica napus*) in Alberta, Canada. *European Journal of Plant Pathology*, doi:10.1007/s10658-016-0888-8.

Strelkov, S. E., Manolii, V. P., Harding, M. W., Sieusahai, G., Hwang, S. F., Shi, Y., Lawson, A., Fisher, J., Cornelsen, J., Daniels, G. C., & Feindel, D. (2016b). Occurrence and spread of clubroot on canola in Alberta in 2015. *Canadian Plant Disease Survey*, *96*: In Press.

Strelkov, S. E., Tewari, J. P., & Smith-Degenhardt, E. (2006). Characterization of *Plasmodiophora brassicae* populations from Alberta, Canada. *Canadian Journal of Plant Pathology*, *28*(3), 467-474.

Struss, D., Bellin, U., & Röbbelen, G. (1991). Development of B-genome chromosome addition lines of *B. napus* using different interspecific *Brassica* hybrids. *Plant Breeding*, *106*(3), 209-214.

Sundelin, T., Christensen, C. B., Larsen, J., Møller, K., Lübeck, M., Bødker, L., & Jensen, B. (2010). In planta quantification of *Plasmodiophora brassicae* using signature fatty acids and real-time PCR. *Plant Disease*, *94*(4), 432-438.

Takano, S., Matsuda, S., Kinoshita, N., Shimoda, N., Sato, T., & Kato, K. (2014). Genome-wide single nucleotide polymorphisms and insertion–deletions of *Oryza sativa* L. subsp. *japonica* cultivars grown near the northern limit of rice cultivation. *Molecular Breeding*, *34*(3), 1007-1021.

Thaler, J. S., Humphrey, P. T., & Whiteman, N. K. (2012). Evolution of jasmonate and salicylate signal crosstalk. *Trends in Plant Science*, *17*(5), 260-270.

Ueno, H., Matsumoto, E., Aruga, D., Kitagawa, S., Matsumura, H., & Hayashida, N. (2012). Molecular characterization of the *CRa* gene conferring clubroot resistance in *Brassica rapa*. *Plant Molecular Biology*, *80*(6), 621-629.

Vigier, B., Chiang M. S., & Hume D. J. (1989). Source of resistance to clubroot (*Plasmodiophora brassicae* Wor.) in triazine-resistant spring canola (rapeseed). *Canadian Plant Disease Survey*, *69*(2), 113-115.

Voorrips, R. E. (1996). Production, characterization and interaction of single-spore isolates of *Plasmodiophora brassicae*. *European Journal of Plant Pathology*, *102*(4), 377-383.

Wallenhammar, A. C. (1996). Prevalence of *Plasmodiophora brassicae* in a spring oilseed rape growing area in central Sweden and factors influencing soil infestation levels. *Plant Pathology*, *45*(4), 710-719.

Walley, J. W., Kliebenstein, D. J., Bostock, R. M., & Dehesh, K. (2013). Fatty acids and early detection of pathogens. *Current Opinion in Plant Biology*, *16*(4), 520-526.

Wang, H. Z., & Dixon, R. A. (2012). On–off switches for secondary cell wall biosynthesis. *Molecular Plant*, *5*(2), 297-303.

Werner, S., Diederichsen, E., Frauen, M., Schondelmaier, J., & Jung, C. (2008). Genetic mapping of clubroot resistance genes in oilseed rape. *Theoretical and Applied Genetics*, *116*(3), 363-372.

Yoshikawa, H. (1993). Studies on breeding of clubroot resistance in cole crops. *Bulletin of the National Research Institute of Vegetables, Ornamental Plants and Tea*, *7*, 1-165.

Yu, F., Zhang, X., Huang, Z., Chu, M., Song, T., Falk, K. C., Deora, A., Chen, Q., Zhang, Y., McGregor, L., Gossen, B. D., McDonald, M. R., & Peng, G. (2016). Identification of genome-wide variants and discovery of variants with *Brassica rapa* clubroot resistance gene *Rcr1* through bulked segregant RNA sequencing. *PloS One*, *11*(4), doi: 10.1371/journal.pone.0153218.

8. APPENDICES

Appendix 1: The list of primers designed for RT-qPCR experiment. Primers from PI 219576 and BRA192/78 are indicated.

RT-qPCR primers from PI 219576							
Gene	**Primer sequence (5'-3')**	**Start**	**Stop**	**Length**	**Tm**	**GC%**	**Product size**
Bra022633	CCTTGGAGGCTTGACAGTATT (Forward)	987	1008	21	62	47.6	129
	TGAGAGCTTGCGCCATAAA (Reverse)	1097	1116	19	62	47.4	
Bra008334	GACCGAGAAATGAAGGGATGT (Forward)	98	119	21	62	47.6	97
	GACCTCAGCTGACGAATATGAG (Reverse)	173	195	22	62	50	
Bra018770	GCTTTGATCTCGTGGACGTATAG (Forward)	322	345	23	62	47.8	102
	ATGAGCAGCCGTGTCAAATA (Reverse)	404	424	20	62	45	
Bra020010	CAGATATGAGAACCGTGAGTGAG (Forward)	500	523	23	62	47.8	88
	CCAGAGACAGCAATGCAAATAC (Reverse)	566	588	22	62	45.5	
Bra019865	CGCTTCGGACAGAGTAGATATG (Forward)	216	238	22	62	50	107
	CTCTCCTTCTCACGGAAGTAAAC (Reverse)	300	323	23	62	47.8	
Bra019824	CCAGGAGGAATGTGGTCAATAC (Forward)	12	34	22	62	50	101
	GGGTTACGAGGTCTAAGCAAAG (Reverse)	91	113	22	62	50	
Bra025396	GCTTCCATCATCACTCTCCTATTC (Forward)	13	37	24	62	45.8	97
	GCTTCCTCTTGCATGACTTCT (Reverse)	89	110	21	62	47.6	
Bra014062	TCCTCCACCGTCAGATATGTA (Forward)	69	90	21	62	47.6	107
	GGATTCTTCGCAGGCTATCTC (Reverse)	155	176	21	62	52.4	
Bra010461	GAAACCTGCTGGTAGGAAGAA (Forward)	102	123	21	62	47.6	95
	TCACTTCACACACCCACTTAC (Reverse)	176	197	21	62	47.6	
Bra030668	CTCAAGGTGTTGTCCGTAAGA (Forward)	167	188	21	62	47.6	93
	GCAAGGAAGAGTCAGGAATGA (Reverse)	239	260	21	62	47.6	
Bra037376	CACTCTCGAACCGTCCAATC (Forward)	66	86	20	62	55	101
	AGATCCAGCCCAGGGTAATA (Reverse)	147	167	20	62	50	

RT-qPCR primers from BRA192/78							
Gene	**Primer sequence (5'-3')**	**Start**	**Stop**	**Length**	**Tm**	**GC%**	**Product size**
Bra023748	Forward TATGACCGAGCCGCCTATAA	178	198	20	62	50	129
	Reverse TGAACCCGAAGATGAAGAAGAAG	284	307	23	62	43.5	
Bra023404	Forward AAGCTACGTCGCATCTCAATC	411	432	21	62	47.6	110
	Reverse GCCACGAAGTAGGAGTAGTAGA	499	521	22	62	50	
Bra020254	Forward AGTGGGCATTGAAGCTACTC	527	547	20	62	50	105
	Reverse TCCACTGACCTCCAACAAATC	611	632	21	62	47.6	
Bra033138	Forward CCCACAGGCTACTCAAAGTT	550	570	20	62	50	109
	Reverse CGACCCTGTAGTTAGTGTTACTG	636	659	23	62	47.8	
Bra032981	Forward CGATACCATGGAGGAGGATTTG	411	433	22	62	50	92
	Reverse GAGGCAACCCAGGTGTTATT	483	503	20	62	50	
Bra012578	Forward GATGTGCTCTCGGTTGTGATA	496	517	21	62	47.6	111
	Reverse GGTGATACGGTAGTGGAAGAAC	585	607	22	62	50	
Bra017612	Forward GGTGAGGATGGAGATACGATTG	1024	1046	22	62	50	100
	Reverse TAGGAGCTGTGAATGGATGATG	1102	1124	22	62	45.5	
Bra014807	Forward GTTCTCCACCACCACCATATT	1385	1406	21	62	47.6	117
	Reverse GTACGTTCGGAGAAGGTGAATAG	1479	1502	23	62	47.8	

Bra031330	Forward GTGAACCAGCACTACCAGTAT	916	937	21	62	47.6	97
	Reverse CTCTCAGCCAAGTCTCCATTAG	991	1013	22	62	50	
Bra010542	Forward CCGCCAGGATTAGGAAACTT	202	222	20	62	50	95
	Reverse CGAGCTAGAGGTTCTGAGAGATA	274	297	23	62	47.8	
Bra010551	Forward CAAATGTTTCCGAGGAGGTAGA	5507	5529	22	62	45.5	95
	Reverse GCCTTCTCTCACCACAAGATAG	5580	5602	22	62	50	
Bra016325	Forward TCTGTCACCGACGACATAAAC	58	79	21	62	47.6	111
	Reverse CGAGATTAGACAGCCACTCATT	147	169	22	62	45.5	
Bra016723	Forward GAGTTCTTCAGCCGAGTATAGC	572	594	22	62	50	79
	Reverse GTCGCCCTATCCAACTCATAAA	629	651	22	62	45.5	
Bra035908	Forward GCTTCACACCATCGGAGATAA	180	201	21	62	47.6	98
	Reverse TGAGAACAAGACCGGGAAAG	258	278	20	62	50	
Bra026922	Forward AGGTGTCCAACACTGAGAAAG	109	130	21	62	47.6	110
	Reverse TCCCATTGAACCCTTTCTATCC	197	219	22	62	45.5	
Bra032469	Forward ATCTCTGCGTCGATAAGAAAGG	290	312	22	62	45.5	112
	Reverse ATGCCGGAGTTGATCCATATC	381	402	21	62	47.6	

Bra003069	Forward CAACGAACAGAGAGGGATGTAG	241	263	22	62	50	124
	Reverse CAACGAAGGCAAGAGTTTCATC	343	365	22	62	45.5	
Bra002117	Forward GTCTCTGGCCTCAGCTTATTT	2261	2282	21	62	47.6	110
	Reverse TATTCCCGTACTGCTCAAATCC	2349	2371	22	62	45.5	
Bra009033	Forward CTCCCTGGCTATATGTGAAACC	384	406	22	62	50	110
	Reverse TGGCAACACCGGAAAGAA	476	494	18	62	50	
Bra008381	Forward TGCAGAGTCGTCTCTCATATCT	636	658	22	62	45.5	93
	Reverse TTACCCTTCCCGTCCCTTAT	709	729	20	62	50	
Bra017261	Forward GAGGAGACACGAGGAGTTTATG	807	829	22	62	50	99
	Reverse ATTCCGACAGCTACCGTTAAA	885	906	21	62	42.9	
Bra037740	Forward ATCGCAACGACAGACCTATG	146	166	20	62	50	116
	Reverse CGCTCCCATTCCTTGGATTA	242	262	20	62	50	
Bra029994	Forward TGATGGGTGGAGGACTAAGA	915	935	20	62	50	98
	Reverse CATCAACACCGATAAGCACAAG	991	1013	22	62	45.5	
Bra025770	Forward TCTACAACAAGCTGAGGAAGAG	237	259	22	62	45.5	97
	Reverse CATGGAGGTGACAATGGATTTG	312	334	22	62	45.5	
Bra025857	Forward GGGACTATTCCTCTTCTGACTTTG	856	880	24	62	45.8	102
	Reverse CACAACCACCACGCCTAATA	938	958	20	62	50	

Bra012231	Forward GAAGTGGAGGAAGTGGAAGAAG	48	70	22	62	50	116
	Reverse CCATCTCCAGTAAGACCTGAAAG	141	164	23	62	47.8	
Bra009535	Forward CCACCGAGAACCTTCATATACC	74	96	22	62	50	92
	Reverse CACGAACCTCTTGTAACTCTCC	144	166	22	62	50	
Bra023486	Forward GGACGGATAACCAAGTGAAGAA	278	300	22	62	45.5	113
	Reverse GAGACTGGTCTGATGGACAATATC	367	391	24	62	45.8	
Bra034855	Forward TCCGACCACAGACCAAATTC	141	161	20	62	50	94
	Reverse CTCCTCCTCGTCTTCGTTTATTT	212	235	23	62	43.5	
Bra015844	Forward GTTATCGACATGGGTGGTAGAG	97	119	22	62	50	105
	Reverse GATCTCAGAGGACAAGCCAAA	181	202	21	62	47.6	
Bra008670	Forward GTCATCAACGAGAGAAAGCAAAG	202	225	23	62	43.5	102
	Reverse CTGCGAGAGAAGCTCATCTAAG	282	304	22	62	50	
Bra007739	Forward CGCGAACATTCTCCCTCTATAC	186	208	22	62	50	150
	Reverse CATCTCTCTGCCTCAATCTTTCT	313	336	23	62	43.5	
Bra001387	Forward ACGAAGGACAACGAGTCTTAAA	633	655	22	62	40.9	96
	Reverse GTCTTCACTCCACTAGCAATACA	706	729	23	62	43.5	
Bra005210	Forward GTCCTGTTAAGAAGAGCGTAGAG	572	595	23	62	47.8	104
	Reverse CAGAAGAGGACAAGAGGGTTTAG	653	676	23	62	47.8	

Appendix 2: Genotype scoring data of the 42 polymorphic SSR markers tested on 32 BC_1 plants from PI 219576. Genemapper 5.0 software was used to analyze the fragment bands of R and S samples so that the alleles from R parents were scored as '+' and the alleles from S parents were scored as '−'. A scoring of '0' was ambiguous data.

	1	2	3	4	5	7	8	9	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	30	31	32	35	36	37
Phenotype	-	-	-	-	-	-	-	-	-	+	-	+	+	+	-	-	-	+	-	+	+	+	+	-	+	-	+	+	+	+	+	+
sB2112a	+	+	+	-	+	+	-	-	+	-	+	+	-	-	+	+	-	-	+	+	-	+	+	0	+	+	+	-	+	+	+	+
sB2112b	-	-	-	-	-	-	+	+	-	+	-	+	+	+	-	-	+	+	-	-	-	-	-	0	-	-	-	-	-	-	-	-
sB1956a	+	+	-	-	-	+	-	-	+	-	+	+	-	+	+	+	-	-	+	-	-	+	+	-	+	-	+	-	+	+	+	-
sB1956b	-	-	-	-	-	-	+	-	+	+	+	-	-	-	+	-	+	+	-	-	+	+	-	-	-	+	-	+	-	+	+	-
sB1956c	-	+	-	-	-	-	+	-	-	+	-	-	-	+	-	-	+	+	+	-	+	-	-	-	+	+	-	+	-	-	-	-
sJ7148a	-	-	-	-	-	-	-	-	-	+	-	+	+	+	-	-	-	+	-	+	+	+	+	-	+	-	+	+	+	+	+	+
sJ7148b	-	+	-	-	+	-	+	+	-	+	-	+	+	+	-	-	+	+	+	-	+	-	-	+	+	-	-	+	-	-	-	-
sJ7190	-	-	+	-	-	-	+	+	+	+	+	+	+	+	+	-	+	+	-	+	+	+	+	+	-	+	-	+	-	+	+	+
sJ1361	-	-	+	-	-	-	+	+	+	+	+	-	+	-	+	-	+	+	-	+	+	+	-	+	-	+	-	+	-	+	+	+
sB2750Ia	-	-	+	-	-	-	+	+	+	+	+	-	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	0	+	+	+	+
sB2750Ib	-	+	-	+	+	-	+	+	-	+	-	+	+	+	-	-	+	+	+	-	+	-	-	-	+	-	-	0	-	-	-	-
sJ7104	-	-	+	+	-	-	+	+	+	+	+	-	+	+	+	-	+	+	-	+	+	+	-	+	-	+	-	0	-	+	+	+
sJ6847a	-	+	-	+	+	-	+	+	-	+	-	+	+	+	-	-	+	+	+	-	+	+	+	+	+	+	-	0	+	-	-	-
sJ6847b	+	+	-	-	+	+	-	-	-	-	-	+	-	+	-	+	-	-	+	-	-	+	+	-	+	-	+	0	+	-	-	-
sB1725Fa	-	+	+	+	-	-	+	+	+	+	+	+	+	+	+	-	+	-	-	+	+	+	-	-	-	+	-	+	-	+	+	+
sB1725Fb	+	+	-	+	+	+	+	-	-	+	-	+	+	+	-	-	+	+	-	-	+	-	-	-	-	+	+	-	-	-	+	-
sJ8228a	-	+	-	-	+	+	-	+	-	+	-	-	-	-	-	-	-	+	-	+	-	-	+	+	+	+	-	0	+	-	-	+
sJ8228b	+	+	+	+	-	+	+	-	+	-	+	+	+	+	+	+	+	+	+	-	+	+	-	-	+	+	+	0	-	+	+	-
sJ8211a	+	+	+	-	+	-	+	0	+	-	+	-	+	-	-	+	+	-	-	+	+	-	+	0	-	+	-	-	-	-	+	-
sB3868a	+	+	+	-	+	-	+	0	+	-	+	-	+	-	-	+	+	-	-	+	+	-	+	0	-	+	-	-	-	-	0	-
sJ4216a	+	+	+	+	+	+	+	0	+	-	+	+	-	+	+	+	+	+	+	-	+	+	+	0	-	+	-	+	+	-	+	+
sJ4216b	-	+	-	-	+	-	+	0	+	+	-	+	+	+	+	+	+	+	+	+	+	+	-	0	+	-	+	+	+	+	+	-
sJ6346	-	+	+	+	+	+	-	0	+	+	+	-	-	-	+	-	+	+	-	-	-	-	-	-	-	-	+	-	-	-	-	-
sJ5080II	+	-	+	-	-	-	-	0	-	+	-	+	-	-	-	+	-	+	-	+	+	+	+	+	-	-	+	-	+	-	-	+
sB2145	-	+	+	+	+	+	-	0	+	+	+	-	-	-	+	+	+	+	+	-	-	-	-	-	-	-	+	-	+	-	-	-
sJ3659Fa	+	+	-	-	+	-	+	-	-	-	+	+	-	+	-	-	+	-	-	+	+	+	-	-	-	-	+	-	-	-	-	-
sJ3659Fb	+	-	+	-	-	-	+	-	-	-	+	-	+	+	-	+	+	-	+	+	+	-	+	+	+	-	-	+	+	+	+	-
sJ34109a	+	-	+	-	+	-	-	-	+	+	-	+	+	+	-	+	+	+	+	-	+	+	+	-	-	+	-	-	-	-	+	+
sJ34109b	+	+	-	+	-	+	+	+	+	-	+	-	-	+	+	-	+	+	-	+	+	-	+	+	+	+	+	+	+	+	+	-
sB5119a	+	+	+	-	-	+	-	0	+	-	+	-	-	-	+	+	+	+	+	+	+	+	+	0	+	+	+	+	+	+	+	+
sB5119b	+	-	+	-	-	-	-	0	+	+	-	+	+	+	-	+	+	+	+	-	+	+	+	0	-	+	-	-	-	-	+	+
sB5079	-	-	-	-	-	-	+	-	+	+	-	+	-	-	+	+	-	-	-	+	-	-	-	-	+	-	-	0	-	+	+	-
sJ01106	+	+	+	-	-	-	+	-	+	+	-	+	+	-	+	+	+	-	+	+	+	+	+	0	+	+	+	+	+	-	-	-
sB1554	-	+	+	+	-	-	-	+	-	-	-	-	+	-	+	-	+	-	+	-	+	-	+	0	-	-	+	+	+	-	-	+
sJ2971	+	-	-	-	+	+	+	-	+	+	+	+	-	+	-	+	-	+	-	+	-	+	-	0	+	+	-	-	-	+	+	+
sB4727F	-	-	+	-	+	+	-	+	-	-	-	+	+	+	+	+	-	+	-	+	+	+	+	+	-	-	-	+	+	-	+	-
sJ8111	-	+	-	+	-	+	+	-	-	+	-	-	+	+	-	-	-	+	+	-	+	+	+	+	+	-	-	-	-	-	+	+
sJ2480R	-	-	-	-	-	-	+	-	+	+	-	+	-	-	+	-	-	-	-	-	-	-	-	-	+	-	-	+	-	+	+	-
sJ3485R	+	+	+	0	+	+	-	-	-	+	+	-	+	+	+	-	+	+	+	+	0	+	+	+	+	-	+	-	+	+	+	+
sB0860Aa	+	+	+	0	+	+	-	+	-	-	+	-	+	+	+	-	+	+	+	-	0	+	+	+	-	+	+	+	+	-	-	+
sB0860Ab	-	-	+	0	-	-	+	+	+	-	-	+	-	-	+	+	-	-	-	+	0	-	-	-	+	+	-	+	-	+	-	-
sB5162	+	-	-	0	-	-	+	+	+	-	-	+	-	-	+	-	+	-	+	-	0	+	+	-	+	+	+	+	+	+	+	+

Appendix 3: Best hit to *Arabidopsis thaliana* genome and gene annotation in the *Rcr6a* and *Rcr6b* target regions.

Gene name	***A. thaliana* annotation**
Bra034556	NBS-LRR
Bra010248	scpl29; scpl29 (serine carboxypeptidase-like 29); serine-type carboxypeptidase
Bra010272	scpl29; scpl29 (serine carboxypeptidase-like 29); serine-type carboxypeptidase
Bra010384	binding
Bra010441	ABI1; ABI1 (ABA INSENSITIVE 1); calcium ion binding / protein serine/threonine phosphatase
Bra010551	NBS-LRR
Bra010552	NBS-LRR
Bra010588	NBS-LRR
Bra010589	NBS-LRR
Bra040085	"LUG, RON2; LUG (LEUNIG); protein binding / protein heterodimerization/ transcription repressor"
Bra040069	ATPANK2; ATPANK2 (PANTOTHENATE KINASE 2); pantothenate kinase
Bra010234	"coatomer beta subunit, putative / beta-coat protein, putative / beta-COP, putative"
Bra010355	PHR1, AtPHR1; PHR1 (PHOSPHATE STARVATION RESPONSE 1); transcription factor
Bra010398	Unknown protein
Bra010400	PPI1; PPI1 (PROTON PUMP INTERACTOR 1); protein binding
Bra010419	aconitate hydratase/ copper ion binding
Bra010432	WD-40 repeat family protein
Bra010453	MATE efflux family protein

Bra010472	"oxidoreductase, 2OG-Fe(II) oxygenase family protein"
Bra010484	ubiquitin-associated (UBA)/TS-N domain-containing protein / octicosapeptide/Phox/Bemp1 (PB1) domain-containing protein
Bra010559	ubiquitin-protein ligase activity
Bra010564	NOL6_XENLA Nucleolar protein 6 OS=Xenopus laevis GN=nol6 PE=2 SV=2
Bra010574	ATAPP1, APP1; ATAPP1; N-1-naphthylphthalamic acid binding / aminopeptidase

Appendix 4: The list of annotated up- and down-regulated DEGs for PI 219576 and BRA192/78.

PI = 11 DEGs at 95% confidence

Name	Fold Change	Chr.	Annotations
Bra019865	59.556 up	A06	S-locus lectin protein kinase family protein
Bra014062	103.805 down	A08	N/A
Bra018770	39.391 up	A06	Glycine-rich protein/oleosin
Bra022633	36.800 up	A02	F-box family protein
Bra008334	75.969 up	A02	Unknown protein
Bra037376	43.411 up	A09	FUNCTIONS IN: Molecular function unknown
Bra030668	38.675 up	A08	Unknown protein
Bra010461	31.042 up	A08	DREB1A, CBF3, ATCBF3/DREB1A (DEHYDRATION RESPONSE ELEMENT B1A)
Bra025396	78.783 up	A06	PDF1.2b (plant defensin 1.2b)
Bra020010	31.718 down	A06	Emp24/gp25L/p24 family protein
Bra019824	38.396 down	A06	Myb family transcription factor

BRA = 382 DEGs at 95% confidence

Name	Chr.	Fold change	Annotations
Bra011893	A01	1.683 up	"PRP4, ATPRP4; PRP4 (PROLINE-RICH PROTEIN 4)"
Bra011861	A01	111.346 up	WRKY DNA-BINDING PROTEIN 13
Bra011799	A01	3.571 down	"ELI3-1, ELI3, ATCAD7, CAD7; ELI3-1 (ELICITOR-ACTIVATED GENE 3-1)
Bra011417	A01	1.286 down	Unknown protein
Bra011346	A01	1.242 down	Peptidyl-prolyl cis-trans isomerase cyclophilin-type family protein
Bra011296	A01	1.149 down	WIN2; WIN2 (HOPW1-1-INTERACTING 2)
Bra011215	A01	83.070 down	Zinc finger (C3HC4-type RING finger) family protein
Bra013332	A01	51.038 down	"ATGSL07, gsl07, atgsl7; ATGSL07 (glucan synthase-like 7)
Bra013441	A01	39.429 down	M4DAD1_BRARP Uncharacterized protein OS=Brassica
Bra013549	A01	1.255 up	Unknown protein
Bra013645	A01	1.730 up	Pentatricopeptide (PPR) repeat-containing protein
Bra013724	A01	40.700 up	ATCHX17, CHX17; ATCHX17 (CATION/H+ EXCHANGER 17)
Bra013743	A01	1.491 down	Unknown protein
Bra013791	A01	2.967 down	Ubiquitin thiolesterase
Bra013856	A01	1.213 down	Unknown protein

Bra039768	A01	1.685 up	Glycosyl hydrolase family 10 protein/carbohydrate-binding domain-containing protein
Bra026414	A01	1.384 down	"WRKY20, AtWRKY20; WRKY20; transcription factor"
Bra026395	A01	1.787 down	MAPKKK16; MAPKKK16
Bra038449	A01	88.902 down	Metal ion binding
Bra033795	A01	32.542 up	Glycosyl hydrolase family 3 protein
Bra029968	A01	1.468 down	AtRLP44; AtRLP44 (Receptor Like Protein 44)
Bra029994	A01	34.346 down	VFB2; VFB2 (VIER F-BOX PROTEINE 2)
Bra031483	A01	37.336 down	IPR002213; UDP glucuronosyl/UDP-glucosyltransferase
Bra023748	A01	110.521 down	IPR001471; Pathogenesis-related transcriptional factor
Bra023775	A01	113.855 down	Pentatricopeptide (PPR) repeat-containing protein
Bra023859	A01	1.317 down	PEX22; PEX22 (peroxin 22); protein binding
Bra023922	A01	96.015 down	EMB3011; EMB3011 (embryo defective 3011)
Bra021261	A01	42.555 down	SON1; SON1 (SUPPRESSOR OF NIM1-1 1)
Bra021225	A01	89.938 down	Invertase/pectin methylesterase inhibitor family protein
Bra021087	A01	44.074 down	Unknown
Bra040513	A01	2.605 down	Armadillo/beta-catenin repeat family protein
Bra040517	A01	1.169 up	Oxygen evolving enhancer 3 (PsbQ) family protein
Bra028873	A02	1.327 down	"SETH2; SETH2; transferase, transferring glycosyl groups"
Bra028833	A02	1.364 down	Unknown
Bra028786	A02	168.621 down	KCS19; KCS19 (3-KETOACYL-COA SYNTHASE 19)
Bra028751	A02	1.713 down	IPR017972; Cytochrome P450, conserved site
Bra028716	A02	1.348 down	Mitochondrial transcription termination factor-related / mTERF-related
Bra028618	A02	2.313 down	"ATXR5, SDG15; ATXR5; DNA binding / protein binding"
Bra028594	A02	1.182 up	K03535; major facilitator transporter ; MFS transporter, ACS family, glucarate transporter
Bra023404	A02	116.133 down	IPR001471; Pathogenesis-related transcriptional factor
Bra023486	A02	64.514 down	WER, ATMYB66, WER1; ATMYB66 (MYB DOMAIN PROTEIN 66)
Bra023611	A02	17.270 down	K09338; homeobox-leucine zipper family protein
Bra023618	A02	1.250 down	LUC7 N_terminus domain-containing protein
Bra020230	A02	39.581 down	"ATNRT2.3, NRT2.3; ATNRT2.3; nitrate transmembrane transporter"
Bra020254	A02	1.376 down	DEM1_BOVIN Defects in morphology protein 1 homolog

Bra020268	A02	1.655 down	Unknown
Bra020336	A02	1.284 down	NRPB7; NRPB7; DNA directed RNA polymerase/RNA binding
Bra022645	A02	1.232 down	UBC10; UBC10 (ubiquitin conjugating enzyme 10); ubiquitin-protein ligase
Bra022534	A02	1.411 up	SDH5_CAEBR Succinate dehydrogenase subunit 5
Bra033980	A02	1.660 down	Pectate lyase family protein
Bra008021	A02	138.253 down	Plastocyanin-like domain-containing protein
Bra008125	A02	53.734 down	SS3; SS3 (STRICTOSIDINE SYNTHASE 3); strictosidine synthase
Bra008124	A02	38.900 down	SS3; SS3 (STRICTOSIDINE SYNTHASE 3); strictosidine synthase
Bra008381	A02	73.286 up	"Wall-associated kinase, putative"
Bra008553	A02	1.427 up	anac069; anac069 (Arabidopsis NAC domain containing protein 69)
Bra018507	A02	169.092 down	MED21; MED21 (MEDIATOR 21)
Bra033138	A02	56.970 down	IPR001938; Thaumatin, pathogenesis-related
Bra026621	A02	2.460 up	GAD4; GAD4 (glutamate decarboxylase 4); calmodulin binding
Bra026517	A02	266.115 down	Unknown
Bra032981	A02	1.172 down	IPR013989; Development/cell death domain
Bra020653	A02	1.571 down	IPR019787; Zinc finger, PHD-finger
Bra031868	A02	1.185 down	LMBR1 integral membrane family protein
Bra031828	A02	1.297 up	MAPKKK5; MAPKKK5; ATP binding / kinase/ protein kinase/ protein serine/threonine kinase
Bra005664	A03	7.902 down	Unknown
Bra005788	A03	68.118 down	Unknown
Bra005959	A03	1.372 down	Myosin heavy chain-related
Bra006072	A03	1.377 down	Aminoacyl-tRNA hydrolase/ protein tyrosine phosphatase
Bra006184	A03	1.187 up	PBS1; PBS1 (avrPphB susceptible 1); kinase/ protein serine/threonine kinase
Bra006464	A03	1.566 down	IPR009637; Transmembrane receptor, eukaryota
Bra022796	A03	2.385 up	"ATOFP2, OFP2; OFP2 (OVATE FAMILY PROTEIN 2)"
Bra000032	A03	1.430 down	YOS1_SCHPO Protein transport protein
Bra000053	A03	1.956 up	Unknown
Bra000231	A03	1.592 down	GPDHC1; GPDHC1; NAD or NADH binding / glycerol-3-phosphate dehydrogenase (NAD+)
Bra000264	A03	1.811 down	Zinc finger (B-box type) family protein
Bra000291	A03	68.420 down	DNA binding / transcription factor

Bra000373	A03	58.874 up	"ATSPX3, SPX3; SPX3 (SPX DOMAIN GENE 3)"
Bra000435	A03	78.328 down	Unknown
Bra000477	A03	1.722 down	"Tropinone reductase, putative / tropine dehydrogenase, putative"
Bra000519	A03	67.085 down	IPR010847; Harpin-induced 1
Bra000775	A03	125.587 up	IPR003612; Plant lipid transfer protein/seed storage/trypsin-alpha amylase inhibitor
Bra000931	A03	1.409 down	RNA recognition motif (RRM)-containing protein
Bra000952	A03	1.137 down	Unknown
Bra000962	A03	1.576 down	BSK3; BSK3 (BR-SIGNALING KINASE 3)
Bra001095	A03	39.260 down	KPYC_SOYBN Pyruvate kinase
Bra001189	A03	1.975 down	"Protein phosphatase 2C, putative / PP2C, putative"
Bra001222	A03	145.649 down	Unknown
Bra001387	A03	44.993 down	Transcriptional factor B3 family protein
Bra001421	A03	1.231 down	Unknown
Bra001580	A03	1.441 down	Unknown
Bra001723	A03	13.805 down	Protein kinase family protein
Bra013253	A03	1.182 down	"AT-HSFC1, HSFC1; AT-HSFC1; DNA binding / transcription factor"
Bra013130	A03	23.181 down	K00517; CYP705A2; CYP705A2 (cytochrome P450, family 705, subfamily A, polypeptide 2)
Bra013006	A03	1.977 down	Zinc finger (C3HC4-type RING finger) family protein
Bra012912	A03	1.145 down	IPR007307; Low temperature viability protein
Bra012911	A03	57.916 down	Unknown
Bra012761	A03	1.191 up	Protein binding
Bra012635	A03	1.304 down	IPR005031; Streptomyces cyclase/dehydrase
Bra012594	A03	3.955 down	Unknown
Bra012578	A03	1.163 up	CESA8, IRX1, ATCESA8, LEW2; IRX1 (IRREGULAR XYLEM 1)
Bra012558	A03	1.167 down	Protein kinase, lipid binding
Bra038792	A03	162.361 down	basic helix-loop-helix (bHLH) family protein
Bra024136	A03	1.247 down	Unknown protein
Bra023979	A03	1.323 down	"Protein phosphatase 2C, putative / PP2C, putative"
Bra017612	A03	1.457 down	IPR001471; Pathogenesis-related transcriptional factor\/ERF, DNA-binding
Bra017742	A03	1.314 down	CSDP1; CSDP1 (cold shock domain protein 1)
Bra040421	A04	1.318 down	MBD4; MBD4; DNA binding / DNA demethylase/ methyl-CpG binding
Bra014393	A04	1.088 up	WIPI3_XENLA WD repeat domain phosphoinositide-interacting protein
Bra014404	A04	2.112 up	Anion exchange family protein
Bra014480	A04	1.995 down	Phosphatidate cytidylyltransferase family protein

Bra014516	A04	1.640 down	Pentatricopeptide (PPR) repeat-containing protein
Bra014618	A04	1.259 down	Aspartate/glutamate/uridylate kinase family protein
Bra014727	A04	69.500 down	Unknown protein
Bra014807	A04	15.128 down	GO:0009664; plant-type cell wall organization
Bra039805	A04	223.117 up	ATIREG1; ATIREG1 (IRON-REGULATED PROTEIN 1); transporter
Bra032725	A04	1.932 down	Unknown protein
Bra028102	A04	89.655 down	Unknown protein
Bra025625	A04	171.081 up	60S acidic ribosomal protein P2 (RPP2E)
Bra025608	A04	2.515 down	IPR000232; Heat shock factor (HSF)-type, DNA-binding
Bra025571	A04	2.224 down	Transducin family protein
Bra034264	A04	269.637 up	Unknown protein
Bra021636	A04	1.355 down	ATC4H, C4H, CYP73A5; C4H (CINNAMATE-4-HYDROXYLASE)
Bra021757	A04	104.945 down	NRPD1B, DRD3, ATNRPD1B, DMS5, NRPE1; NRPD1B; DNA binding / DNA-directed RNA polymerase
Bra021770	A04	75.447 down	GCN5-related N-acetyltransferase (GNAT) family protein
Bra017261	A04	100.805 down	F-box family protein
Bra016934	A04	1.104 down	ERD15, LSR1, CID1; ERD15 (EARLY RESPONSIVE TO DEHYDRATION 15); protein binding
Bra037740	A04	249.245 down	F-box family protein
Bra037669	A04	1.166 down	Late embryogenesis abundant family protein / LEA family protein
Bra040335	A04	37.867 down	ATHB4, ATHB-4; ATHB4 (ARABIDOPSIS THALIANA HOMEOBOX-LEUCINE ZIPPER PROTEIN 4)
Bra004427	A05	1.290 down	Zinc finger (CCCH-type) family protein
Bra004558	A05	34.837 down	Oxidoreductase family protein
Bra004637	A05	89.351 down	MIZ1; MIZ1 (mizu-kussei 1)
Bra004738	A05	1.399 down	NPG1; NPG1 (no pollen germination 1); calmodulin binding
Bra004801	A05	1.247 down	Ribosomal protein L2 family protein
Bra005140	A05	1.785 up	"IRX12, LAC4; IRX12 (IRREGULAR XYLEM 12); laccase"
Bra005210	A05	37.469 down	G2, ATWRKY44, WRKY44, DSL1; TTG2 (TRANSPARENT TESTA GLABra 2); transcription factor"
Bra005352	A05	1.320 down	Protein binding / zinc ion binding
Bra005380	A05	110.847 down	CLE42; CLE42 (CLAVATA3/ESR-RELATED 42)
Bra005478	A05	1.331 up	RD21, RD21A; RD21 (responsive to dehydration 21)
Bra005558	A05	2.867 down	"SUM5, SUMO5, ATSUMO5; SUMO5 (SMALL UBIQUITINRELATED MODIFIER 5); protein tag"

Bra005653	A05	50.207 down	Protein kinase family protein
Bra018255	A05	1.775 down	emb1381; emb1381 (embryo defective 1381)
Bra018313	A05	1.198 down	HSI2, VAL1; HSI2 (HIGH-LEVEL EXPRESSION OF SUGAR-INDUCIBLE GENE 2)
Bra018333	A05	1.222 down	UDP-glucoronosyl/UDP-glucosyl transferase family protein
Bra018374	A05	272.890 up	Transferase family protein
Bra018465	A05	1.733 down	Unknown protein
Bra038102	A05	153.422 up	ATPER1; ATPER1; antioxidant/ thioredoxin peroxidase
Bra038098	A05	2.548 down	SPA4; SPA4 (SPA1-RELATED 4); protein binding / signal transducer
Bra030397	A05	1.175 up	Unknown protein
Bra030422	A05	1.314 down	RWD domain-containing protein
Bra030484	A05	76.970 down	CYP705A27; CYP705A27; electron carrier/ heme binding / iron ion binding / monooxygenase/ oxygen binding
Bra034425	A05	1.290 up	"Oxidoreductase, 2OG-Fe(II) oxygenase family protein"
Bra034444	A05	1.255 down	IP5PI, AT5P1, ATIP5PI, AT5PTASE1; IP5PI (INOSITOL POLYPHOSPHATE 5-PHOSPHATASE I)
Bra031330	A05	1.219 down	FHY3; FHY3 (FAR-RED ELONGATED HYPOCOTYLS 3); transcription activator/ transcription factor
Bra031305	A05	1.829 down	APO4; APO4 (ACCUMULATION OF PHOTOSYSTEM ONE 4)
Bra035761	A05	1.606 up	Ethylene-responsive protein - related
Bra022441	A05	1.414 down	Magnesium transporter CorA-like family protein (MRS2-3)
Bra022397	A05	23.676 down	Leucine-rich repeat family protein
Bra022274	A05	27.216 down	HON4; HON4; DNA binding
Bra022250	A05	1.446 down	RLK902; RLK902; ATP binding / kinase/ protein serine/threonine kinase
Bra027252	A05	1.441 up	"UBC25, PFU1; UBC25 (UBIQUITIN-CONJUGATING ENZYME 25); small conjugating protein ligase"
Bra027431	A05	23.448 down	"Chaperonin, putative"
Bra034855	A05	134.413 down	WOX5; WOX5 (WUSCHEL RELATED HOMEOBOX 5); transcription factor
Bra029890	A05	1.233 down	Unknown protein
Bra029720	A05	1.441 down	Unknown protein
Bra029710	A05	1.332 down	"leucine-rich repeat transmembrane protein kinase, putative"
Bra029632	A05	1.312 down	Unknown protein
Bra029599	A05	1.235 down	Unknown protein
Bra020769	A05	31.319 down	"ATMKK9, MKK9; MKK9 (MAP KINASE KINASE 9); MAP kinase kinase/ kinase/ protein kinase activator"
Bra039149	A05	1.324 down	GAUT13; GAUT13 (Galacturonosyltransferase 13)
Bra038020	A06	1.407 down	Unknown protein
Bra018746	A06	1.267 down	"PEX3, PEX3-2; peroxin-3 family protein"
Bra018639	A06	1.546 down	Glycosyl transferase family 29 protein / sialyltransferase family protein

Bra019942	A06	137.079 up	"DNAJ heat shock protein, putative"
Bra019750	A06	2.098 up	"VND4, EMB2749, ANAC007; ANAC007 (ARABIDOPSIS NAC 007); transcription factor"
Bra019747	A06	1.709 up	Flavin-containing monooxygenase family protein / FMO family protein
Bra019698	A06	2.116 up	Remorin family protein
Bra019676	A06	112.861 down	"2-oxoglutarate-dependent dioxygenase, putative"
Bra026166	A06	58.653 down	IAA34; IAA34 (INDOLE-3-ACETIC ACID INDUCIBLE 34); transcription factor
Bra025698	A06	75.607 down	Protein binding / zinc ion binding
Bra025761	A06	1.099 down	Pentatricopeptide (PPR) repeat-containing protein
Bra025770	A06	134.758 down	"Auxin-responsive protein, putative"
Bra025784	A06	1.599 up	HVA22H; HVA22H (HVA22-LIKE PROTEIN H (ATHVA22H))
Bra025793	A06	1.886 up	Zinc finger (FYVE type) family protein
Bra025857	A06	23.954 down	Auxin efflux carrier family protein
Bra018005	A06	1.107 down	Unknown protein
Bra018016	A06	1.705 down	"Lipase class 3 family protein / calmodulin-binding heat-shock protein, putative"
Bra018112	A06	1.242 down	Aldose 1-epimerase family protein
Bra018127	A06	1.358 down	Transcription regulator/ zinc ion binding
Bra018203	A06	1.296 down	UDP-GLUCOSYL TRANSFERASE 76E11
Bra019433	A06	1.117 up	Peptidase M1 family protein
Bra019536	A06	1.260 down	Pentatricopeptide (PPR) repeat-containing protein
Bra019540	A06	108.889 down	MA3 domain-containing protein
Bra019566	A06	2.546 down	DNA-binding protein-related
Bra038632	A06	62.778 down	CYP94B1; CYP94B1; electron carrier/ heme binding / iron ion binding / monooxygenase/ oxygen binding
Bra024449	A06	2.250 up	HMA4; HMA4; cadmium ion transmembrane transporter
Bra024498	A06	1.076 up	Unknown protein
Bra009701	A06	1.587 down	Sterile alpha motif (SAM) domain-containing protein
Bra009754	A06	1.294 up	"ATIRE1-1, IRE1-1; IRE1-1 (INOSITOL REQUIRING 1-1)"
Bra009846	A06	1.407 down	"Serine/threonine protein phosphatase 2A (PP2A) regulatory subunit B', putative"
Bra037468	A06	14.112 down	Unknown protein
Bra037531	A06	1.285 up	SOUL heme-binding family protein
Bra025366	A06	1.443 down	"Avirulence-responsive protein, putative / avirulence induced gene (AIG) protein, putative"
Bra025258	A06	1.607 down	Unknown protein
Bra025230	A06	36.506 down	Unknown protein
Bra025190	A06	51.169 down	CYP71B20; CYP71B20; electron carrier/ heme binding / iron ion binding / monooxygenase/ oxygen binding
Bra025180	A06	1.134 up	Plastid-lipid associated protein PAP / fibrillin family protein
Bra024825	A06	2.569 up	AtPP2-B1; AtPP2-B1 (Phloem protein 2-B1); carbohydrate binding

Bra024858	A06	1.227 down	"Leucine-rich repeat protein kinase, putative"
Bra024872	A06	102.954 down	Unknown protein
Bra024896	A06	1.113 up	Nucleotidyltransferase
Bra024897	A06	2.058 down	"Leucine-rich repeat transmembrane protein kinase, putative"
Bra024958	A06	1.487 up	Vesicle-associated membrane family protein / VAMP family protein
Bra033745	A06	167.895 down	"Calcium-binding protein, putative"
Bra033631	A06	17.038 down	GGT1; GGT1 (GAMMA-GLUTAMYL TRANSPEPTIDASE 1)
Bra039027	A07	2.627 down	Haloacid dehalogenase-like hydrolase family protein
Bra039028	A07	1.196 up	"Ceramide glucosyltransferase, putative"
Bra039003	A07	1.215 down	TAG1, AS11, ABX45, DGAT1, RDS1, ATDGAT; TAG1 (TRIACYLGLYCEROL BIOSYNTHESIS DEFECT 1)
Bra039604	A07	38.586 down	"ATOFP7, OFP7; OFP7 (ARABIDOPSIS THALIANA OVATE FAMILY PROTEIN 7)"
Bra039624	A07	1.513 up	Unknown protein
Bra001982	A07	1.595 down	SSRA_ARATH Translocon-associated protein subunit alpha
Bra015109	A07	171.260 down	Nucleic acid binding / zinc ion binding
Bra014911	A07	1.330 down	IPR015798; Copper amine oxidase, C-terminal
Bra012241	A07	1.978 down	RWP-RK domain-containing protein
Bra012231	A07	127.338 down	Auxin-responsive family protein
Bra011968	A07	63.177 down	Lipin family protein
Bra003593	A07	1.295 down	MIF4G domain-containing protein / MA3 domain-containing protein
Bra003665	A07	2.023 down	ARF17; ARF17 (AUXIN RESPONSE FACTOR 17); transcription factor
Bra003718	A07	1.713 down	IPR002068; Heat shock protein Hsp20
Bra004090	A07	1.561 down	Rubber elongation factor (REF) family protein
Bra004190	A07	1.872 down	Glycine-rich protein
Bra004280	A07	1.552 down	Zinc finger (C3HC4-type RING finger) family protein
Bra004300	A07	64.820 down	zinc finger protein-related
Bra016081	A07	1.237 up	"leucine-rich repeat transmembrane protein kinase, putative"
Bra016001	A07	23.097 down	scpl3; scpl3 (serine carboxypeptidase-like 3); serine-type carboxypeptidase
Bra015929	A07	2.197 down	AtRLP14; AtRLP14 (Receptor Like Protein 14); protein binding
Bra015844	A07	162.753 down	BLH11; BLH11 (BEL1-LIKE HOMEODOMAIN 11); transcription factor
Bra015796	A07	1.627 up	CPK29; CPK29; ATP binding / calcium ion binding / calmodulin-dependent protein kinase

Bra015661	A07	1.691 down	Nucleoside-triphosphatase/ nucleotide binding
Bra035193	A07	280.514 down	Zinc ion binding
Bra030855	A08	1.638 down	TBP2, ATTBP2; TBP2 (TATA BINDING PROTEIN 2)
Bra030943	A08	60.271 down	Encodes a ECA1 gametogenesis related family protein
Bra014353	A08	24.533 down	Replication protein-related
Bra034875	A08	1.367 down	Phosphoinositide binding
Bra038434	A08	162.546 down	Unknown protein
Bra038413	A08	28.543 down	CYP96A1; CYP96A1; electron carrier/ heme binding / iron ion binding / monooxygenase/ oxygen binding
Bra021021	A08	109.337 up	"lil3 protein, putative"
Bra020966	A08	1.326 down	"ATRLI2, RLI2; ATRLI2; transporter"
Bra020887	A08	2.385 up	AAT1, CAT1; AAT1 (AMINO ACID TRANSPORTER 1)
Bra020851	A08	1.806 down	SWIB complex BAF60b domain-containing protein
Bra034667	A08	2.371 down	"Peptidyl-prolyl cis-trans isomerase, putative / cyclophilin, putative / rotamase, putative"
Bra040078	A08	1.741 down	Agenet domain-containing protein
Bra010239	A08	2.317 down	DNA-directed DNA polymerase/ damaged DNA binding
Bra010247	A08	1.146 down	Protein kinase family protein
Bra010253	A08	1.573 down	Unknown protein
Bra010272	A08	1.905 up	scpl29; scpl29 (serine carboxypeptidase-like 29); serine-type carboxypeptidase
Bra010326	A08	2.223 down	Mitochondrial ATP synthase g subunit family protein
Bra010360	A08	4.462 up	RabGAP/TBC domain-containing protein
Bra010499	A08	1.385 down	"VEP1, AWI31; VEP1 (VEIN PATTERNING 1); binding / catalytic"
Bra010542	A08	1.864 down	IPR004265; Plant disease resistance response protein
Bra010551	A08	3.319 down	"Disease resistance protein (TIR-NBS-LRR class), putative"
Bra010577	A08	1.305 down	"BLH2, SAW1; BLH2 (BEL1-LIKE HOMEODOMAIN 2); DNA binding / transcription factor"
Bra010625	A08	12.173 down	"ELI3-1, ELI3, ATCAD7, CAD7; ELI3-1 (ELICITOR-ACTIVATED GENE 3-1); binding / catalytic/ oxidoreductase/ zinc ion binding"
Bra010626	A08	11.368 down	ELICITOR-ACTIVATED GENE 3-1
Bra010658	A08	2.046 up	"Trehalose-6-phosphate phosphatase, putative"
Bra010694	A08	2.182 down	"Isoflavone reductase, putative"
Bra010724	A08	172.883 up	RNA recognition motif (RRM)-containing protein

Bra010727	A08	2.117 up	"PRP4, ATPRP4; PRP4 (PROLINE-RICH PROTEIN 4)"
Bra010817	A08	1.345 up	Protein kinase family protein
Bra010860	A08	81.415 down	AtRABA1i; AtRABA1i (Arabidopsis Rab GTPase homolog A1i); GTP binding
Bra010899	A08	53.943 down	Sulfotransferase family protein
Bra010984	A08	1.896 down	Unknown protein
Bra016280	A08	75.366 down	Zinc finger (C2H2 type) family protein
Bra016298	A08	1.898 down	Peptidyl-prolyl cis-trans isomerase cyclophilin-type family protein
Bra016325	A08	232.628 down	GO:0006952; defense response; Biological Process
Bra016469	A08	185.721 down	ATTIM17-1; ATTIM17-1; P-P-bond-hydrolysis-driven protein transmembrane transporter/ protein transporter
Bra016532	A08	1.601 down	Aminoacyl-tRNA synthetase family
Bra016547	A08	1.283 down	IBS1; IBS1 (IMPAIRED IN BABA-INDUCED STERILITY 1)
Bra016708	A08	1.934 down	"2-oxoglutarate-dependent dioxygenase, putative"
Bra016723	A08	1.499 down	DRL1; DRL1 (DEFORMED ROOTS AND LEAVES 1); calmodulin binding / purine nucleotide binding
Bra016729	A08	1.255 up	"GAPC, GAPC-1, GAPC1; GAPC1 (GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE C SUBUNIT 1)"
Bra016788	A08	2.080 down	Unknown protein
Bra030740	A08	2.488 down	"AIP3, PDF4; AIP3 (ABI3-INTERACTING PROTEIN 3); unfolded protein binding"
Bra030713	A08	33.117 down	ATNRT2:1, NRT2, NRT2;1AT, ACH1, LIN1, ATNRT2.1, NRT2.1; ATNRT2:1 (NITRATE TRANSPORTER 2:1)
Bra030684	A08	1.098 down	Transcription factor IIA large subunit / TFIIA large subunit (TFIIA-L)
Bra030672	A08	2.454 down	"STN1, ATSTN1; STN1; nucleic acid binding"
Bra030662	A08	2.094 down	RNA polymerase Rpb7 N-terminal domain-containing protein
Bra030648	A08	62.036 down	Unknown protein
Bra030644	A08	1.348 down	Unknown protein
Bra030638	A08	5.138 up	Fatty acid desaturase family protein
Bra030631	A08	1.251 down	Unknown protein
Bra030540	A08	1.405 up	POR C, PORC; POR C (PROTOCHLOROPHYLLIDE OXIDOREDUCTASE)
Bra037305	A09	1.443 down	CAT9; CAT9 (CATIONIC AMINO ACID TRANSPORTER 9); cationic amino acid transmembrane transporter
Bra037382	A09	1.618 down	"GBF2, ATBZIP54; GBF2 (G-BOX BINDING FACTOR 2); DNA binding / sequence-specific DNA binding / transcription factor"
Bra037392	A09	82.014 down	Unknown protein
Bra036230	A09	22.621 down	Ubiquitin family protein
Bra039084	A09	1.332 up	G6PD5; G6PD5 (glucose-6-phosphate dehydrogenase 5); glucose-6-phosphate dehydrogenase

Bra036203	A09	2.089 up	"CYP85A2, BR6OX2; BR6OX2 (BraSSINOSTEROID-6-OXIDASE 2); monooxygenase/ oxygen binding"
Bra026495	A09	1.202 down	Unknown protein
Bra035938	A09	1.158 down	Zinc finger protein-related
Bra035908	A09	169.092 down	GO:0016998; cell wall macromolecule catabolic process; Biological Process
Bra037749	A09	1.321 down	"Oxidoreductase, zinc-binding dehydrogenase family protein"
Bra037275	A09	1.280 down	ARID/BRIGHT DNA-binding domain-containing protein
Bra037237	A09	1.407 down	Unknown protein
Bra027758	A09	165.431 down	Unknown protein
Bra027071	A09	1.610 down	Unknown protein
Bra027867	A09	1.478 up	"RXF12, ATXYN1; RXF12; endo-1,4-beta-xylanase/ hydrolase, hydrolyzing O-glycosyl compounds"
Bra027921	A09	51.340 down	AtRLP46; AtRLP46 (Receptor Like Protein 46); kinase/ protein binding
Bra027941	A09	1.337 down	emb1075; emb1075 (embryo defective 1075); carboxy-lyase/ catalytic/ pyridoxal phosphate binding
Bra027497	A09	1.507 up	"2-oxoglutarate-dependent dioxygenase, putative"
Bra017585	A09	1.740 up	"ATPTR3, PTR3; PTR3 (PEPTIDE TRANSPORTER 3); dipeptide transporter/ transporter/ tripeptide transporter"
Bra017583	A09	2.844 up	"ATPTR3, PTR3; PTR3 (PEPTIDE TRANSPORTER 3); dipeptide transporter/ transporter/ tripeptide transporter"
Bra017395	A09	2.176 down	PHO1_ARATH Putative phosphate transporter
Bra017392	A09	1.456 down	"DNA-binding protein, putative"
Bra017333	A09	4.855 up	CYP86C2; CYP86C2; electron carrier/ heme binding / iron ion binding / monooxygenase/ oxygen binding
Bra035269	A09	1.252 down	Unknown protein
Bra029539	A09	1.276 up	"Lectin protein kinase, putative"
Bra028030	A09	50.182 up	Alliinase family protein
Bra032294	A09	2.745 down	Unknown protein
Bra032920	A09	199.836 down	TPR2; TPR2 (TOPLESS-RELATED 2)
Bra024770	A09	1.532 down	Acidic ribosomal protein P0-related
Bra024766	A09	2.605 down	Unknown protein
Bra024705	A09	118.285 down	CLE9; CLE9 (CLAVATA3/ESR-RELATED 9); protein binding / receptor binding
Bra024660	A09	1.678 down	Galactosyltransferase family protein
Bra024630	A09	32.257 down	Unknown protein
Bra031365	A09	43.238 down	Structural constituent of ribosome
Bra036043	A09	1.401 down	Protein binding / signal transducer
Bra036786	A09	1.359 up	Unknown protein
Bra006884	A09	1.832 down	Unknown protein
Bra007119	A09	1.801 down	"PLP7, PLA IIIA; PLA IIIA (PATATIN-LIKE PROTEIN 6)"

Bra007372	A09	1.268 down	Unknown protein
Bra007390	A09	1.524 up	Meprin and TRAF homology domain-containing protein / MATH domain-containing protein
Bra007515	A09	4.261 down	ATL4; ATL4; protein binding / zinc ion binding
Bra007595	A09	1.444 down	DNA-binding family protein
Bra007624	A09	4.356 down	Unknown protein
Bra007739	A09	72.902 down	HFA7B_ARATH Heat stress transcription factor
Bra007772	A09	64.514 down	Bet v I allergen family protein
Bra007787	A09	1.403 down	IPR002885; Pentatricopeptide repeat
Bra007824	A09	80.216 down	CPI1; CPI1 (CYCLOPROPYL ISOMERASE); cycloeucalenol cycloisomerase
Bra039221	A09	1.369 down	"ATCNGC6, CNGC6; ATCNGC6; calmodulin binding"
Bra039184	A09	39.847 down	Unknown protein
Bra031233	A09	1.120 up	PUB13_ARATH U-box domain-containing protein
Bra031151	A09	1.320 down	TPC2L_DICDI Trafficking protein particle complex subunit 2-like protein
Bra026708	A09	1.618 down	HAC12; HAC12 (histone acetyltransferase of the CBP family 12)
Bra026861	A09	1.945 down	"FUT6, ATFUT6; FUT6 (FUCOSYLTRANSFERASE 6); fucosyltransferase/ transferase, transferring glycosyl groups"
Bra026922	A09	131.109 down	RPS5_ARATH Disease resistance protein
Bra026946	A09	1.192 down	Importin-related
Bra026968	A09	1.623 down	ATCCS; ATCCS (COPPER CHAPERONE FOR SOD1); superoxide dismutase/ superoxide dismutase copper chaperone
Bra031611	A09	19.088 down	ATNRT2:1, NRT2, NRT2;1AT, ACH1, LIN1, ATNRT2.1, NRT2.1; ATNRT2:1 (NITRATE TRANSPORTER 2:1)
Bra031600	A09	1.572 down	"ATRNL, RNL; RNL (RNA LIGASE); 2',3'-cyclic-nucleotide 3'-phosphodiesterase/ RNA ligase (ATP)/ polynucleotide kinase"
Bra032446	A09	54.797 down	Unknown protein
Bra032469	A09	4.501 up	Pectinesterase family protein
Bra032483	A09	1.524 down	Unknown protein
Bra015620	A10	81.325 down	Unknown protein
Bra033399	A10	1.161 down	Leucine-rich repeat family protein
Bra033328	A10	2.531 down	Glycoside hydrolase family 28 protein / polygalacturonase (pectinase) family protein
Bra033303	A10	39.902 down	ANAC004; ANAC004 (Arabidopsis NAC domain containing protein 4); transcription factor
Bra033277	A10	2.147 down	"ATRABA3, ATRAB-A3; ATRABA3 (ARABIDOPSIS RAB GTPASE HOMOLOG A3); GTP binding"
Bra028298	A10	1.194 down	"ATARLB1, ATGB1; ATGB1 (ARABIDOPSIS THALIANA GTP-BINDING PROTEIN 1); GTP binding"
Bra003080	A10	1.663 up	"Protein phosphatase 2C, putative / PP2C, putative"
Bra003069	A10	1.314 down	PDV1; PDV1 (PLASTID DIVISION1)

Bra002990	A10	3.612 down	Unknown protein
Bra002915	A10	137.204 down	Long-chain-alcohol O-fatty-acyltransferase family protein / wax synthase family protein
Bra002814	A10	24.148 down	FAD-binding domain-containing protein
Bra002692	A10	1.491 up	"FLP1, YRE, CER3, WAX2; CER3 (ECERIFERUM 3); binding / catalytic/ iron ion binding / oxidoreductase"
Bra002539	A10	95.419 down	HSP18.2; HSP18.2 (heat shock protein 18.2)
Bra002501	A10	22.894 down	Zinc finger (C3HC4-type RING finger) family protein
Bra002327	A10	1.118 down	ARFG_ARATH Auxin response factor
Bra002310	A10	4.301 down	F-box family protein
Bra002138	A10	1.538 down	Unknown protein
Bra002117	A10	11.227 down	"Disease resistance protein (TIR-NBS-LRR class), putative"
Bra008670	A10	179.877 down	K09060; GBF3; GBF3 (G-BOX BINDING FACTOR 3); transcription factor ; plant G-box-binding factor
Bra008828	A10	290.523 down	Unknown protein
Bra008836	A10	1.154 down	IPR004993; GH3 auxin-responsive promoter
Bra008916	A10	1.367 down	Pentatricopeptide (PPR) repeat-containing protein
Bra009033	A10	1.999 down	IPR002528; Multi antimicrobial extrusion protein MatE
Bra009061	A10	3.388 down	Unknown protein
Bra009291	A10	1.216 down	SNX2b; SNX2b (SORTING NEXIN 2b); phosphoinositide binding / protein binding
Bra009228	A10	38.778 up	Myb family transcription factor
Bra009411	A10	1.563 down	Vacuolar protein sorting 36 family protein / VPS36 family protein
Bra009535	A10	225.800 down	Auxin-responsive family protein
Bra009580	A10	1.423 down	Short-chain dehydrogenase/reductase (SDR) family protein
Bra009618	A10	1.724 up	"LSN, PNY, HB-6, BLR, RPL, BLH9, VAN; RPL (REPLUMLESS); Sequence-specific DNA binding / transcription factor"