A Multi-omics approach for genetic and molecular analysis of clubroot resistance in *Brassica napus* canola

by

Kawalpreet Kaur

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Plant Science

Department of Agricultural, Food and Nutritional Science
University of Alberta

Abstract

Clubroot disease, caused by *Plasmodiophora brassicae*, is a significant threat to canola production. Growing of clubroot resistant canola cultivars is the most efficient and environment-friendly way of managing this disease. The evolution of new pathotypes eroded some of the resistance available in canola. This necessitates the finding of new resistance and introgression into canola. In this PhD thesis research, clubroot resistance of the European fodder turnip cv. Debra (ECD 01) was introgressed into *B. napus* canola through interspecific hybridization between these two species, and an advanced generation of *B. napus* canola recombinant inbred line (RIL) population (BC_1F_8) was developed. This population carried resistance to multiple pathotypes including pathotype 3H, 3A, 2B and 5X. Whole-genome resequencing of the RILs carrying resistance to pathotype 3H or 3A was performed to identify the genomic regions contributing to clubroot resistance. Based on this finding, and through different QTL mapping approaches, three genomic regions located at 1-4 Mb region of the chromosome A02, 0.98-2.0 Mb region of A03 and 38.01-41.5 Mb region of A09 were identified showing resistance to pathotype 3; among these, the A09 locus exerted the greatest contribution for resistance to pathotype 3A.

To understand the molecular basis of the Debra-resistance in canola, near-isogenic clubroot resistant and susceptible lines were developed based on this resistance and were used for transcriptome and proteome analysis. These analyses identified several putative genes/proteins related to cell surface receptors (receptor-like kinases, receptor-like proteins), phytohormone signalling (auxin, salicyclic acid, jasmonic acid), calcium-mediated signalling (calmodulin-gated calcium channels, calcium-dependent protein kinases), reactive oxygen species, glucosinolate biosynthesis pathway as well as genes/proteins related to disease resistance.

In addition to this, an investigation on the C2H2-zinc finger protein coding genes (transcription factors) of the *B. napus* genome was carried out. This study identified 267 genes carrying cis-acting elements in their promoter regions; and all these genes showed distribution across the 19 chromosomes of *B. napus*. Expression analysis of 20 C2H2-zinc finger genes demonstrated that this class of genes are involved in defense response against abiotic and biotic stresses, including clubroot disease. This analysis further demonstrated that some of the genes may be involved in defense response against multiple stresses and in response to more than one phytohormones.

Thus, this thesis research provided solid evidence that the clubroot resistance of the turnip accession ECD 01 can be used for broadening the genetic base of clubroot resistance in *B. napus* canola for durable resistance to this disease. The multi-omics approaches applied in this research helped in the identification of the genomic regions, genes and proteins playing a role in the molecular mechanisms involved in clubroot resistance. These results also suggested that the C2H2-zinc finger genes can be targeted for the improvement of stress tolerance in canola.

Preface

This dissertation is submitted by Kawalpreet Kaur for the degree of Doctor of Philosophy. Kawalpreet Kaur conducted all the experiments, collected and analyzed all the data, compiled the draft of all the chapters of this dissertation and incorporated the valuable comments and suggestions from her supervisor Dr. Habibur Rahman and committee members Drs. Urmila Basu, Nat Kav and Glen Uhrig.

A version of Chapter 1 of this dissertation has been published as Kaur, K., Liu, Y., and Rahman, H. 2022. Introgression of resistance to multiple pathotypes of *Plasmodiophora brassicae* from turnip (*Brassica rapa* ssp. *rapifera*) into spring *B. napus* canola. Agronomy. 12(5):1225. doi:10.3390/agronomy12051225.

Kawalpreet Kaur carried out the experiments, collected and analyzed the data and drafted the manuscript. Yingyi Liu carried out the experiment for resistance to multiple pathotypes. Dr. Habibur Rahman conceived the research idea, supervised the experiments and edited the manuscript.

A version of Chapter 6 has been published as Kaur, K., Megha, S., Wang, Z., Nat, N. V. Kav., and Rahman, H. 2023. Identification and expression analysis of C2H2-zinc finger protein genes reveals their role in stress tolerance in *Brassica napus*. Genome. 66(5): 91-107. doi: 10.1139/gen-2022-0100.

Kawalpreet Kaur conducted the research experiments including data curation and writing the first draft of the manuscript. Dr. Swati Megha conducted biotic, abiotic and hormone application experiments. Zhengping Wang did the analysis for syntenic relationships. Dr. Nat Kav

provided input to cold and sclerotinia stress experiments and gave feedback on the manuscript. Dr. Habibur Rahman helped to shape the research, data analysis and finalizing the manuscript.

In addition, Dr. Urmila Basu provided valuable feedback, reviewed and edited the draft version for Chapter 4. Dr. Glen Uhrig conducted the LC-MS/MS analysis and wrote the corresponding materials and methods for Chapter 5. Dr. Dinesh Adhikary helped in data analysis for Chapter 5.

Acknowledgements

First and foremost, I am deeply indebted to my supervisor, Dr. Habibur Rahman for believing in me, giving me insightful feedback and pushing me to develop as a scientist. My heartfelt gratitude also goes to supervisory committee members: Drs. Urmila Basu, Dr. Nat Kav and Dr. R. Glen Uhrig, who generously provided knowledge and expertise. I offer my most sincere appreciation to Dr. Stephen Strelkov and Dr. Enrico Scarpella for being in my candidacy examination committee.

Many others have been instrumental in this process. I am also grateful to Dr. Swati Megha, Dr. Aleya Ferdausi, Dr. Berisso Kebede, Dr. Muhammad Jakir Hasan and other members of Canola Breeding Program at University of Alberta. This endeavour would not have been possible without their ever-willing help and technical support. I would like to acknowledge my friends, Dr. Harleen Kaur and Karanjot Singh for an emotional support and always cheering me on. I would also like to thank Nikki Scott and Kelly Dunfield for their support in my greenhouse experiments. I would also like to acknowledge Canola Agronomic Research Program (CARP) (Alberta Canola, SaskCanola and Manitoba Canola Growers), Alberta Innovates Bio Solutions and Agriculture and Agri-Food Canada for funding my research project.

Lastly, I would be a remiss in not mentioning my loving family, especially my parents. Thank you for encouraging me to pursue my passions. I hope I have made you proud. Special thanks to my husband, Ramanjeet Singh Brar, for his constant reassurance and celebrating my accomplishments alongside me. Your support is what has gotten me through when I wanted to give up.

Table of Contents

List of Tables

List of Figures

List of Symbols and Abbreviations

=	Equal
2*n*	Diploid number of chromosomes
>	Greater than
%	Percent
°C	Degree Celsius
Δ	Delta
/	Per
μmol g^{-1}	Micromoles per gram
g^{-1}	Per gram
cv.	cultivar
ABA	Abscisic acid
ANOVA	Analysis of variance
BC_1	First backcross generation
BC_1F_x	X^{th} generation of BC_1-derived population
bp	Base pairs
BR	Brassinosteroid
BSA	Bulked segregant analysis
BSR-seq	Bulked segregant RNA sequencing
CC	Coiled coil
CCD	Canadian clubroot differential
cm	Centimetre
CPDS	Canadian plant disease survey
CR	Clubroot resistance
dai	Days after inoculation
DEGs	Differentially expressed genes

DGE	Digital gene expression
DNA	Deoxyribonucleic acid
dpi	Days post inoculation
DSI	Disease severity index
ECD	European clubroot differential
ETI	Effector-triggered immunity
F_1	First generation
GC	Guanine cytosine
hai	Hours after inoculation
Indels	Insertion/deletions
JA	Jasmonic acid
kg/ha	Kilogram per hectare
LRR	Leucine-rich repeats
MAS	Marker assisted selection
Mb	Million base pair
mL	Millilitre
MMT	Million metric tons
MR	Moderately resistant
MS	Moderately susceptible
NBS	Nucleotide-binding site
NGS	Next generation sequencing
NIL	Near-isogenic line
NIRS	Near-infrared spectroscopy
ORF	Open reading frame
PCA	Principal Component Analysis
PCR	Polymerase chain reaction
PFGE	Pulsed field gel electrophoresis

PTI	PAMP-triggered immunity
QTL	Quantitative trait loci
R genes	Resistance genes
RIL	Recombinant inbred line
RLK	Receptor like kinase
RLP	Receptor like protein
RNA	Ribonucleic acid
S	Susceptible
SA	Salicylic acid
SNP	Single nucleotide polymorphism
TFs	Transcription factors
TIR	Toll-interleukin-1
UTR	Untranslated region
WGRS	Whole-genome resequencing
ZFPs	Zinc-finger proteins

Chapter 1. Literature review

1.1 Introduction to *Brassica* genus

Oilseed rape, also known as rapeseed or canola, belongs to the genus *Brassica* of the family *Brassicaceae* (*Cruciferae*) that comprises about 372 genera and 3700 species, found mainly in northern hemisphere. Many *Brassica* species have been cultivated since prehistoric times for their edible parts. *Brassica napus* L. ($2n = 38$, AACC genome), *Brassica rapa* L. ($2n = 20$, AA genome) and *Brassica juncea* L. Czern. ($2n = 36$, AABB genome) are the most important oilseed crops of the family *Brassicaceae*. The A-genome species *B. rapa* includes turnip, Chinese cabbage, and oilseed type. Members of *B. napus* mainly include rutabaga and oilseed type, while *B. juncea* mainly includes condiment mustard and oilseed (Indian brown mustard) types. *B. rapa* and *B. napus* are collectively often called rapeseed. Based on the growth habit, *B. napus* is categorised into winter, spring and semi-winter types. The winter type requires vernalization for about eight weeks for flowering and is primarily grown in Europe; the semi-winter type requires a shorter period of vernalization (about four weeks) and is primarily cultivated in China, while the spring type does not require vernalization for flowering and is mainly grown in Canada, Australia, and the USA (Raymer 2002).

Historical evidence suggests that *Brassica* crops have been cultivated and consumed in India since 1500 to 2000 BC (Gupta and Pratap 2007). Its introduction in Europe dates back to 13th century due to its ability to grow at relatively low temperatures. Seeds of traditional *Brassica* oilseed contained a high level of erucic acid (>40% of total fatty acids) in oil and a high level of glucosinolates (GSL) (>100 μmol g^{-1} seed) in seed meal; this made the oil unsuitable for use in edible purposes and the protein-rich seed meal unsuitable for use in animal feed (Hung et al. 1977;

Mawson et al. 1993). Therefore, the traditional oil was primarily used as a lubricant in Europe; however, this oil has been used as edible oil in Asia since the Middle Ages (5^{th} to 15^{th} century). During the Second World War, when the European and Asian oil supplies were blocked, *Brassica* oilseed crops were introduced in Canada as an emergency measure. Production of this crop in Canada increased quickly in a few years; however, soon after the war, it almost disappeared from the country due to the replacement of steam engines by diesel engines (Casséus 2009). The ability of this crop to withstand and grow at lower temperatures makes it suitable for cultivation in Canada (Shahidi 1990). To use this oil for edible purposes and the seed meal for animal feed, it was essential to remove or lower the two main impediments, viz., the erucic acid and GSL contents, from the seeds. Intensive breeding and research by Canadian researchers led to the development of canola (Canadian Oil Low Acid) quality cultivars in 1970's. Western Canadian Oilseed Crushers trademarked the name 'Canola' (less than 2% erucic acid in oil and less than 30 μmol GSL g^{-1} of seed meal) in 1978. 'Double low' or canola quality spring type cv. Tower carrying the low erucic acid genes introduced from German spring forage rape cv. Liho (for review, see Przybylski and Mag 2011) and low GSL genes from Polish forage cv. Bronowski (Olsen and Sorenson 1980) was released in 1974 (for review, see Friedt et al. 2018). Canola, developed in Canada about 50 years ago, is now at the centre of a burgeoning industry with an international reach, accounting for 86.31 million metric tons (MMT) of global production in 2022-23 (Figure 1.1) (USDA-FAS 2023).

Canola production can be affected by various diseases, such as, clubroot, alternaria black spot, blackleg, sclerotinia stem rot, and white rust caused by *Plasmodiophora brassicae* Woronin, *Alternaria brassicae* (Berk.) Sacc., *Leptosphaeria maculans* (Sowerby) P. Karst., *Sclerotinia sclerotiorum* (Lib.) de Bary, *Albugo candida* (Pers.) Kuntze, respectively. The occurrence and

severity of these diseases depends on the extent of resistance in the plants, and virulence and diversity of the pathogen races as well as on environmental conditions. The strategies for containment of the diseases can vary depending on the pathogen; however, a combination of cultivar resistance and crop production practices are generally used in an integrated pest management practice to secure the production of the crop. Among the different diseases, as mentioned above, the soil-borne disease clubroot poses the most serious threat to canola production in Canada (for review, see Howard et al. 2010); this crop contributes about $29.9 billion to the national economy (Canola Council of Canada 2020). To manage this disease, an integrated multifaceted approach has been suggested by Peng et al. (2014a); in this approach and growing of clubroot resistant cultivars has been considered to play the most important role (Peng et al. 2014a).

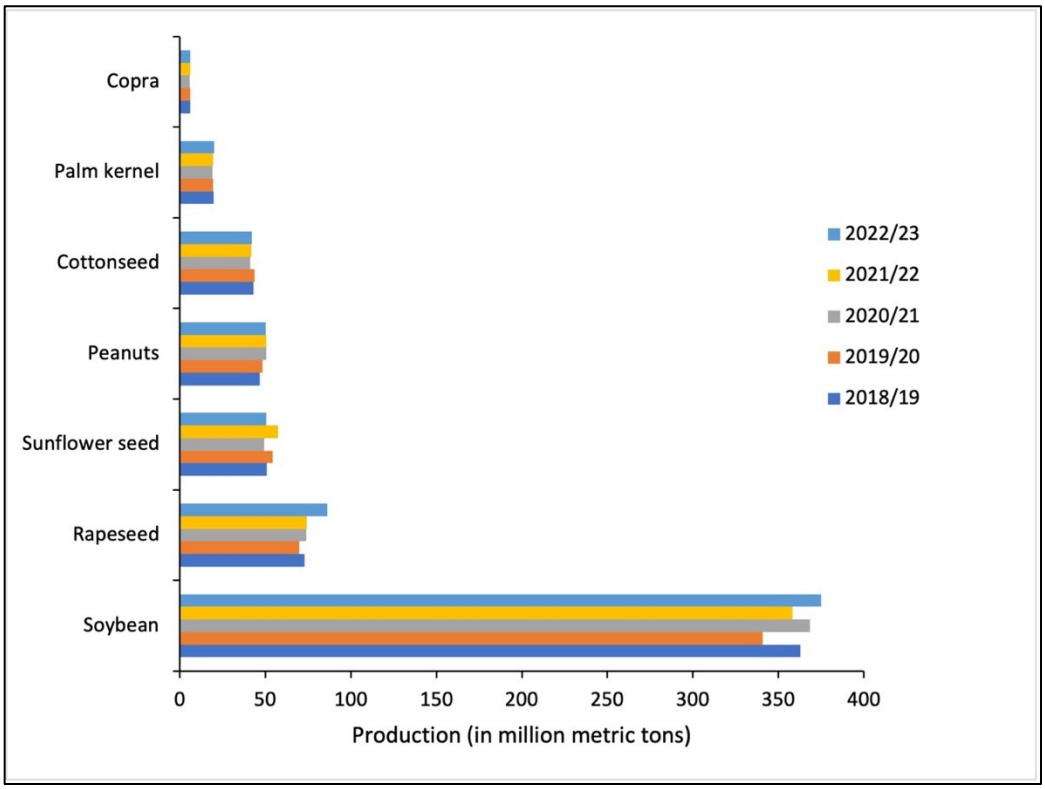


Figure 1.1 The global production of the main oilseed crops in 2022/23 (adapted from USDA-FAS 2023) including canola (rapeseed).

1.2 Clubroot

1.2.1 Introduction to clubroot

Clubroot disease is known to impact cruciferous crops since 13th century. Its causal agent *Plasmodiophora brassicae* Woronin is an obligate biotrophic protist belonging to the supergroup Rhizaria within the phylum Cercozoa and the Endomyxa (for review, see Neuhauser, Kirchmair and Gleason 2011). This pathogen attacks the root hair, which ultimately results in the malformed roots called 'galls'. These club shaped galls severely reduce the uptake of nutrients and water by the roots resulting in wilting of the plants (Kageyama and Asano 2009).

1.2.2 History of clubroot

Clubroot is one of the oldest diseases of *Brassica* crops. Pallidus reported, for the very first time, the occurrence of spongy roots on rape, turnip, and radishes in Italy in 4th century (cited by Watson and Baker 1969). The second record of this disease was provided by Diaz de Isla in 1539 in Spain (for review, see Howard et al. 2010). These records provide the evidence that clubroot disease might have originated in the Mediterranean region, close to the centre of origin of the genus *Brassica*. According to Dixon (2009), the historical information on the occurrence of clubroot disease in Europe can be traced back to 13th century. In late 19th century, a severe epidemic of this disease destroyed a large proportion of the cabbage fields in St. Petersburg. This disease is known to be present in about 60 countries causing about 10-15% yield loss on a global scale (for review, see Dixon 2009). Apart from the yield losses, clubroot also adversely affects the quality of Brassica oilseeds by decreasing the seed oil content by about 2-6% (Engqvist 1994; Pageau et al. 2006) and increasing the chlorophyll content by 50% (Engqvist 1994).

The exact time of introduction of the pathogen *P. brassicae* in Canada is unknown; however, it is thought to have been introduced with fodder turnips brought by the early European

settlers to feed their livestock. This disease has been reported in vegetable production areas of British Columbia, Quebec, Ontario, and Atlantic provinces since 1920s (for review, see Howard et al. 2010). In 1956, clubroot disease gained a national attention and was mentioned in Canadian Plant Disease Survey (CPDS) as a 'noteworthy disease' (for review, see Howard et al. 2010). The first incidence of this disease in commercial canola fields was reported in Sturgeon Country, northwest of Edmonton, Alberta in 2003 (Tewari et al. 2005). Field survey indicates that clubroot disease can be widespread and become a serious problem in cruciferous crops in a short period of time, and this raised a significant concern to Canadian canola growers due to its economic impact. Yield loss due to this disease in canola has been estimated at about 30%; however, complete failure of the crop has also been reported in heavily infested fields in Alberta (Tewari et al. 2005; for review, see Strelkov and Hwang 2014).

1.2.3 Host Range and taxonomy of *P. brassicae*

P. brassicae infects most of the species of the family Brassicaceae; however, only a few studies have been conducted outside the genera *Brassica*, *Raphanus* and *Arabidopsis* (Dixon 2009). The universal hosts of this pathogen are the cultivated *Brassica* species, including *B. oleracea* Brussels sprouts, cabbage, cauliflower, kale, and kohlrabi), *B. rapa* (turnip, turnip rape and Chinese cabbage), *B. napus* (rutabaga, oilseed rape or canola), *Raphanus sativus* L., *B. carinata* L., *B. nigra* L. and *B. juncea*. Some of the cruciferous weeds, such as shepherd's purse (*Capsella bursapastoris* L.), field pepperwort (*Lepidium latifolium* L.), flixweed (*Descurainia sophia* L.) and stinkweed (*Thlaspi arvense* L.) are also susceptible to *P. brassicae* (Buczacki and Ockendon 1979). This can reduce the effectiveness of crop rotation as a strategy for the control of clubroot disease.

Historically, *P. brassicae* was placed in the genus *Plasmodiophora* of the order *Plasmodiophorales* together with seven other genera (Karling 1969). Other species of the genus

Plasmodiophora are *P. diplantherae*, *P. halophilae*, *P. fici-repentis* and *P. bicaudate*. Later, *P. brassicae* was classified in the order *Plasmodiophorida* based on the protein sequence data (Neuhauser et al. 2010). Currently, it is included in the supergroup *Rhizaria*, phylum *Cercozoa*, class *Phytomexa*, and order *Plasmodiophorida* (Neuhauser et al. 2010). Due to the incomplete understanding of the *P. brassicae* life cycle, accurate taxonomic classification has been a challenge (for review, see Donald and Porter 2014).

1.2.4 Life Cycle of *P. brassicae*

P. brassicae exhibits a complex life cycle comprising of three different stages: Survival as resting spores, primary infection, and secondary infection stages. The resting spores, which are the source of primary inoculum for infection, can survive in soil for about 17 years with an approximate half-life of about four years (Wallenhammer 1996). The long viability of the resting spores is due to the unique composition of their cell wall which is composed of 25% chitin, protein, and lipids (Moxham et al. 1983; Rolfe et al. 2016).

Germination of the resting spore results in biflagellate motile spore, called primary zoospore. Germination of the resting spores may be enhanced by the presence of root exudates (Friberg et al. 2005). The zoospores require water films in soil to swim to the root hair of the host plants. They penetrate the cell wall of the root hair or the root epidermal cells and inject their protoplast into cytoplasm causing primary infection, where they undergo endomitotic divisions and form multinucleate primary plasmodium (Liu et al. 2020). The primary plasmodium, in turn, cleaves into zoosporangia, each containing 4-16 secondary zoospores which are released into the soil (Naiki et al. 1984). Primary infections do not cause macroscopic symptoms or significant yield and quality losses.

The secondary zoospores invade the cortical tissue of the main root and develop into a plasmodium (Figure 1.2). The formation of secondary plasmodia alters the level of plant hormones causing hyperplasia and hypertrophy and results in a recognizable symptom of the disease i.e., club-shaped malformed roots, which ultimately disintegrate and release millions of resting spores in the soil (for review, see Howard et al. 2010; Hwang et al. 2012). It has been estimated that as many as 16 billion resting spores can be produced in one large gall (Hwang et al. 2012). The activation of auxin (Jahn et al. 2013), cytokinin (Lan et al. 2019) and possibly the brassinosteroid (BR) signaling pathways (Schuller et al. 2014) is reported when the plasmodia pass and establish themselves in the plant cells. The abnormal cell division and enlargement of roots leads to increased allocation of photosynthetic products from leaves to the roots (Ludwig-Müller et al. 2009). The pathogen usually completes one cycle per season and cannot spread rapidly in the soil as motility of the zoospores is limited. *P. brassicae* spreads over long distances via movement of the resting spores through soil or infected plant material (for review, see Hwang et al. 2014).

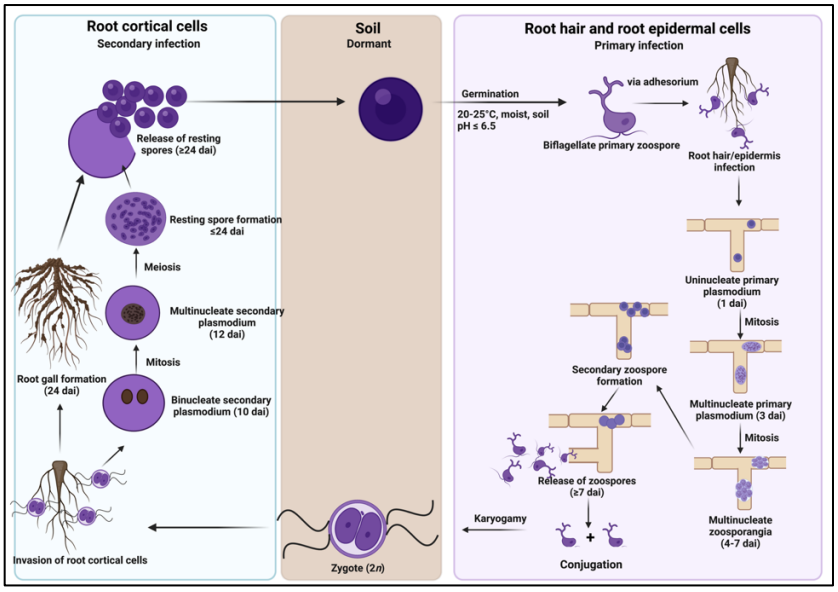


Figure 1.2 The life cycle of clubroot disease causal pathogen *Plasmodiophora brassicae* in the cruciferous plants. Dai: days after inoculation; 2*n* = diploid number of chromosomes.

1.2.5 Pathotype classification of *P. brassicae*

Pathotypes refer to the isolates of *P. brassicae* collected from infested soil or plant material that exhibit different virulence on different hosts (Buczacki and Ockenson 1979). Many pathotypes or races exist in clubroot pathosystem due to mutation or genetic changes which occur in chromosomes due to crossover during meiosis. This demands the need for differentiating the pathotypes for their pathogenic diversity while developing clubroot resistant cultivars. Various host-plant differential sets have been proposed over the past 60 years to identify the pathotypes of *P. brassicae*. Williams (1966), Some et al. (1996) and the European Clubroot Differential (ECD) are the most widely used systems (Buczacki et al. 1975). The differential hosts of Williams include two rutabaga (*B. napus* var. *napobrassica*) cvs. Laurentian and Wilhelmsburger, and two cabbage (*B. oleracea* var. *capitata*) cvs. Badger Shipper and Jersey Queen. These hosts can, theoretically, distinguish a maximum of 16 pathotypes, however, they may not be suitable to differentiate all the pathotypes collected from a canola field (for review, see Strelkov and Hwang 2014). The ECD set was proposed in 1970s, which includes 15 hosts – five from each of *B. rapa*, *B. napus* and *B. oleracea* (Buczacki et al. 1975). Due to inclusion of several redundant hosts, this set has not been used extensively outside Europe to characterize the pathotypes. The third system, developed in France by Some et al. (1996), includes only three *B. napus* genotypes; it also lacks the capacity of differentiating a large number of *P. brassicae* pathotypes (Strelkov et al. 2018).

At the initial stage of the establishment of clubroot disease in canola in Canada, five pathotypes, *viz.*, 2, 3, 5, 6 and 8 have been identified (Strelkov et al. 2006, 2007; Xue et al. 2008; Cao et al. 2009). Recently, several new pathotypes have been isolated from canola fields in western Canada (Strelkov et al. 2016b). It was difficult to distinguish all new pathotypes from the old ones by using the above-mentioned differential sets; therefore, a new Canadian Clubroot Differential

(CCD) set was developed by using 13 hosts (differentials of Williams et al. 1966 and Some et al. 1996, and few selected hosts of ECD, and canola cultivars Mendel, Westar and 45H29) (Strelkov et al. 2018). In this system, a total of 17 pathotypes *viz.*,2B, 2F, 3A, 3D, 3H, 3O, 5C, 5G, 5I, 5K, 5L (D-G3), 5X (LG-1, LG-2, LG-3), 6M, 8E, 8J, 8N and 8P have been listed. Based on the CCD set, a total of 37 (5 original and 32 novel) *P. brassicae* pathotypes were detected in Western Canada (Strelkov et al. 2020; Askarian et al. 2021; Hollman et al. 2021). The CCD set include all hosts of the Williams and Some et al. (1996)'s set, therefore, it is also possible to assign pathotype designations according to these systems. For example, pathotypes A and H of the CCD system share similar virulence patterns but differ for their ability to cause disease to the clubroot resistant canola cv. 45H29 (pathotype A is virulent to 45H29, while H is avirulent to this host) (Strelkov et al. 2018). Both pathotypes, however, were classified as pathotype 3 on the Williams differentials (1966); as such, these two pathotypes in the CCD system referred to as pathotypes 3A and 3H.

1.2.6 Molecular biology of *P. brassicae*

The whole genome of *P. brassicae* was first sequenced by Schwelm et al. (2015) by using genomic DNA of resting spores from the single spore isolate e3. The size of this genome was reported to be 25.5 Mb which was assembled *de novo* into 165 scaffolds with N50 size of 473 kb. The GC content in the genome is 58.5%. With the support of seven transcriptome libraries, a total of 9,730 genes were predicted representing 66% of the entire genome with an average of 404 genes per Mb. Each gene contains about 4 introns of 60 bp size.

To understand the genome diversity of various *P. brassicae* pathotypes, the genomes of Pb3 and Pb6 pathotypes were assembled *de novo,* and the pathotypes Pb2, Pb5 and Pb8 were re-sequenced (Rolfe et al. 2016). The whole genome shotgun sequencing using a Roche 454 GS FLX Titanium platform assembled the sequence reads into 109 scaffolds for Pb3 and 356 scaffolds for

Pb6. The genome size of these two pathotypes was found to be similar (approximately 24.2 Mb); the total number of predicted genes was 10,851 in Pb3 and 10,070 in Pb6. Most of the genes in Pb6 were orthologous to Pb3 except for 14 unique genes. Through alignment of the Pb6 genome and the re-sequenced pathotypes against the Pb3 reference genome, a total of 173,134 single nucleotide polymorphisms (SNPs) were detected. A comparison of SNP profile of Pb3 with four pathotypes showed a clear difference between the Pb3 and Pb6 while the SNP profile of Pb2, Pb5 and Pb8 was very similar (Rolfe et al. 2016).

1.2.7 Epidemiology of *P. brassicae*

Several environmental factors govern successful infection of the host plants by *P. brassicae*; uneven distribution of these environmental conditions can result in scattered infested patches in field.

Temperature and pH: A very early study on the effect of temperature indicated that clubroot disease can develop at soil temperatures ranging from 9 °C to 30 °C, where the optimum temperature is 25 °C (for review, see Dixon 2009); maximum disease severity in cabbage, Chinese cabbage, mustard and radish has been found at 21-22 °C. Studies conducted under controlled environmental conditions have also depicted that the optimum temperature for primary infection to disease symptom development is about 25 °C (Sharma et al. 2011a, 2011b). Therefore, growing of Brassica plants under a condition where soil temperature at seedling stage is low may help to reduce the severity of this disease.

Soil pH less than 7.0 is favourable for clubroot disease development. Therefore, application of lime to raise soil pH has been recommended to control this disease (Webster and Dixon 1991). The optimum pH for maximum disease incidence and severity is 5.0-6.5 when the temperature and moisture conditions are optimum.

Soil type and compaction: Soil organic matter may not influence the severity of clubroot disease as high severity of this disease has been reported in soils with both low and high organic matter (Wallenhammer 1996). Compact soil with poor drainage capacity and high moisture content for a longer period can increase the severity of this disease. The high moisture in soil is also needed for the movement of primary and secondary zoospores to the nearby roots.

1.3 Management of clubroot disease

1.3.1 Cultural practices

Sanitation: Sanitation of farm machinery and equipment, which have been in contact with clubroot infested soil, can reduce the movement of this disease from one field to the other (for review, see Howard et al. 2010). The three main steps for proper sanitation of the machineries are: (i) removal of bulk soil and crop debris, (ii) pressure washing or scrubbing to remove the residual soil, and (iii) application of disinfectant.

Crop rotation: It is the most effective cultural approach for the containment of this disease; a 3- to 4-year rotation of canola and non-*Brassica* crop is generally recommended (Peng et al. 2015; Ernst et al. 2019). However, a rotation of canola in every four years, or even a longer rotation, may not significantly reduce the resting spore population in soil when weeds susceptible to clubroot are present in the field (Strelkov et al. 2006; for review, see Howard et al. 2010); however, the severity of this disease under this condition can be substantially reduced (Peng et al. 2015).

Soil amendments: As mentioned above, soil pH lower than 7.0 is favourable for clubroot disease development; therefore, increasing soil pH by adding lime, wood ash or calcium cyanamide has been found to be effective for the management of this disease (Dixon 2012a). The application of 7.5-8.5 tons (t) lime per hectare (ha) has been found to reduce the severity of this disease in Central Alberta (Hwang et al. 2012). It has also been reported that repeated application of lime can reduce

the uptake of various nutrients, such as phosphorus, boron, magnesium, manganese, and zinc by the plants (McDonald et al. 2004); therefore, this factor also needs to be considered while applying limes. Mineral nutrients such as calcium, boron and nitrogen can also affect clubroot severity (Tremblay et al. 2005). Calcium cyanamide, a slow-release nitrogen fertilizer and herbicide, has also been used in the last 60 years to control this disease in vegetable Brassicas in different countries including Europe (Dixon 2017); a high content of calcium increases soil pH, and this makes the soil less favourable for pathogen growth (Tremblay et al. 2005).

Roguing and use of bait crops: Rouging (removal of infected plants) was recommended in early studies from the 1940s and 1950s (Donald and Porter 2009). Removal of the diseased roots can help to prevent the release of additional inoculums in soil.

As mentioned above, *P. brassicae* resting spores can survive in soil for more than 15 years; however, the germinated spores survive only for a shorter period in the absence of host plant (Suzuki et al. 1995). Therefore, inducing germination of the resting spores by growing a bait crop and ploughing the crop down before completion of the pathogen's life cycle can help to reduce the inoculum load in soil (Friberg et al. 2006). Different crops such as leek (*Allium ampeloprasum* L. var. *porrum* Gay), rye (*Secale cereale* L.), perennial ryegrass (*Lolium perenne* L.) and red clover (*Trifolium pratense* L.) can be used for this purpose as their root exudate can stimulate germination of the resting spores (Friberg et al. 2005). However, planting of bait crops may not be economical in a canola production system (Ahmed et al. 2011).

1.3.2 Chemical control

Fungicides: Based on the mode of action, fungicides can be classified as systemic and contact types (Dixon 2001; for review, see Tilman et al. 2002). Contact fungicides exert an effect on multiple metabolic processes of the target pathogen; however, they may also be toxic to different

microflora. Most of the newer systemic fungicides are less toxic and their target in the biochemical processes is highly specific. The majority of the fungicides tested were found to be uneffective against *P. brassicae*. A drench application of quintozene (Terraclor) was found to be efficient against this pathogen in Cole crops; however, drench application is not feasible in canola. Two other relatively new fungicides, fluazinam and cyazofamid, were also found useful in reducing disease severity in cruciferous crops with an efficiency of more than 50% (Gossen et al. 2012; Liao et al. 2021). However, the level of control achieved with these fungicides cannot be justified for use in a canola production system (Hwang et al. 2012, 2014).

Soil fumigation: Soil fumigants are toxic chemicals that form a gas following application, and thus help to control the pests, including *P. brassicae*. Several soil fumigants such as dazomet (tetrahydro-3, 5-dimethyl 2H-1, 3, 5- thiodiazine-2-thione) and Vapam (N- methyldithiocarbamic acid) were found to reduce clubroot disease severity in canola (for review, see Hwang et al. 2014; Hwang et al. 2017a); however, their application in farmers canola fields may not be economically viable.

1.3.3 Biological control

The use of micro-organisms has been considered as one of the most environment-friendly approaches for the control of clubroot disease (Peng et al. 2011a). However, availability and the consistency of their effect as biocontrol agents is limited. These types of agents can produce anti-microbial metabolites or colonize in the cortical tissue of the roots and, thus, provide resistance against clubroot (Hashiba et al. 2003). Three endophytic fungal isolates and four rhizobacterial isolates are found to reduce the severity of this disease by more than 75%; for example, *Heteroconium chaetospira* was found to colonize the cortical tissues of the root and reduce disease severity by > 80% (Lahlali et al. 2014). An isolate of *Phoma glomerate*, which generates the

mycotoxin epoxydon, appeared to suppress the development of *P. brassicae* by inducing anti-auxin activity (Arie et al. 1998, 1999). The endophytic fungus *H. chaetospira* also found to reduce clubroot in Chinese cabbage when inoculum concentration is low (Narisawa et al. 2005). In Canada, several registered commercial biocontrol products like *Bacillus subtilis* (Serenade), *Gliocladium catenulatum* (Prestop), *Streptomyces griseoviridis* (Mycostop), *S. lydicus* (Actinovate), *Trichoderma harzianum* (Root Shield), *Bacillus velezensis* have been evaluated for their efficacy against clubroot on crucifer vegetable crops (Peng et al. 2011a; Zhou et al. 2014; Zhu et al. 2020). However, these did not show significant effect in field trials, therefore, are not used in Canada as a tool for the management of clubroot disease in canola.

1.4 Clubroot Resistance and Breeding

Brassica species

The foundation of interspecific hybridization for the improvement of Brassica crop species is largely constituted from knowledge of the "Triangle of U" which illustrates the evolutionary and genetic relationships between the six important Brassica species (U 1935). Most of the globally important Brassica crops belong to *B. rapa*, *B. napus, B. juncea* and *B. oleracea*. *B. rapa* includes turnips (*B. rapa* var. *rapifera*), Chinese cabbage (*B. rapa* var. *pekinensis*), bok choy (*B. rapa* var. *chinensis*) and oilseed types. *B. oleracea* includes cabbage (*B. oleracea* var. *capitata*), kale (*B. oleracea* var. *acephala*), broccoli (*B. oleracea* var. *italica*) and cauliflower (*B. oleracea* var. *botrytis*). *Brassica napus* was derived from hybridisation of *B. rapa* and *B. oleracea* and includes rapeseed or canola (*B. napus* var. *napus*) and rutabaga (*B. napus* var. *napobrassica*).

Clubroot resistance

Breeding for clubroot resistance has been an important objective to the breeders for more than 50 years. Several germplasms of the genus *Brassica* have been investigated to identify resistance to

this disease. The close evolutionary relationship of *Brassica* species, as depicted by the triangle of U (U 1935) and, more recently, through analysis of the evolutionary pathway of the three *Brassica* genomes suggests that it is possible to transfer clubroot resistance from one species to the other or from one genome to the other through interspecific hybridization. *B. rapa* especially the European turnips and *B. napus* var. *napobrassica* often carry dominant resistance genes which have been used in *B. napus* canola breeding (for review, see Rahman et al. 2014b). Most of these resistances are pathotype-specific and controlled by Mendelian genes; however, pathotype non-specific quantitative resistance have been found in *B. oleracea* (Diederichsen et al. 2009; Farid et al. 2020).

In Canada, 45H29 was the first clubroot resistant canola cultivar released in 2009 by Pioneer Hi-Bred, Caledon, ON. Since then, several clubroot resistant cultivars have been released and as of November 2022, there are eight companies offering 63 clubroot resistant canola cultivars (Canola Council of Canada 2021); however, many of these cultivars are not completely resistant to the newly evolved pathotypes (Strelkov et al. 2016b, 2018). Furthermore, changes in the pathogen population structure and increasing inoculum concentrations pose a risk to the effectiveness and durability of resistance based on a single major gene in most of the *Brassica* crops (Hirai 2004; Strelkov et al. 2016b; Peng et al. 2018b); therefore, identification of new resistance genes is needed. Pyramiding or use of different resistance genes has been reported to provide resistance to multiple pathotypes (Matsumoto et al. 2012; Wang et al. 2022), and thus can increase durability of the resistance.

1.4.1 Clubroot resistance in *B. rapa*

Among the A and C genome *Brassica* species, *B. rapa* exhibit stronger resistance to a specific pathotype as compared to *B. oleracea*. In case of the B genome, *B. nigra*, often carry resistance, however, the amphidiploid species *B. juncea* and *B. carinata,* which also carry the B-genome, are

susceptible to this disease (Hasan et al. 2012). In *B. rapa*, resistance is mostly found in European fodder turnip (*B. rapa* spp. *rapifera*) cultivars like Gelria R, Siloga, Debra and Milan White as well as in the European Clubroot Differential host ECD 04. Major dominant genes of these sources have been used extensively in the breeding of the clubroot resistant cultivars. However, a rapid loss of resistance of the resistant-categorized cultivars has been observed due to the ever-evolving multiple pathotypes of *P. brassicae* (Hatakeyama et al. 2004; Strelkov et al. 2018, 2020). Therefore, pyramiding of multiple resistance genes with specificity to different pathotypes might be a strategy to develop a cultivar for durable resistance to this disease.

The European fodder turnip cultivars, such as, Gelria R, Siloga, Debra, Milan White and ECD 04 were used for introgression of clubroot resistance (CR) genes into susceptible vegetable *B. rapa* crops. Resistance in different *B. rapa* accession is controlled by different Mendelian genes. In the past two decades, genetic analysis and quantitative trait loci (QTL) mapping has identified several CR loci on A01, A02, A03, A05, A06, A07 and A08 chromosomes of *B. rapa*, and this includes *Crr2*, *CR6a* and *PbBa1.1* on A01 (Lee et al. 2002; Suwabe et al. 2003, 2006); *CRc* and *Rcr8* on A02 (Sakamoto et al. 2008; Yu et al. 2017); *Bcr1*, *BraA.CR.a*, *CRa/CRb*, *CRd*, *CRk*, *CRq*, *CR6b*, *PbBa3.1*, *PbBa3.2*, *PbBa3.3*, *Rcr1*, *Rcr2*, *Rcr4*, *Rcr5* and $Rcr10^{ECD01}$ on A03 (Lee et al. 2002; Hirai et al. 2004; Piao et al. 2004; Saito et al. 2006; Cho et al. 2008; Sakamoto et al. 2008; Ueno et al. 2012; Chen et al. 2013; Kato et al. 2013; Chu et al. 2014, Yuan et al. 2015; Yu et al. 2016, 2017, 2021; Hatakeyama et al. 2017; Huang et al. 2017, 2019; Hirani et al. 2018, Pang et al. 2018, Zhang et al. 2022); *CrrA5* on A05 (Nguyen et al. 2018); *Crr4* on A06 (Suwabe et al. 2006); *qBrCR38-1* on A07 (Zhu et al. 2019); and *Bcr2*, *BraA.CR.b*, *Crr1a*, *Crr1b*, *CRs*, *PbBa8.1*, $PbBrA08^{Banglim}$, *qBrCR38-2*, *Rcr3*, *Rcr9*, $Rcr9^{ECD01}$ on A08 (Suwabe et al. 2003, 2006; Chen et al. 2013; Hatakeyama et al. 2013; Yu et al. 2017, 2022; Hirani et al. 2018; Laila et al. 2019; Zhu et

al. 2019; Choi et al. 2020; Karim et al. 2020; Zhang et al. 2022). A comprehensive review on these CR loci has been published by Hasan et al. (2021a). Most of these loci are detected by using different genetic materials and different *P. brassicae* pathotypes; therefore, the use of different names for the same locus cannot be ruled out. For example, the DNA sequence of the functional ORF of *CRb* was found to be identical to *CRa*, demonstrating that these two loci are the same (Hatakeyama et al. 2017). On the other hand, *Crr1* was originally identified as a single locus, however, fine-mapping revealed that this locus comprises two loci — *Crr1a* with a major effect and *Crr1b* with a minor effect (Hatakeyama et al. 2013).

At least five classes of resistance (R) genes have been reported to date, viz. TIR-NBS-LRR (Toll/interleukin receptor- nucleotide binding site–leucine rich repeat), CC-NBS-LRR (Coiled coil- nucleotide binding site–leucine rich repeat), RLKs (Receptor like kinases), RLPs (Receptor like proteins) and *RPW8* (powdery mildew resistance gene in *Arabidopsis*) (for review, see Meyers et al. 2005). Among these, the Toll/interleukin receptor (TIR)-nucleotide binding site (NBS)–leucine rich repeat (LRR) class is known to be the largest one. Ueno et al. (2012) and Hatakeyama et al. (2013) identified TIR-NBS-LRR genes in the genomic regions of *CRa* and *Crr1a* through molecular analysis of these two genomic regions. It was generally thought that a single CR gene is resided in a locus; however, fine mapping of the genomic region of *Crr1* identified one major (*Crr1a*) and a minor (*Crr1b*) gene in this region; among these, the *Crr1a* encodes a TIR-NBS-LRR protein (1225 amino acid residues) which confer resistance to clubroot disease.

1.4.2 Clubroot resistance in *B. oleracea*

Genetic analysis of clubroot resistance revealed that this trait in *B. oleracea* is under polygenic control (Yoshikawa 1993); this includes the involvement of recessive genes to dominant genes with large to small effect on the trait. Laurens and Thomas (1993) reported loci exhibiting

dominance effect, whereas other researchers such as Chiang and Crete (1970) and Voorrips and Visser (1993) reported multiple recessive loci in the control of clubroot resistance in cabbage. Crisp et al. (1989) evaluated about 1,047 *B. oleracea* accessions for resistance to two *P. brassicae* isolates from the UK; and among these, kale (*B. oleracea* var. *viridis*), Brussels sprouts (*B. oleracea* var. *gemmifera*) and cabbages (*B. oleracea* var. *capitata*) were found to be the most resistant. Clubroot resistance in cabbage (Hasan et al. 2012), kale and two Brussels sprout (Fredua-Agyeman et al. 2019) have been reported by other researchers as well, and the occurrence of strong resistance in var. *acephala* (kale) has also been reported by Farid et al. (2020).

To date, QTL from almost all chromosomes of *B. oleracea* were detected by different researchers (for review, see Hasan et al. 2021a). For example, Landry et al. (1992) reported two QTL *CR2a* and *CR2b* contributing, respectively, 58% and 15% of the total phenotypic variance for resistance to *P. brassicae* race 2. Similarly, three QTL showing resistance to race 7 were identified in broccoli (Figdore et al. 1993). Moriguchi et al. (1999) identified three QTL, viz. *QTL1*, *QTL3* and *QTL9*, in the kale accession K269. By using genotypic data from 60K Brassica SNP array, Peng et al. (2018b) reported a total of 23 QTL, including *DIC.I-1* and *DIC.II-1* associated with CR in *B. oleracea*, where individual QTL explained 6.91- 17.7% of the phenotypic variance. Dakouri et al. (2018) reported a CR locus '*Rcr7*' exhibiting resistance to pathotype 3 by using a segregating population derived from a resistant cabbage cultivar Tekila crossed with susceptible line T010000DH3. This locus has been confirmed by two SNPs and using Kompetitive Allele Specific PCR markers. Nagaoka et al. (2010) identified one major QTL *Pb-Bo* (*Anju*) *1* located on linkage group 2 of *B. oleracea* explaining about 47% of the total phenotypic variance for resistance to race 4 of *P. brassicae*. The two QTL, *CRQTL-GN_1* and *CRQTL-GN_2*, for resistance to pathotype 9 and one QTL, *CRQTL_YC*, for resistance to pathotype 2 have been

identified by Lee et al. (2016). Among these, the *CRQTL_YC* was found to overlap *CRQTL-GN_2* implying that a single genomic region may confer resistance to more than one pathotype. These two QTL were found to be located at the same position where QTL for resistance to pathotype 4 has been reported by Nagaoka et al. (2010). Farid et al. (2020) identified 10 QTL (*PbC3*, *PbC4.1*, *PbC4.2*, *PbC6*, *PbC7.1*, *PbC7.2*, *PbC7.3*, *PbC8*, *PbC9.1*, *PbC9.2*) from six C-genome chromosomes associated with resistance to pathotypes 3A and 5X, where individual QTL explained about 8-10% of the total phenotypic variance.

1.4.3 Clubroot resistance in *B. napus*

Studies on clubroot resistance in *B. napus*, depending on the materials being investigated, showed simple Mendelian or oligogenic control of this trait. The resynthesized *B. napus* (resistance combined from *B. rapa* and *B. oleracea*) exhibits a complex inheritance (Diederichsen et al. 1996). Clubroot resistance in many commercial canola cultivars has been introduced via intra- or inter-specific crosses involving *B. rapa*, *B. oleracea* and rutabaga (*B. napus* spp. *napobrassica*) (for review, see Rahman et al. 2014b; Fredua-Agyeman and Rahman 2016; Fredua-Agyeman et al. 2018). Rutabaga and canola share the same genome ($2n = 38$, AACC); therefore, introgression of clubroot resistance from rutabaga to canola does not face the pre- and post-fertilization crossing barrier challenges. It has also been reported that rutabagas possess resistance against the old and new pathotypes of *P. brassicae* (Hasan et al. 2012; Fredua-Agyeman et al. 2019). Ayers and Lelachur (1972) found that the rutabaga cultivars York and Wilhemsburger are resistant to pathotypes 2 and 3 (as defined on the differentials of Williams 1966) and the cultivar Ditmars S2 is resistant to pathotype 3. Since then, rutabagas are being used in the breeding of clubroot resistant canola cultivars (for review, see Rahman et al. 2014b; Hasan and Rahman 2016; Hasan et al. 2021b). In a recent study, nine *B. napus* cultivars from different geographic origins (six European

winter oilseed rape cultivars Aviso, Bristol, Darmor-bzh, Express, Montego and Samourai; three spring oilseed rape cultivars Cresor from Europe, Yudal from Korea and Grouse from Australia) were tested for resistance to Canadian field isolates of *P. brassicae* (Aigu et al. 2020). It was found that under full nitrogen fertilization, these cultivars exhibit isolate-specific partial resistance to the pathogens.

Diederichsen and Sacristán (1996) reported that resistance in resynthesized *B. napus* (developed from *B. rapa* ssp. *rapifera* cv. ECD 04 and *B. oleracea* ssp. *capitata* cv. ECD 15) is controlled by at least two dominant unlinked genes. Werner et al. (2008) developed a mapping population by using the above-mentioned resynthesized *B. napus* and reported 19 QTL from eight chromosomes conferring resistance to seven field isolates of *P. brassicae*. One major locus of linkage group DY4 (*Pb-Bn1*), and two small effect QTL of linkage groups DY4 and DY15 were found to be contributing to the resistance to the isolates PB137-522 and K92-16, respectively (Manzanares-Dauleux et al. 2000). Zhang et al. (2016a) identified the locus *CRa* on A03 by using an oilseed *B. napus* DH population. One major and a minor locus, respectively, on chromosome A08 and A03 were identified by Hasan and Rahman 2016 and Hasan et al. (2021b). Hejna et al. (2019) found that two major QTL located on A02 and A03 provide resistance to pathotype ECD 17/31/31.

Resistance in majority of the canola cultivars initially developed in Canada was derived from the winter canola cv. Mendel (Rahman et al. 2011); resistance in this cultivar is controlled by a major locus located on chromosome A03 (Fredua-Agyeman and Rahman 2016). The pedigree of the cv. Mendel carries the resynthesized *B. napus* line produced by crossing *B. rapa* var. *rapifera* ECD 04 and *B. oleracea* cv. Bohmerwaldkohl (Diederichsen et al. 2006). This ‘single gene controlled’ resistance in canola became ineffective in four years with the evolution of new virulent

P. brassicae pathotypes (Strelkov et al. 2016b). Therefore, it is necessary to have multiple resistance genes in a canola cultivar for durable resistance to this disease (for review, see Hasan et al. 2021a). Clubroot resistant *B. rapa* lines pyramided with three major genes, *Cra*, *CRk* and *CRc*, were developed by Matsumoto et al. (2012) through marker-assisted selection (MAS), and the pyramided genes provided resistance to six *P. brassicae* isolates. Similar approach was followed by Tomita et al. (2013) to pyramid five QTL in *B. oleracea*. Shah et al. (2019) and Wang et al. (2022) also found that two CR loci together can provide resistance to multiple pathotypes in *B. napus*.

1.5 Whole-genome resequencing (WGRS) for identification of clubroot resistance genes/loci

The dramatic advances in the next-generation sequencing (NGS) technologies have resulted in the genome sequencing of five *Brassica* species *viz.*, *B. rapa* (Wang et al. 2011; Cai et al. 2017), *B. oleracea* (Parkin et al. 2014), *B. nigra* (Yang et al. 2016), *B. napus* (Chalhoub et al. 2014; Bayer et al. 2017; Rousseau-Gueutin et al. 2020; Song et al. 2020) and *B. juncea* (Yang et al. 2016a) as well as the model plant *Arabidopsis thaliana* (Schranz and Mitchell-Olds 2006). Availability of the genome sequences have enabled the identification of QTL and candidate genes controlling different traits such as the yield, quality, and disease/pest resistance (Delourme et al. 2018). Thus, the use of genome sequences is expected to accelerate the breeding of *Brassica* crop species for the target traits as well as will help to understand the molecular aspects of plant-pathogen interactions in host resistance.

Whole-genome resequencing (WGRS) is an omics technique which involves the alignment and comparison of the sequence reads of the materials of interest to the reference genome for the identification of SNPs and insertion/deletions (Indels). In molecular plant breeding, bulked segregant analysis (BSA) has been widely used for identification of molecular markers linked to a

gene of interest. In this method, two bulks of DNA (extremely opposite phenotypes from a mapping population) are genotyped to identify polymorphic markers, and the polymorphic markers are tested on the whole mapping population for linkage association of the markers with the trait (Michelmore et al. 1991). Identification of a QTL through WGRS also employ a similar principle. In this case, SNP indices of the two bulks for each SNP position is calculated by sequencing the bulks and mapping the sequence reads of each of the bulks individually to a reference genome. SNP index can be defined as the ratio between the numbers of alternative reads (SNPs) to the total number of reads corresponding to that SNP position. The knowledge of the SNP indices of the two bulks can be used to locate the genomic region responsible for the trait of interest. For this, ΔSNP index is calculated by subtracting SNP index of the two bulks (e.g., resistant and susceptible), where ΔSNP index of 1.0 or -1.0 for a SNP would indicate the SNP is associated with the trait of interest (Pandey et al. 2017).

SNP index analysis of an F_2 population derived from crossing a clubroot resistant (85-74) and susceptible (BJN3-1) Chinese cabbage lines was used to locate the dominant locus '*CRd*' conferring resistance to race 4. This locus was mapped between the markers *yau389* and *yau376* of A03 located upstream of the locus *Crr3* (Pang et al. 2018). Two novel QTL, '*qBrCR38-1*' and '*qBrCR38-2*' and seven candidate genes to be involved in resistance to race 7, have also been identified in Pakchoi (*B. campestris* ssp. *chinensis*) by using the NGS-BSA technique (Zhu et al. 2017). This technique has been used by Hasan et al. (2021a, 2021b) and Wang et al. (2022) for QTL mapping of resistance located on A03 and A08 of *B. napus* var. *napobrassica* and to identify the putative candidate genes involved in CR. Wang et al. (2022) also demonstrated that interaction of the A03 and A08 loci contribute to the resistance to multiple pathotypes including 3H, 3A and 3D. Hasan et al. (2021b) identified a single major locus on chromosome A03 of *B. rapa* var.

pekinensis cv. Bilko to be associated with CR in *B. napus*. One of the major challenges with the above-mentioned approaches is the difficulty of identifying the candidate genes to develop gene-based SNP markers for use in breeding. To counter this difficulty, transcriptomics has emerged as a great potential for increasing the power of detection of genomic loci and genes controlling the traits of interest.

1.6 Transcriptomics and its application in *Brassica*-clubroot pathosystem

High-throughput RNA-sequencing is a powerful tool for transcriptome profiling which can detect the genes controlling a trait, including quantification of their expression (for review, see Wang et al. 2009). It also permits the discovery of novel genes or transcripts and improves the gene annotations by extending transcriptional boundaries of the genes (for review, see Ozsolak and Milos 2011). RNA-seq has also been used to identify the alternative splicing events and 3' untranslated regions (3' UTRs) (for review, see Cloonan and Grimmond 2008; Sultan et al. 2008; Trapnell et al. 2012). Most RNA-Seq experiments involve shearing a sample of purified RNA, converting these into cDNA prior to sequencing on a high-throughput sequencing platform such as Illumina, SOLiD or Roche 454. This generates millions of short reads of 25-300 bp from one end of the cDNA fragments or paired end reads. These short reads are then mapped on the genome, and assembled into a gene-level, exon-level, or transcriptome-level expression summaries. The summarized data is normalized in accordance with statistical testing for differential expression leading to the list of candidate genes with associated *p*-values and fold changes. The biological insights of these transcripts can finally be gained by following systems biology approaches (Oshlack et al. 2010). Data generated from RNA-seq is highly reproducible (Marioni et al. 2008), and this technique has been used for identification of SNPs that can be used as molecular markers as well as for the analysis of host-pathogen interactions. Another technical innovation called 'dual-

RNA seq' enables the simultaneous analysis of gene expression in the pathogen and its host (for review, see Westermann et al. 2012).

Recently, a combination of bulked segregant analysis and RNA-seq, named 'bulked segregant RNA-seq (BSR-Seq)', has been used to map the QTL and genes of specific trait of interest like disease resistance (Du et al. 2017; Huang et al. 2017; Fu et al. 2019b). BSR-Seq, which uses the NGS technique, has been employed to locate and fine map the clubroot resistance loci in *Brassica napus* (for review, see Zhou et al. 2021). This method, through identification of the polymorphic variants, has been used to map the clubroot resistance gene *Rcr7* located within a physical interval of 41- 44 Mb of C07 in *B. oleracea* genome (Dakouri et al. 2018). Chang et al. (2019) mapped the locus *Rcr6* on chromosome A08 of *B. rapa* and by using the BSR-Seq approach, they identified its homoeologous copy on the chromosome B07 of *B. nigra*. Of the 190 genes annotated, five genes (*Bra010552*, *Bra010588*, *Bra010589*, *Bra010590* and *Bra010663*) were found to encode for TIR-NBS-LRR class of proteins. By using SNP markers identified through BSR-Seq approach, Huang et al. (2017, 2019) fine mapped the A03 clubroot resistance loci *Rcr2* and *Rcr5* of the Chinese cabbage cvs. Jazz and Purple Top White Globe.

Transcriptome sequencing together with digital gene expression (DGE) analysis has been used to elucidate the host-pathogen interactions in Brassicas such as *B. napus*, *B. rapa* and *B. oleracea* (Chu et al. 2014; Jia et al. 2017; Dakouri et al. 2018; Fu et al. 2019; Qasim et al. 2020; Adhikary et al. 2022b). Jia et al. (2017) conducted RNA-Seq analysis on resistant (R635-10) and susceptible (S177-47) *B. rapa* lines before and after infection by *P. brassicae* to identify the differentially expressed genes (DEGs) to be associated with resistance. They identified 2,089 DEGs upregulated in the resistant line; these genes were predicted to be involved in various biological pathways, such as disease resistance, calcium ion influx and cell wall thickening. Mei

et al. (2019) conducted RNA-seq analysis in a clubroot resistant line ZHE-226 and susceptible cv. Zhongshuang 11 of *B. napus* at 0, 3, 6, 9 and 12 days after inoculation with *P. brassicae*. A significant promotion of DNA repair, protein synthesis, signalling and activation of host intercellular G (Guanine nucleotide-binding) proteins, production of reactive oxygen species etc. was witnessed as soon as the root hair infection occurred in ZHE-226. Similar comparative transcriptomics studies at early- and late-infection stages of *P. brassicae* have been conducted by different researchers to identify the DEGs and to understand the molecular mechanisms of resistance (Chen et al. 2016; Zhang et al. 2016; Summanwar et al. 2021; Adhikary et al. 2022b); however, they still lack their functional characterization.

1.7 Proteomics and its application in *Brassica*-clubroot pathosystem

Most of the currently available knowledge related to the plant responses to pathogen infection has been obtained by using genomic and transcriptomic approaches. These approaches have contributed greatly to our understanding of the host-pathogen interactions; however, correlation between mRNA expression and protein abundance levels is often poor as the protein levels are also dependent on the host's fundamental biological factors in the transcription and translation processes (for review, see Feussner and Polle 2015). Therefore, proteomics would be one of the most significant approach to characterize a biological system. The term 'proteome' refers to the entire protein content of a cell that can be characterized in its location, their interactions, and post-translational modifications at a particular time in a cell. Wilkins et al. (1996) coined the term 'proteomics' in 1996 to refer to the "PROTein complement of a genOME". The field of proteomics started making its impact after the whole genome sequences of many organisms became available, extending our knowledge from the gene to the phenotype. There have been tremendous technological advancements in this field owing to the innovations in protein extraction,

purification and separation, mass spectrometry analysis, protein identification and quantification over the past 20 years, thus, making it easier to quantify the protein accumulation changes between conditions and genotypes (for review, see Aslam et al. 2017). Based on platforms and equipment being used, proteomics analysis has moved from first to fourth generation, originally based on 2-DE protein separation along with MS analysis of spots, moving to liquid chromatography-based shotgun methods, labelled and label-free quantitation approaches, and finally to selected or multiple reaction monitoring strategies (for review, see Jorrin-Novo 2014).

High-throughput proteomics analysis has been conducted in Brassica-clubroot pathosystem to investigate the post-transcriptional changes in the host proteins in response to infection (Cao et al. 2007; Song et al. 2016; Ji et al. 2018, 2021; Su et al. 2018; Lan et al. 2019; Moon et al. 2020; Adhikary et al. 2022a). One of the earliest studies on host-*P. brassicae* interaction was conducted by Devos et al. (2006) using the model plant *A. thaliana* inoculated with clubroot pathogen and identified a number of proteins (35 increased and 11 decreased in abundance) involved in host cell wall modifications in this interaction. Cao et al. (2007) found differential abundance of 20 proteins related to lignin biosynthesis, plant hormone metabolism, calcium signaling and reactive oxygen species (ROS) detoxification in the inoculated susceptible *B. napus* at 12, 24, 48 and 72 hours after inoculation (hai). Proteome analysis at early infection stage (3-5 dai) of *B. rapa* ssp. *pekinensis* by *P. brassicae* confirmed the role of salicyclic acid (SA) and jasmonic acid (JA) pathways in defense response (Ji et al. 2018, 2021). iTRAQ-based quantitative proteomics technique was used by Su et al. (2018) to investigate the role of BR pathway in defense response in *B. rapa*. Lan et al. (2019) used the same technique at secondary stage of infection (14, 21, 28, 35 and 42 dai) and reported 5,003 proteins in resistant and susceptible *B. rapa*; among these, accumulation of 487 proteins changed over the course of infection indicating

the occurrence of a change in physiology of the plant due to *P. brassicae* infection. They also found an increased abundance of early nodulin protein in the susceptible *B. rapa* plants which resulted in hormonal imbalance causing uncontrolled cell division in the infected roots. The involvement of abscisic acid (ABA) in the resistance mechanism against this disease has been reported in *B. oleracea* (Moon et al. 2020). Adhikary et al. (2022a) identified 73 putative proteins, which are orthologous to clubroot resistance proteins, from the A03 and A08 CR QTL region. In addition, they also observed putative proteins related to calcium signalling, reactive oxygen species, dehydrins, lignin, thaumatin and phytohormones.

1.8 C2H2-Zinc finger proteins

Zinc-finger proteins (ZFPs) are one of the largest families of transcription factors (TFs) in plants having multiple roles in growth and development, metabolic processes, hormone signal transduction and resistance mechanism against several biotic and abiotic stresses (for review, see Takatsuji et al. 1999; Tian et al. 2010; Feurtado et al. 2011; Giri et al. 2011). Their role in the regulation of stress responses can be postulated from the presence of zinc-finger DNA binding domain in the nucleotide binding sequences-leucine rich repeats (NBS-LRR) class of resistance (R) proteins. A 'zinc-finger' can be defined as a small compact freely folded domain having cysteine (Cys) and histidine (His) residues for coordination of divalent zinc ions to stabilize its three-dimensional structure (for review, see Laity et al. 2001; for review, see Kielbowicz-Matuk 2012). They can be categorised into nine classes based on their structural variations (number and order of Cys and His) and this includes C2H2 (TFIIIA, Kruppel-like zinc-finger proteins), C2HC (Retroviral nucleocapsid), C2HC5 (LIM domain), C3HC4 (RING finger), C4HC3 (Requium), C2C2, C4 (GATA-1), C6 (GAL4) and C8, where the C and H represents the cysteine and histidine residues (for review, see Laity et al. 2001; for review, see Liu et al. 2022). Among these, the C2H2-

type ZFPs are the most studied DNA-binding motifs, accounting for approximately 0.7% of *A. thaliana* genes, 0.8% of yeast genes and 3% of dipteran and mammalian genes (Böhm et al. 1997; Englbrecht et al. 2004; for review, see Ciftci-Yilmaz and Mittler 2008). The C2H2-ZFPs are composed of about 30 amino acids having two Cys and two His residues coordinated together with one zinc ion in a tetrahedral manner (for review, see Laity et al. 2001). Their sequence can be represented as $CX_{2-4}CX_3FX_5$-$LX_2HX_{3-5}H$, where the X denotes any amino acid, and the subscript denotes the number of amino acids (Kubo et al. 1998). This family of ZFPs have two classes, C-type and Q-type. The Q-type C2H2-ZFPs are plant specific having a conserved motif 'QALGGH' while the C-type is present in other organisms as well (Agarwal et al. 2007). The identification of DNA-binding protein *ZPT2-1* (previously called *EPF1*) in Petunia was the pioneering study in discovering the C2H2-type ZFPs in plants (Takatsuji et al. 1992). Since then, many C2H2-type ZFPs have been reported in many plant species including 176 in *A. thaliana* (Englbrecht et al. 2004), 189 in *Oryza sativa* (Agarwal et al. 2007), 321 in *Glycine max* (Yuan et al. 2018), 301 in *B. rapa* (Alam et al. 2019) and 122 in *Triticum aestivum* (Faraji et al. 2018).

1.8.1 The role of C2H2-ZFPs in abiotic stress tolerance

Functional analyses of C2H2-ZFP domains revealed their importance in regulation of transcription in biological processes, DNA-RNA binding, growth and development, root hair development, and biotic and abiotic stress responses (Kim et al. 2004; Luo et al. 2012; Sharma et al. 2021; Yu et al. 2021; Zheng et al. 2021). Several C2H2-ZF proteins reported to increase salt tolerance in plants, and this includes *ZFP252* and *ZFP179* of rice (Xu et al. 2008; Sun et al. 2010), *GmWRKY49* of soybean (Xu et al. 2018), *Bra009464* orthologous to *AtZAT6* of *B. rapa* (Alam et al. 2019), and *ZFP3* and *AtSIZ1* of *A. thaliana* (Zhang et al. 2016; Han et al. 2019b). C2H2-ZFPs have also been reported to play a role in enhancing cold and drought tolerance in different crops such as *ZFP245*

of rice (Huang et al. 2009), *SCOF-1* of soybean (Kim et al. 2001), *BcZAT12* of tomato (Rai et al. 2013) and *ZFP1* of *A. thaliana* (Luo et al. 2012a). *GhSTOP1* has been found to increase root growth and counteract the aluminium and proton stress in *A. thaliana* (Kundu et al. 2019) whereas *AZF2*, an orthologous gene of *Bra022436*, functions as a transcription repressor and inhibits plant growth under stress conditions (Alam et al. 2019). Another C2H2-ZFP named *SIZF2* alters the flower and leaf shape in *Arabidopsis* and increases salt tolerance in tomato (Hichri et al. 2014). Tao et al. (2017) reported the involvement of *BnLATE* zinc-finger protein in silique shattering in *B. napus*.

1.8.2 The role of C2H2-ZFPs in biotic stress tolerance

Only a few studies have been carried out to depict the role of C2H2-ZFPs in response to pathogen infection in plants (Kim et al. 2004; Tian et al. 2010; Yin et al. 2020; Sharma et al. 2021). An over-expression of the grapevine C2H2-ZFP, *VvZFP11*, enhanced the resistance to *Golovinomyces cichoracearum* in *A. thaliana* (Yu et al. 2016). Yin et al. (2020) identified 129 C2H2-ZFPs in *Cucumis sativus*, and transient over-expression of *Csa1G085390* and *Csa7G071440* in *Nicotiana benthamiana* increased or decreased leaf necrosis in response to pathogen infection. *ZFT1* and *StZFP1* enhanced tolerance to Tobacco mosaic virus and *Phytophthora infestans* in tobacco and potato, respectively (Uehara et al. 2005; Tian et al. 2010). Zhao et al. (2020) found a change in expression of eight out of 32 C2H2-ZF genes under biotic stresses in tomato. Recently, 18 C2H2-ZF genes were found to exhibit differential expression in *Capsicum annuum* after infection with *Colletotrichum truncatum* causing anthracnose disease (Sharma et al. 2021).

1.9 Research objectives

The ultimate goal of this PhD thesis research was to develop canola-quality clubroot resistant *B. napus* lines, and to get a deeper insight into the *B. napus*-*P. brassicae* pathosystem for use of the

knowledge in molecular breeding. For this, the following studies were undertaken:

1) Investigate the feasibility of introgression of resistance to multiple *P. brassicae* pathotypes from turnip into spring canola and the development of canola-quality *B. napus* lines (Chapter 2).
2) Map the clubroot resistance genes introgressed into canola for the development of molecular markers for resistance (Chapter 3).
3) Decipher the molecular cues of host-pathogen interactions through transcriptomics and proteomics studies and investigate the prospects of using this knowledge in molecular breeding (Chapter 4 and 5).
4) Identify the C2H2-zinc finger protein genes in *B. napus* and investigate their roles in biotic and abiotic stress tolerance including resistance to clubroot disease (Chapter 6).

Chapter 2. Introgression of resistance to multiple pathotypes of *Plasmodiophora brassicae* from turnip (*Brassica rapa* ssp. *rapifera*) into spring *B. napus* canola

2.1 Introduction

Clubroot, a soil-borne disease, caused by *Plasmodiophora brassicae* Woronin, has been known since the 13th century in cruciferous oilseed and vegetable crops such as *Brassica napus* (2*n* = 38, AACC), *Brassica rapa* (2*n* = 20, AA), and *Brassica oleracea* (2*n* = 18, CC) (for review, see Dixon 2009). *P. brassicae* is an obligate biotrophic protist which attacks root hair and results in abnormal cell division, and thus, forms club shaped roots or 'galls'. These clubbed roots inhibit the uptake of water and nutrients and result in stunted growth, wilting, chlorosis, and even the death of the plant (Kageyama and Asano 2009). This disease is known to be present in about 60 countries and causes 10–15% yield loss of *Brassica* crops on a global scale (for review, see Dixon 2009). In the case of canola, yield loss due to this disease has been reported in many countries including Canada (Tewari et al. 2005; for review, see Rahman et al. 2014), China (for review, see Chai et al. 2014), India (for review, see Bhattacharya et al. 2014), Europe (for review, see Diederichsen et al. 2014; for review, see Wallenhammar et al. 2014), and Australia (for review, see Donald and Porter 2014). This disease can result in about 30% yield loss in canola (Tewari et al. 2005); however, under extreme conditions, complete failure of the crop has been reported (for review, see Howard et al. 2010). Apart from yield loss, this disease can result in about a 6% decrease in seed oil (Pageau et al. 2006).

The resting spores of *P. brassicae* can survive in soil for more than 15 years (Wallenhammar 1996) and various pathotypes of this pathogen generally occur in soil (Manzanares-Dauleux et al. 2001; Strelkov et al. 2006, 2016b). The structure of the pathotype population of *P. brassicae* in soil can change in a short period of time (Strelkov et al. 2016b). For

example, in the initial stage of the occurrence of this disease in Canada, pathotype 3 or 3H is the most virulent and prevalent. However, new virulent pathotypes, such as 3A, 2B, 3D, and 5X evolved in canola fields after growing resistant cultivars for about 10 years rendering some of the resistances ineffective (Strelkov et al. 2016b, 2018). Strelkov et al. (2021) reported that 3H was the most common pathotype in the Peace Country of Canada in 2017, and this pathotype was almost absent in 2018; however, novel pathotypes such as 8C were detected in this region. Control of this disease by crop management practices such as cultural, chemical, and agro-technical treatments was not sustainable (for review, see Hasan et al. 2021a); therefore, the deployment of resistant cultivars has been considered the most effective, economic, and environment-friendly approach for the long-term management of this disease (for review, see Rahman et al. 2014).

To date, a total of 24 clubroot-resistant loci conferring resistance to different pathotypes have been mapped on the A01, A02, A03, A05, A06, A07, and A08 chromosomes of the A genome of *B. rapa* and *B. napus* (for review, see Hasan et al. 2021a). The majority of the loci including *CRa*/*CRb* (Matsumoto et al. 1998; Piao et al. 2004; Cho et al. 2008; Kato et al. 2012, 2013; Hatakeyama et al. 2017), *CRd* (Pang et al. 2018), *CRk* (Sakamoto et al. 2008), *Crr3* (Hirai et al. 2004; Saito et al. 2006), *CRq* (Yuan et al. 2015), *Rcr1* (Chu et al. 2014), *Rcr2* (Huang et al. 2017), *Rcr4* (Yu et al. 2017), and *Rcr5* (Huang et al. 2019) have been mapped at the 15.5–16.3 Mb and 23.8–26.8 Mb regions of A03 of the reference genome *B. rapa* Chiifu-401 (for review, see Hasan et al. 2021a). Among the other chromosomes, A08 carries five loci while the remaining chromosomes carry one or two loci. These loci often confer race-specific resistance to pathotypes, such as *Rcr1*, *Rcr2*, and *Rcr5* which confer resistance to pathotype 3, *Crr3* and *CRq* which confer resistance to pathotype 2, and *CRb* which confers resistance to pathotypes 2, 3, 4, and 8 (for review, see Hasan et al. 2021a). On the other hand, clubroot resistance in the C genome of *B. oleracea* is

often under polygenic control with the involvement of recessive to dominant gene loci with a small to large effect (Farid et al. 2020; for review, see Hasan et al. 2021a). Of the major clubroot-resistant loci of the A genome, only a few of them have been used to develop *B. napus* canola cultivars in Europe and Canada (for review, see Rahman et al. 2014; for review, see Hasan et al. 2021a). For example, in Canada, the major clubroot resistance genes were obtained from the European winter oilseed *B. napus* cv. Mendel (Rahman et al. 2011), and *B. rapa* subsp. *chinensis* cv. Flower Navina (Chu et al. 2013) and Chinese cabbage (*B. rapa* ssp. *pekinensis*) (Hirani et al. 2016; Hasan et al. 2021c); among these, resistance of the cv. Mendel has been used extensively in breeding. The cv. Mendel carries a race-specific resistance originating from resynthesized *B. napus* from a partially resistant cabbage (*B. oleracea* ssp. *capitata*) accession ECD 15 and a resistant turnip (*B. rapa* ssp. *rapifera*) accession ECD 04 (Diederichsen et al. 1996). However, the major resistance gene of this cultivar conferring resistance to pathotype 3H is located on A03 (Fredua-Agyeman and Rahman 2016); this gene is apparently derived from the turnip accession ECD 04. The European turnips were also used to develop race- or pathotype-specific clubroot-resistant rutabaga (*B. napus* var. *napobrassica*) cultivars (Ayers and Lelacheur 1972; Bradshaw et al. 1997). The clubroot resistance of rutabaga cv. Brookfield, located on A08, has been introgressed into Canadian canola (Hasan and Rahman 2016). A cultivar carrying a single major clubroot-resistant gene has been reported to become ineffective only after a few years of cultivation; this has been reported in several countries, including Canada (Strelkov et al. 2016b) and Japan (Hirai et al. 2004). This is primarily due to the change in the population structure of this pathogen in soil and the evolution of new virulent pathotypes (LeBoldus et al. 2012; Strelkov et al. 2016b, 2018; Sedaghatkish et al. 2019). As mentioned above, the A-genome of *B. rapa* carries several clubroot resistant loci (for review, see

Hasan et al. 2021a); in this regard, introgression of additional loci from this species into canola will broaden the genetic base of clubroot resistance in this crop.

As discussed above, a few clubroot-resistant loci have been introgressed into *B. napus* from *B. rapa* (Frauen et al. 1999; Chu et al. 2013; Hirani et al. 2016). However, detailed information of this, including the breeding behavior of the *B. napus* × *B. rapa* interspecific cross progeny and the prospect of the introgression of resistance to multiple pathotypes cannot be found in the literature. The turnip accession ECD 01 carry resistance to multiple pathotypes including pathotype 3H and 3A. The objectives of this study were to investigate the feasibility of the introgression of resistance to multiple pathotypes from a turnip into a spring *B. napus* canola, and the development of a canola-quality *B. napus* line carrying resistance to these pathotypes. The knowledge from this research could be used for efficient introgression of additional resistances from the A genome of *B. rapa* into *B. napus* canola.

2.2 Materials and Methods

2.2.1 Plant material and population development

The parents used in this study include a clubroot-susceptible spring *B. napus* (2*n* = 38, AACC) canola-quality (zero erucic acid in oil and low GSL in seed meal) line A04-73NA and a clubroot-resistant *B. rapa* var. *rapifera* turnip accession of the European Clubroot Differential (ECD) set (accession ECD 01 or cv. Debra). ECD 01 was completely resistant to multiple pathotypes, including pathotype 3H and 3A (disease score = 0, disease severity index or DSI = 0%). The *B. napus* × *B. rapa* interspecific cross was made by using A04-73NA as the female, and F_1 seeds (>2 seeds/pollination) of this cross could be obtained without the application of the cell and tissue culture technique (Attri and Rahman 2018). High meiotic anomalies were expected in the allotriploid (AAC) F_1 plants; therefore, the F_1 plants were backcrossed to A04-73NA to achieve

canola-quality clubroot-resistant euploid ($2n$ = 38) *B. napus* lines. The BC_1 plants were self-pollinated for seven generations with selection for spring growth habit, canola-quality traits, and resistance to the *P. brassicae* pathotype 3H in the early generations, and resistance to 3A and other pathotypes in later generations.

2.2.2 Greenhouse assay for clubroot resistance

Single spore isolates of *P. brassicae* pathotypes 3H, 3A, 2B, and 5X used in this study were obtained from Dr. Stephen Strelkov, University of Alberta. These isolates were preserved in galls of a susceptible *B. napus* cv. Hi-Q at −18 °C until use. The inoculum of the resting spores was prepared following a modified version of Williams (Williams 1966) technique as described in Hasan et al. (2012). In brief, approximately 37 g galls were soaked in 1000 mL distilled water for 2 h and homogenized in a blender (Ninja Professional Blender 1100 W) at a medium speed for 2 minutes. The homogenate was filtered through cheesecloth and the concentration of the spore suspension was adjusted to 1×10^7 spores/mL and stored at 4 °C overnight before inoculation.

The BC_1F_2 to BC_1F_7 generation plants of the (*B. napus* × *B. rapa*) × *B. napus* interspecific cross were grown together with the susceptible check *B. napus* cv. Hi-Q in a greenhouse of the University of Alberta in 32-cell (7 cm × 7 cm × 9 cm, L × W × D) trays filled with Sunshine Professional Growing Mix (Sunshine Horticulture, 15831 N.E., Bellevue, WA, USA); the greenhouse conditions were 20–22/15 °C day/night, 16 h photoperiod, and light intensity of 450 μE (mV) m^2 s^{-1}. The BC_1F_2 to BC_1F_6 generation populations were tested for resistance to pathotype 3H, while the BC_1F_5 to BC_1F_7 populations were tested for resistance to pathotype 3A. The BC_1F_8 generation population was also tested for resistance to pathotypes 3A, 2B, and 5X in 2 replications (8 plants per replication). Inoculation was carried out at 7–10 days after germination with 1 mL inoculum following the pipette method (Voorrips and Visser 1993). In this method, the

inoculum was pipetted at the base of the seedling, and the inoculation was repeated the following day to ensure successful infection. After inoculation, soil in the cells was kept well saturated for 14 days to ensure sufficient moisture for infection; afterwards, the inoculated seedlings were watered daily and fertilized once a week. During the advancement of generation from BC_1F_2 to BC_1F_7, five to ten buds of the main raceme of the plants were self-pollinated by bag isolation and the plants were scored for resistance at maturity; seeds were harvested for growing the next generation population. The BC_1F_8 generation seeds were produced by growing the BC_1F_7 plants in 5-inch size pots and through self-pollination of individual plants by bag isolation.

Scoring of the plants for clubroot disease severity was carried out using a 0–3 scale (for details, see Kuginuki et al. 1999, Strelkov et al. 2006), where 0 = no gall, 1 = small galls on less than 1/3rd of the roots, 2 = small to medium-sized galls on 1/3rd to 2/3rd of the roots, and 3 = severe galling with medium to large-sized galls on more than 2/3rd of the roots. The disease severity index (DSI) of different generation families was calculated using the following formula:

$$\text{DSI (\%)} = \frac{\sum(n \times 0 + n \times 1 + n \times 2 + n \times 3)}{N \times 3} \times 100$$

where, *n* is the number of plants in each disease severity class and 0, 1, 2, and 3 are the disease symptom severity classes, and *N* is the total number of plants (Hasan et al. 2012). A DSI of less than 20% was considered resistant (R) and more than 70% was considered susceptible (S) while DSIs of >20 to 40% and >40 to 70% were considered as moderately resistant (MR) and moderately susceptible (MS), respectively. While growing the BC_1F_2 to BC_1F_7 generation plants in the greenhouse, selection was performed for clubroot resistance and zero erucic acid in seed oil.

2.2.3 Field evaluation for clubroot resistance

A total of 341 BC_1F_8 lines exhibiting resistance to pathotype 3H and containing zero erucic acid in seed oil were tested in a *P. brassicae*-infested field in Spruce Grove, Alberta in 2020 for clubroot

resistance. This field carries multiple *P. brassicae* pathotypes, including 5X (Alberta Agriculture and Forestry, personnel communication). For this, a 0.5 g seed was seeded in a 3 m long single-row plot in one replication with 70 cm space between the rows; the susceptible check cv. Hi-Q was seeded at every 10th row as a check. Clubroot resistance was assessed at the end of flowering by uprooting 25 plants from each row and scoring the plants on a 0–3 scale, as described above.

2.2.4 Field trials for agronomic and seed quality traits

A total of 452 BC_1F_8 lines, including the 341 lines tested in a *P. brassicae*-infested field, were also evaluated in field nursery plots in 2020 at the South Campus Research Station of the University of Alberta for agronomic and seed quality traits. For this, seeding was performed in 3 m long single row plots in one replication, as described above. The canola line A03-73NA was seeded at every 20th row as a check. Open-pollinated seeds harvested by hand were used for chemical analysis.

Of the above-mentioned BC_1F_8 lines, 112 selected lines (BC_1F_9) were tested in replicated field trials in 2021 together with a recently released open-pollinated canola cultivar UA CountyGold for agronomic and seed quality traits. These trials were placed at three locations: the ERS-Michener, ERS-West 240, and St. Albert research farms of the University of Alberta within a 30 km radius of Edmonton, Alberta. Seeding was performed with a plot seeder in 7.5 × 1.8 m plots using 8 g seeds per plot, and the plots were trimmed to 5.0 × 1.8 m after emergence. All trials were laid in a modified randomized complete block design where each replication was divided into four blocks; the number of replications for each trial was two.

The following agronomic and seed quality traits were recorded: days to flowering, days to maturity, seed yield (kg/ha), seed oil (%), protein (%), and GSL (μmol/g seed) contents. Days to flowering data were recorded when about 50% of the plants in a plot had at least one open flower. Days to maturity data were collected when the silique color changed from green to light yellow or

brown and the seed color in the silique of the main branch changed to brown or black. All yield trial plots were harvested using a plot combine and plot yield data were converted to kilograms per hectare.

2.2.5 Chemical analysis

The fatty acid profiles of the seed oil of the BC_1F_5 and BC_1F_6 generation plants grown in greenhouse and of the seeds harvested from field trials were estimated by the gas chromatographic technique. Bulk seeds harvested from the field plots were analyzed for oil, protein, and GSL contents using the near-infrared spectroscopy (NIRS) technique (FOSS NIRSystems, Model 6500). About 0.5 g seeds of each sample was used for the fatty acid analysis while 2.5 g seeds were used for NIRS analysis. Erucic fatty acid and saturated fatty acids, which include lauric acid, myristic acid, palmitic acid, stearic acid, arachidic acid, behenic acid, and lignoceric acid, were reported as a percent of the total fatty acids; GSL content was reported as μmol g^{-1} seed at 8.5% moisture. All these analyses were carried out in the Analytical Laboratory of the Canola Program of the University of Alberta; this laboratory has been certified by the Canadian Grain Commission for these analyses.

2.2.6 Ploidy analysis

Nuclear DNA contents of the clubroot-resistant BC_1F_6 generation plants grown in the greenhouse were estimated using a flow cytometer (Partec GmbH, Munster, Germany) to assess their approximate chromosome number. For this, two leaf samples (each of about 0.5 cm^2) were collected from each plant at 3–4 weeks after seeding and the samples were chopped in an extraction buffer (Partec GmbH, Munster, Germany). The contents were filtered through a 50 μm Partec CellTrics disposable filter and a 1.6 mL nuclear fluorochrome DAPI staining buffer (4,6-diaminido-2-phenylindole, Sigma, product number D-9542) was added to each sample. After

incubation for 1–2 min, the samples were analyzed by a Partec Ploidy Analyzer (for detail, see Iftikhar et al. 2018). The mean value of the two samples was used for estimation of the approximate chromosome number of the plants.

2.2.7 Statistical analysis

Analysis of variance (ANOVA) was carried out using the software program 'R' (R Core Team 2020). In this, the lmer function of the 'lme4' package was used to fit a linear mixed-effects model for each trait. Least-square mean (lsmean) values were calculated using the 'lsmeans' package of R, and Tukey's test was carried out to compare the mean values for significant difference. Broad-sense heritability (*H*) was estimated using the following formula: $H = \delta^2 g / (\delta^2 g + \delta^2 ge/E + \delta^2 \varepsilon/ER)$, where $\delta^2 g$, $\delta^2 ge$ and $\delta^2 \varepsilon$ are the variance components for genotype, genotype × environment interaction and residual, respectively, and E and R are the number of environments and replications, respectively; the variance components were calculated using 'R'. Other statistical analyses, such as mean and standard error, were calculated using Microsoft Excel. Pearson's correlation, principal component analysis (PCA), and genotype by trait biplot (GT biplot) analysis for the relationship between the lines for resistance to different pathotypes, were carried out using the software program 'R'. The significance of correlations was tested using *t*-tests (Iversen and Gergen 1997).

2.3 Results

2.3.1 Development of B. napus lines carrying resistance to pathotype 3H

Thirty-five BC_1F_1 plants were grown in the greenhouse of which 14 were spring growth habit type and produced seeds under bag isolation; the remaining plants were either winter growth habit type (require vernalization for flowering) or failed to produce seeds under self-pollination. A total of 358 BC_1F_2 plants belonging to nine families (harvested from nine BC_1F_1 plants) were evaluated

of which 213 (59.5%) plants were found to be resistant and 145 (40.5%) susceptible (Table 2.1a). Self-pollinated seeds (BC_1F_3 seeds) of all 213 resistant BC_1F_2 plants were harvested. The BC_1F_3 and subsequent generation populations were grown from self-pollinated seeds harvested from the resistant plants only. In total, 594 BC_1F_3, 930 BC_1F_4, 767 BC_1F_5, and 286 BC_1F_6 plants, respectively, of 44, 108, 191, and 46 families were grown and evaluated for resistance to pathotype 3H. A decline in DSI from 41.4% in BC_1F_2 to 2.1% in the BC_1F_6 generation was observed, demonstrating the effectiveness of the selection for resistance to this pathotype in this interspecific population.

Table 2.1 Occurrence of plants resistant to *Plasmodiophora brassicae* (**a**) pathotype 3H and (**b**) pathotype 3A in different generation populations of the (*B. napus* × *B. rapa*) × *B. napus* interspecific cross.

Generation	Total families	Total plants	No. segregating families	No. non-segregating families	No. plants with disease score 0 to 3, and disease severity index (DSI)					Percent plants *	
					0	1	2	3	DSI (%)	Resistant	Susceptible
(a) Selection for resistance to pathotype 3H:											
BC_1F_2	9	358	5	4	173	40	30	115	41.4	59.5	40.5
BC_1F_3	44	594	23	21	436	43	35	80	19.8	80.6	19.3
BC_1F_4	108	930	53	55	587	96	57	190	27.9	73.4	26.5
BC_1F_5	191	767	20	171	725	9	1	32	4.6	95.6	4.3
BC_1F_6	46	286	0	46	280	0	0	6	2.1	97.9	2.1
(b) Selection for resistance to pathotype 3A:											
BC_1F_5	217	725	59	158	183	64	91	387	64.6	34.0	66.0
BC_1F_6	168	1109	53	115	90	24	36	959	89.3	10.3	89.7
BC_1F_7	91	1151	58	33	976	49	36	90	11.3	89.0	11.0
BC_1F_8	233	1561	11	222	1485	55	11	10	2.3	98.6	1.4

* Plants with disease scores 0 and 1 were classified as resistant while plants with scores 2 and 3 were classified as susceptible.

2.3.2 Development of B. napus lines carrying resistance to pathotype 3A

Due to the high sterility in early generations, the evaluation for resistance to pathotype 3A was delayed until the BC_1F_5 generation when sufficient seeds could be harvested from a good number of plants for the evaluation of resistance to pathotype 3H as well as 3A. A total of 725 BC_1F_5 plants belonging to 217 families were tested for resistance to pathotype 3A, and these families were found

to be resistant (disease score 0) to pathotype 3H. Of the 217 families, 59 (27.2%) were segregated for resistance and 158 (72.8%) were completely susceptible to this pathotype (Table 2.1b). Of the total BC_1F_5 population, 247 (34.0%) plants were found to be resistant to pathotype 3A. Thus, despite the selection performed in the earlier generations for resistance to pathotype 3H, several BC_1F_5 plants were found to carry resistance to pathotype 3A. In BC_1F_5 as well as in subsequent generations, the resistant plants were self-pollinated to grow the next generation population. A total of 1,109 BC_1F_6 plants belonging to 168 families were evaluated where only 114 (10.3%) plants were found to be resistant and the remaining 995 (89.7%) were susceptible. Thus, the proportion of resistant plants in BC_1F_6 generation declined as compared to the BC_1F_5 generation. However, in BC_1F_7 and BC_1F_8, respectively, 89.0% (1025/1151) and 98.6% (1540/1561) plants were found to be resistant (Table 2.1b). Thus, repeated selection for resistance to pathotype 3A reduced the DSI from 64.6% in BC_1F_5 to 2.3% in the BC_1F_8 generation.

2.3.3 Evaluation for resistance to clubroot under field conditions

A total of 341 BC_1F_8 lines, descended from 46 BC_1F_6 families homozygous for resistance to pathotype 3H, were evaluated in a clubroot-infested field (carrying pathotype 5X) in Spruce Grove, Alberta. However, resistance data from 315 lines were obtained; the number of plants for the remaining lines was less than the threshold number of plants needed for evaluation. Of these 315 lines, 168 (53.3%) were R, 10 (3.2%) MR, 6 (1.9%) MS, and 130 (41.3%) S (Figure 2.1A). Thus, the results demonstrated that the inbred *B. napus* lines derived from this interspecific cross carry resistance to multiple pathotypes including 5X.

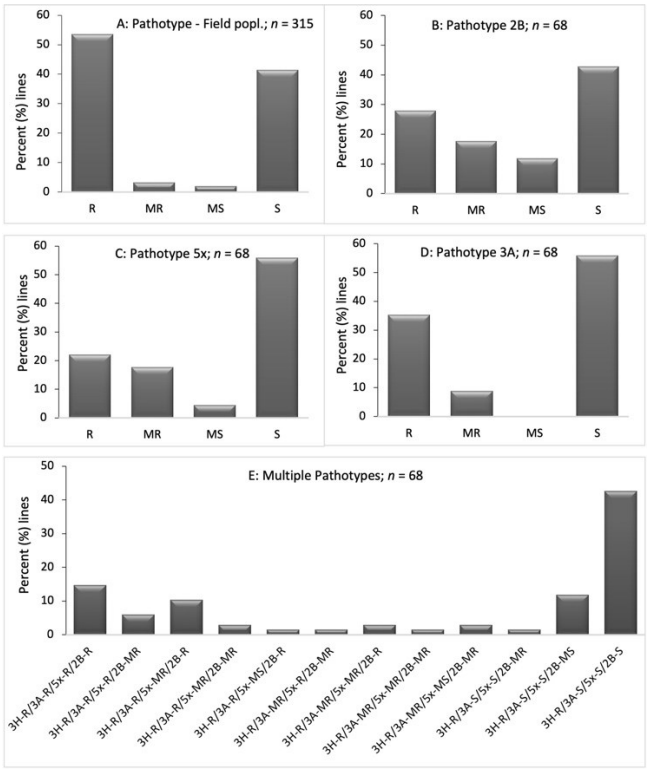


Figure 2.1 Distribution of the BC_1F_8 generation *Brassica napus* lines derived from the (*B. napus* × *B. rapa*) × *B. napus* interspecific cross exhibiting clubroot resistance under field conditions and to different *Plasmodiophora brassicae* pathotypes under greenhouse conditions. R = resistant (DSI ≤ 20%); MR = moderately resistant (DSI ≥ 20–40%); MS = moderately susceptible (DSI ≥ 40–70%); S = susceptible (DSI > 70%). '*n*' indicates the number of lines. (**A**) Field condition in Alberta, (**B**) pathotype 2B, (**C**) pathotype 5X (L-G1), (**D**) pathotype 3A, (**E**) multiple pathotypes (*x*-axis: 3H-R/3A-R/5X-R/2B-R = resistant to all pathotypes (3H, 3A, 5X and 2B); 3H-R/3A-R/5X-R/2B-MR = resistant to pathotypes 3H, 3A and 5X while moderately resistant to pathotype 2B; 3H-R/3A-R/5X-MR/2B-R = resistant to pathotypes 3H, 3A and 2B while moderately resistant to pathotype 5X; 3H-R/3A-R/5X-MR/2B-MR = resistant to pathotypes 3H and 3A while moderately resistant to pathotypes 5X and 2B; 3H-R/3A-R/5X-MS/2B-R = resistant to pathotypes 3H, 3A and 2B while moderately susceptible to pathotypes 2B; 3H-R/3A-MR/5X-R/2B-MR = moderately resistant to pathotypes 3A and 2B while resistant to pathotypes 3H and 2B; 3H-R/3A-MR/5X-MR/2B-R = moderately resistant to pathotypes 3A and 5x while resistant to pathotypes 3H and 2B; 3H-R/3A-MR/5X-MR/2B-MR = moderately resistant to pathotypes 3A, 5X, and 2B while resistant to 3H; 3H-R/3A-MR/5X-MS/2B-MR = resistant to 3H, moderately resistant to pathotypes 3A and 2B while moderately susceptible to pathotypes 5X; 3H-R/3A-S/5X-S/2B-MR = resistant to 3H, susceptible to pathotypes 3A and 5X while moderately resistant to pathotype 2B; 3H-R/3A-S/5X-S/2B-MS = resistant to 3H, susceptible to pathotypes 3A and 5X while susceptible to pathotype 2B; 3H-R/3A-S/5X-S/2B-S = resistant to 3H and susceptible to other three pathotypes).

2.3.4 Evaluation for resistance to multiple P. brassicae pathotypes in the greenhouse

Sixty-eight BC_1F_8 lines, homozygous for resistance to pathotype 3H, were evaluated for resistance to pathotype 2B, 5X and 3A, where nineteen (27.9%) lines were classified as R, eleven (17.6%) as MR, eight (11.8%) as MS and twenty-nine (42.7%) as S for resistance to pathotypes 2B (Figure 2.1B). In case of resistance to pathotype 5X, 15 (22.1%) were R, 12 (17.6%) were MR, 3 (4.4%) were MS, and 38 (55.9%) were S (Figure 2.1C); while for resistance to pathotype 3A, 24 (35.3%) were R, 6 (8.8%) were MR and 38 (55.9%) were S (Figure 2.1D). A joint analysis of the data of the 68 lines for resistance to multiple pathotypes showed that 10 (14.7%) lines were resistant to all four pathotypes; 29 (42.6%) were resistant to pathotype 3H but susceptible to pathotypes 2B, 3A, and 5X; and the remaining 29 (42.6%) lines were resistant to pathotype 3H but moderately resistant or susceptible to pathotypes 2B, 3A, or 5X (Figure 2.1E). This demonstrates that resistance to multiple pathotypes can be introgressed from a turnip into a *B. napus* canola through a *B. napus* × *B. rapa* interspecific cross.

Principal component analysis (PCA) was carried out to understand the nature of genetic variability among the BC_1F_8 lines for resistance to different *P. brassicae* pathotypes. The genotype by trait biplot analysis explained 98.2% of the total variance, where PC1 explained 96.4% of the total variance caused by genotype and PC2 explained 1.8% of the total variance which was due to the resistance to different pathotypes (Figure 2.2). The size of the vector for resistance to pathotype 3A was relatively longer as compared to other pathotypes, suggesting that this pathotype was the major discriminator of the BC_1F_8 population. The vectors for resistance to pathotype 3A and 5X had an angle of nearly zero ($r = 0.97$, $p < 0.01$) indicating that these two pathotypes accounted for a similar type of variation. The trait pairs, resistance to pathotype 3A vs. 2B ($r = 0.94$, $p < 0.01$), and resistance to 5X vs. 2B ($r = 0.95$, $p < 0.01$) had an acute (<90°) angle suggesting that their

variation was positively correlated. The trait pairs' resistance to pathotype 3H vs. 3A ($r = 0.02$, $p < 0.01$), 3H vs. 5X ($r = 0.08$, $p < 0.01$) and 3H vs. 2B ($r = 0.08$, $p < 0.01$) had a near-right angle, indicating that resistance to pathotype 3H was weakly or not correlated with resistance to the other pathotypes.

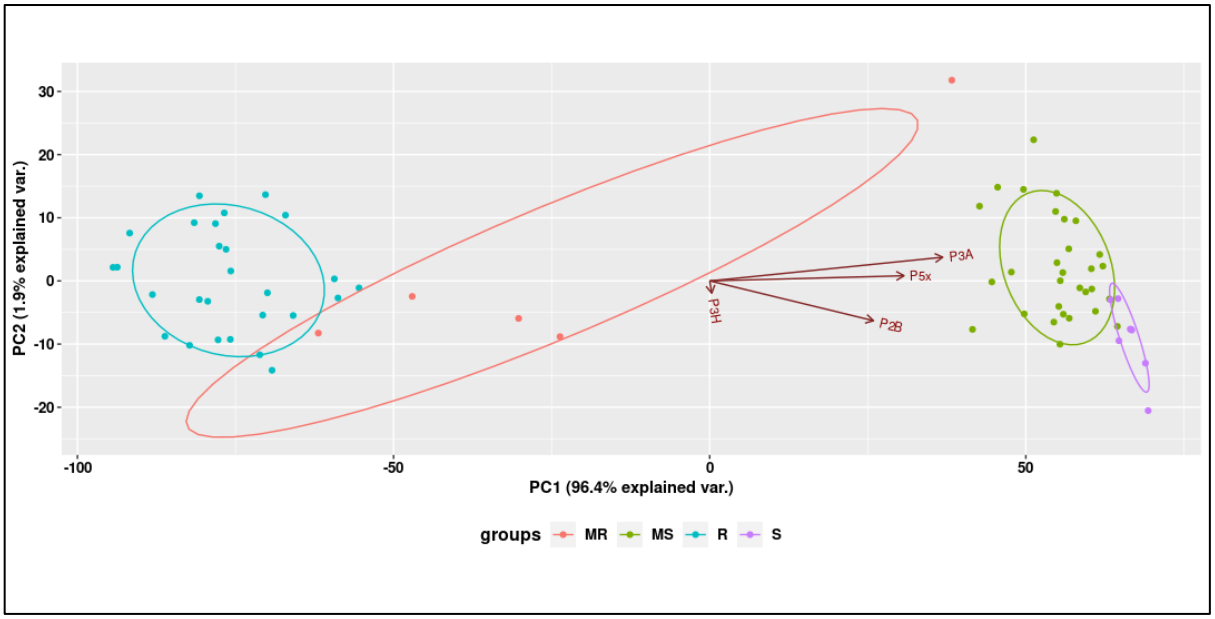


Figure 2.2 Principal component analysis (PCA) of the BC_1F_8 *B. napus* population ($n = 68$) derived from the (*B. napus* × *B. rapa*) × *B. napus* interspecific cross for resistance to *Plasmodiophora brassicae* pathotypes 2B (P2B), 3A (P3A), 3H (P3H), and 5X (P5X). This analysis placed the population into four groups based on disease severity index (DSI%) for these pathotypes.

2.3.5 Flow cytometric analysis

A total of 150 BC_1F_6 plants belonging to 46 families which showed resistance to pathotype 3H were analyzed for relative nuclear DNA content (Partec value); the parents *B. rapa* ECD 01 and *B. napus* A03-73NA were used as controls. Nuclear DNA content of the *B. rapa* parent was 161.5 ± 2.0 S.E. while it was 404.2 ± 1.2 S.E. for the *B. napus* parent (Table 2.2, Supplementary Figure S2.1). Relative nuclear DNA content of the BC_1F_6 population varied from 351.4 to 409.2 with a mean of 383.0 ± 0.9 S.E. which was close to the *B. napus* parent. This suggests that the majority of the BC_1F_6 plants had a chromosome number similar to the *B. napus* parent (Table 2.2).

Table 2.2 Partec values for the relative nuclear DNA content in the BC_1F_6 population of the (*B. napus* × *B. rapa*) × *B. napus* interspecific cross.

Population	Generation	No. plants	Range	Mean ± S.E.
(*B. napus* × *B. rapa*) × *B. napus*	BC_1F_6	150	351.4–409.2	383.0 ± 0.9
B. napus (A04-73NA)	Parent	6	400.3–408.2	404.2 ± 1.2
B. rapa (ECD 01)	Parent	6	154.3–169.2	161.5 ± 2.0

2.3.6 Agronomic and seed quality traits

Erucic acid content in the seed oil of the BC_1F_5 plants grown in a greenhouse varied from 0.01 to 30.9% with a mean of 18.1 ± 0.6% S.E. (Supplementary Figure S2.2). Of the total 240 plants evaluated, 17% (41/240) plants had <1% erucic acid in the seed oil, i.e., were of the zero-erucic acid type. All these zero-erucic acid plants were selected, and selection for additional zero erucic acid plants was performed in the progeny of the families containing <10% erucic acid in oil. All BC_1F_8 lines tested in the field nursery plots in 2020 were of the zero erucic acid type.

Days to flowering in the BC_1F_8 lines tested in the field nursery in 2020 varied from 46 to 60 days with a mean of 51.0 ± 0.2 days, which was about three days later than the spring *B. napus* parent A04-73NA (48.0 ± 0.2 days) (Supplementary Table S2.1a, Supplementary Figure S2.2). About 17% of the BC_1F_8 lines (76/455) flowered earlier than the *B. napus* parent. This population, on average, took 105.8 ± 0.2 days to mature (range 99 to 115 days); however, about 18% of the BC_1F_8 lines (83/455) matured earlier than A04-73NA. The oil content in the seeds of the BC_1F_8 lines varied from 40.6 to 51.5% with a mean of 46.2 ± 0.1% which was lower than the *B. napus* parent A04-73NA (48.7 ± 0.3%). However, the best BC_1F_8 line contained 51.5% oil, which was about 2.8% higher than the *B. napus* parent. In contrast, the mean seed protein of this population was about 1.6% higher (range 19.2 to 30.0%; mean 24.9 ± 0.1 S.E. %) than the *B. napus* parent (23.4 ± 0.3 S.E. %) (Supplementary Table S2.1a, Supplementary Figure S2.2). Glucosinolates in

this population varied from 7.8 to 31.2 µmol/g seed (Supplementary Table S2.1a, Supplementary Figure S2.2) with a mean of 14.4 ± 0.6 S.E. The confidence limits of GSL content for the *B. napus* canola parent were 11.6–12.4 µmol/g seed (n = 24; range 10.2–14.2 and mean = 12.0 ± 0.2 S.E. µmol/g seed). Based on this, about 31% of the BC_1F_8 lines had a GSL content either similar to or lower than the *B. napus* parent.

Based on clubroot resistance under field conditions in Spruce Grove, and the agronomic and seed quality data from the 2020 nursery plots, 112 lines were selected for tests in replicated yield trials in 2021. Analysis of variance of the 112 BC_1F_9 lines tested in the replicated field trials showed significant variation ($p < 0.001$) for seed yield, days to flowering, and days to maturity. The variation due to locations and due to the interaction between genotypes and locations (G × L) was significant ($p < 0.05$) for days to flowering; however, no significant variation was found between the locations for seed yield and days to maturity (Supplementary Table S2.2). Broad sense heritability for seed yield, days to flowering, and days to maturity were 0.67, 0.96 and 0.87, respectively. For the seed quality traits, significant variation was found for seed oil and GSL contents ($p < 0.001$) (Supplementary Table S2.2); G × L was significant ($p < 0.001$) for seed protein and GSL. Broad sense heritability for seed oil, protein and GSL were 0.81, 0.69, and 0.95, respectively.

Days to flowering in the BC_1F_9 population varied from 42 to 55 days with a mean of 48.7 ± 0.2 days when the check cultivar UA CountyGold took 49.0 ± 0.6 days to flower. About 50% of the lines flowered earlier than UA CountyGold and a large majority of the lines also matured earlier than this check. Seed yield of the lines varied from 1008 to 3626 kg/ha with a mean of 1,749 ± 17.8 kg/ha, which was similar to mean yield of the check (1744 ± 93.9 kg/ha); about 20% of the lines yielded similar or greater than the check. On average, the seed oil content in the BC_1F_9

population was lower, and seed protein, GSL, and saturated fatty acid contents were similar to the check; indeed, all 112 lines were of the zero erucic acid type (Figure 2.3). Thus, several canola-quality lines with acceptable agronomic and seed quality traits and exhibiting resistance to multiple *P. brassicae* pathotypes could be obtained from this interspecific cross.

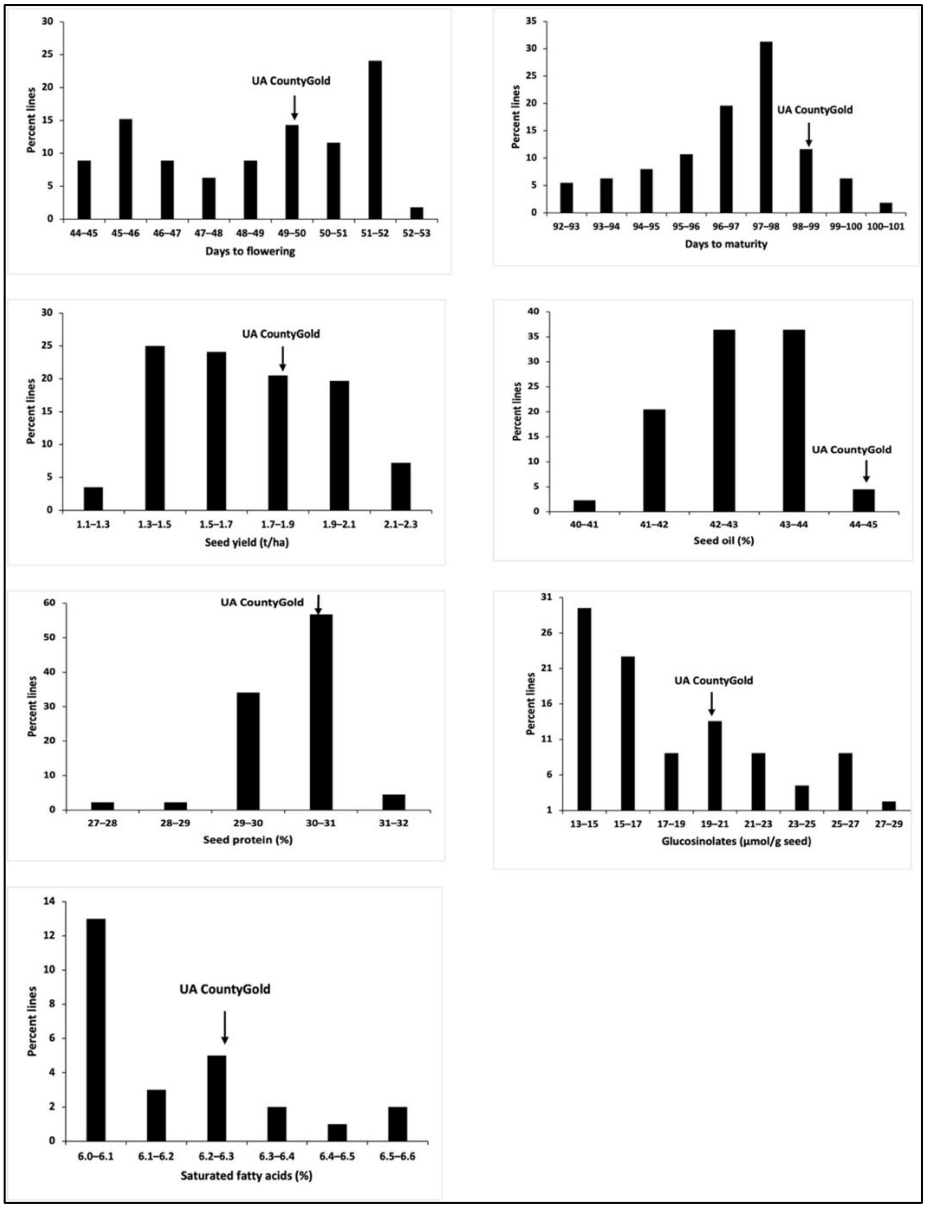


Figure 2.3 Distribution of the BC_1F_9 lines of the (*B. napus* × *B. rapa*) × *B. napus* interspecific cross grown in replicated field trials in 2021, for days to flowering (n = 112 lines), days to maturity (n = 112), and seed yield (n = 112), and seed oil (n = 44), protein (n = 44), glucosinolate (n = 44), and saturated fatty acids (n = 44). The value for the check cultivar UA CountyGold is indicated by the vertical arrow.

2.4 Discussion

Interspecific crosses between *B. napus* and *B. rapa* have been carried out by several researchers for different purposes (Leflon et al. 2006; Attri and Rahman 2018), including the introgression of clubroot resistance from *B. rapa* into *B*. *napus* (Chu et al. 2013; Hirani et al. 2016; Hasan et al. 2021a). Among the different forms of *B. rapa*, several accessions of European turnips (*B. rapa* ssp. *rapifera*) have been reported to carry clubroot resistance (Hasan et al. 2012; Peng et al. 2014b). Some of the European turnips have been used to develop clubroot-resistant Chinese cabbage (*B. rapa* ssp. *chinensis/pekinensis*) (Hirai et al. 2004), oilseed *B. napus* (Diederichsen and Sacristan 1996) and rutabagas (*B. napus* var. *napobrassica*) (Ayers and Lelacheur 1972; Bradshaw et al. 1997). Molecular mapping of clubroot resistance using turnip accessions identified several loci conferring resistance to different *P. brassicae* races or pathotypes. For example, Suwabe et al. (2003, 2006) mapped the loci *Crr1* and *Crr2,* conferring resistance to race two and four by using the *B. rapa* ssp. *rapifera* cv. Siloga. Similarly, by using the turnip cv. Debra two clubroot-resistant loci, *CRk* and *CRc*, conferring resistance to the isolates M85 and K08, and K08, respectively, were identified by Sakamoto et al. (2008). QTL analysis using resistance derived from the turnip cv. Milan White, Hirai et al. (2004) identified the locus *Crr3* conferring resistance to isolate Ano-01 (Saito et al. 2006). The turnip accession ECD 04 carries five clubroot-resistant loci where they mostly provide race-specific resistance to the *P. brassicae* strains *Pb2*, *Pb4*, *Pb7*, and *Pb10* (Chen et al. 2013). Hirani et al. (2018) reported that four turnip accessions of the ECD set, viz. ECD 01, ECD 02, ECD 03, and ECD 04, carry resistance to the Canadian field isolates. Similarly, Fredua-Agyeman et al. (2020) reported two major dominant loci, *CRa*/*CRb*^{Kato} and *Crr1* in ECD 02 exhibiting resistance to pathotypes 5X and 5G. Zhan et al. (2020) fine-mapped the clubroot-resistant locus *PbBa8.1* of ECD 04. Thus, it is apparent that several turnip accessions, including

the cv. Debra (ECD 01) may carry multiple clubroot-resistant loci and confer resistance to multiple pathotypes including 3H and 3A (Hasan et al. 2012; Yu et al. 2022) and the present study. However, possibility of the introgression of resistance to multiple pathotypes from a turnip into a *B. napus* canola through a *B. napus* × *B. rapa* interspecific cross has not been investigated so far. Our study demonstrates that resistance to multiple *P. brassicae* pathotypes can be introgressed from the turnip into the canola even when selection for resistance to a single pathotype is performed in an early segregating population.

According to Attri and Rahman (2018), a loss of about 55% of the *B. rapa* alleles can occur in the progeny of the *B. napus* × *B. rapa* interspecific cross by the F_4 generation. This is apparently due to a high meiotic anomaly in the early generation populations of this interspecific cross. Therefore, we backcrossed the di-genomic triploid F_1 plants (AAC) to the *B. napus* (AACC) canola parent to reduce meiotic anomalies and thus, to recover canola- quality *B. napus* plants without much difficulty. We applied selection for resistance to pathotype 3A first in the BC_1F_5 generation and were still able to develop *B. napus* lines carrying resistance to this pathotype. In fact, following this approach of selection for resistance to pathotype 3H in the early generations and for resistance to pathotype 3A in BC_1F_5, we were able to develop 10 BC_1F_8 canola lines carrying resistance to four pathotypes (2B, 3A, 3H, and 5X). The reconstituted clubroot-resistant lines derived from this *B. napus* × *B. rapa* interspecific cross were of the *B. napus* type as evident from the flow cytometric analysis for nuclear DNA content (Table 2.2). This agrees with the results reported by Attri and Rahman (2018) that most of the plants derived from the *B. napus* × *B. rapa* interspecific cross stabilize into the *B. napus* type.

In the present study, early generation interspecific population was selected for resistance to pathotype 3H; however, some of the lines also exhibited resistance to other pathotypes such as

3A, 2B, and 5X. This could have resulted from the pleiotropic effect of the 3H resistance to resistance to other pathotype, and the co-localization of multiple clubroot-resistant genes in the same genomic region conferring resistance to multiple pathotypes cannot be ruled out. Bi-plot and principal component analyses of the lines carrying resistance to multiple pathotypes revealed that the resistance to pathotypes 3A and 5X was under a similar genetic control, while a different genetic control was involved in the control of resistance to pathotype 3H. This was also evident from a strong correlation (r = 0.94–0.97) between the resistance to the 3A and 5X pathotypes, while there was a weaker correlation (r = 0.01–0.08) between the resistance to the 3H and 3A or 5X pathotypes. Thus, it is highly likely that, even though selection for resistance to pathotype 3H was performed in an early segregating generation, additional clubroot-resistant loci were introgressed into the inbred population derived from this interspecific cross.

The genetic control of the two canola-quality traits, erucic acid in seed oil and GSL in seed meal, is relatively simple compared to the genetic control of many other quantitative traits, such as days to flowering and maturity (Luo et al. 2014) and seed oil and protein contents (Liu et al. 2016). Erucic acid content in *B. napus* is controlled by two major loci located on A08 and C03 (Rahman et al. 2008) and these loci act in an additive (Harvey and Downey 1964) or partly dominant manner (Jönsson 1978); a homozygous recessive condition at both loci results in the zero erucic acid phenotype. Seed GSL content in *B. napus* is controlled by at least four gene loci (Rahman et al. 2001) where about half of the loci are located in the A genome (for review, see Rahman et al. 2014); a homozygous recessive condition at all loci results in the low GSL phenotype. This genetic control of the two canola-quality traits and the breeding technique applied in this study, viz., backcrossing the F_1 to the canola-quality parent ($A^cA^tC^c \times A^cA^cC^cC^c$, where the superscripted *c* indicates the alleles for the canola-quality traits, and *t* indicates the turnip (non-

canola quality) alleles), facilitated achievement of canola-quality clubroot-resistant lines from this interspecific cross.

2.5 Conclusion

The results reported in this chapter demonstrate that a canola-quality *B. napus* line carrying resistance to multiple *P. brassicae* pathotypes can be achieved from a (*B. napus* × *B. rapa*) × *B. napus* interspecific cross through selection for resistance to a single pathotype in the early generation population. The occurrence of resistance to multiple pathotypes in the advanced generation population and the strong correlation between resistance to some of these pathotypes indicates that a common genetic mechanism might be involved in the control of resistance to multiple pathotypes. A pleiotropic effect of a resistance gene conferring resistance to multiple pathotypes also cannot be ruled out; however, all these hypotheses need to be confirmed experimentally. Nevertheless, the knowledge gained from this study could be used by Brassica researchers and breeders for the introgression of clubroot resistance as well as other traits from *B. rapa* into *B. napus* canola.

Chapter 3. Mapping and molecular analysis of clubroot resistance of *Brassica rapa* cv. Debra introgressed into *Brassica napus*

3.1 Introduction

Clubroot disease, caused by *Plasmodiophora brassicae* Woronin, is an ever-increasing global concern for negative economic impact on the crops of the family Brassicaceae (for review, see Dixon 2009). This disease ranks among the top 10 significant threats to Brassica oilseed production (for review, see Zheng et al. 2020). The pathogen *P. brassicae* attacks the roots causing hyperplasia and hypertrophy of the cells resulting in the formation of galls (Kageyama and Asano 2009; Liu et al. 2020). This results in above-ground symptoms like stunted growth, wilting and yellowing of the plants due to the impaired uptake of water and nutrients (Korbas et al. 2009). The seed yield and oil content are adversely affected by this disease (Pageau et al. 2006). Moreover, clubroot also increases the chlorophyll content in the seed oil by about 50% (Engqvist 1994), and thus, adversely affecting the oil quality.

The virulence pattern of the pathogen *P. brassicae* keeps changing with the emergence of new genetically distinct pathotypes in different countries, including Canada (Strelkov et al. 2016b; 2018; Cao et al. 2020). About 17 pathotypes have been identified in Canada based on the Canadian Clubroot Differential (CCD) set developed by Strelkov et al. (2018). Currently, a total of 37 pathotypes have been reported in Canada, all of which show different virulence patterns on canola (Hollman et al. 2021). Among these, the pathotypes 3A and 3D are the most dominant at present although 3H was originally the most prevalent (Askarian et al. 2021; Dakouri et al. 2021; Hollman et al. 2021); 8C and 2B are new virulent pathotypes recently found in Western Canada (Strelkov et al. 2021). Pathotype 3H is still the most prevalent in Saskatchewan and Manitoba (Hollman et al. 2021). The rapid evolution of new pathotypes on the prairies requires adopting effective

strategies for the management of this disease. One of the most reliable methods for the control of clubroot disease is the deployment of resistant cultivars (for review, see Rahman et al. 2014b; Peng et al. 2015). In addition, pyramiding of multiple resistance genes along with adequate crop rotation have been considered the essential components of an integrated approach to the management of this disease (Donald and Porter 2009; Peng et al. 2015). Therefore, introgression of multiple resistance genes into canola to increase the durability of clubroot resistance is extremely important.

Clubroot resistance (CR) has been reported in several germplasm accessions belonging to the "triangle of U" (U, 1935), including the diploid species *B. rapa* (AA, 2*n* = 20), *B. oleracea* (CC, 2*n* = 18) and *B. nigra* (BB, 2*n* = 16) and the allopolyploid species *B. napus* (AACC, 2*n* = 38) (Karling 1969; Johnston 1970; Ayers and Lelacheur 1972; Buczacki et al. 1975; Werner et al. 2008; Hasan et al. 2012; Peng et al. 2014a, 2018; Fredua-Agyeman and Rahman 2016; Hasan and Rahman 2016; Lee et al. 2016; Laperche et al. 2017; Dakouri et al. 2018; Chang et al. 2019; Fredua-Agyeman et al. 2019; Farid et al. 2020). Among these, very limited accessions of *B. napus* carry resistance and, therefore, introgression of CR genes from its allied species, especially from *B. rapa*, has been considered by several researchers (Hirani et al. 2016; Hasan et al. 2021c; Kaur et al. 2022). In *B. rapa*, resistance is most commonly found in the European fodder turnip (*B. rapa* ssp. *rapifera*) cultivars such as Gelria R, Siloga, Debra, Milan White and ECD 04 (for review, see Hirai et al. 2006). Debra has been included in the European Clubroot Differential (ECD) set as ECD 01 (Buczacki et al. 1975; Diederichsen et al. 2009) and has been reported to carry resistance to the recently evolved pathotypes, such as 3A (Kaur et al. 2022). Therefore, there are diverse sources of clubroot resistance genes that may be potentially introgressed into economically important oilseed Brassica such as canola.

To date, about 29 CR loci/quantitative trait loci (QTL) have been mapped on seven A genome chromosomes of *B. rapa* through genetic analysis and QTL mapping approaches (for review, see Hasan et al. 2021a). About half of these loci have been mapped onto two major regions of A03 i.e., at about 15.5-16.3 Mb and 23.8-26.8 Mb of the *B. rapa* Chiifu-401 whole genome assembly v3.0 (for review, see Hasan et al. 2021a). A modifier gene, *Crr2* (Suwabe et al. 2003, 2006) and two QTL *CR6a* (Lee et al. 2002) and *PbBa1.1* (Chen et al. 2013) were identified on A01; a QTL (possibly a major locus) *CRc* (Sakamoto et al. 2008) and a QTL *Rcr8* (Yu et al. 2017) were mapped on A02; seven major loci, viz. *CRd* (Pang et al. 2018), *Crr3* (Hirai et al. 2004; Saito et al. 2006), *Rcr5* (Huang et al. 2019), *Rcr1* (Chu et al. 2014), *CRb* (Piao et al. 2004; Cho et al. 2008; Kato et al. 2012, 2013), *Rcr2* (Huang et al. 2017), *Cra* (Ueno et al. 2012) and seven QTL namely *CRk* (Sakamoto et al. 2008), *Rcr4* (Yu et al. 2017), *CRq* (Yuan et al. 2015; Wei et al. 2022), *CR6b* (Lee et al. 2002), *Bcr1* and *Bcr2* (Zhang et al. 2022), $Rcr10^{ECD01}$ (Yu et al. 2022) located on chromosome A03; a locus *CrrA5* (Nguyen et al. 2018) mapped on A05; a QTL *Crr4* (Suwabe et al. 2006) mapped on A06; a QTL *qBrCR38-1* (Zhu et al. 2019b) mapped on chromosome A07; four loci viz. *CRs* (Laila et al. 2019), *Rcr3* (Karim et al. 2020), *Crr1a* (Suwabe et al. 2003, 2006; Hatakeyama et al. 2013), *Crr1b* (Hatakeyama et al. 2013) and three QTL viz. *Rcr9* (Yu et al. 2017; Karim et al. 2020) and *qBrCR38-2* (Zhu et al. 2019b), $Rcr9^{ECD01}$ (Yu et al. 2022) mapped on chromosome A08. Many different classes of resistance genes have been reported in plants (for review, see Gururani et al. 2012); among these, the Toll/interleukin receptor (TIR)-nucleotide binding site (NBS)-leucine rich repeat (LRR) is the largest class. TIR-NBS-LRR protein genes of *Crr1*, *CRa* and CRb^{kato} have been identified by cloning (Ueno et al. 2012; Hatakeyama et al. 2013, 2017). All the above-mentioned loci/QTL provide resistance against *P. brassicae* either individually or through a combined effect.

Of the above-mentioned genes/QTL, the *B. rapa* ssp. *rapifera* Debra/ECD 01 is reported to carry the locus *CRc* conferring resistance to *P. brassicae* pathotype 4 and *CRk* conferring resistance to pathotype 2 and isolate K04 (Piao et al. 2004; Sakamoto et al. 2008). Recently, Yu et al. (2022) identified two QTL, $Rcr9^{ECD01}$ and $Rcr10^{ECD01}$, on chromosomes A08 and A03 respectively, by using a population derived from ECD 01 (resistant) × *B. napus* canola (susceptible) interspecific cross. These two loci provide resistance to pathotypes 3A, 3D and 3H, while $Rcr9^{ECD01}$ alone provide resistance to pathotype 5X (Yu et al. 2022). Hirani et al. (2018) also reported two loci, *BraA.CR.a* and *Bra.CR.b*, on A03 and A08 respectively, while working with ECD 01. This implies that ECD 01 carries CR loci which can confer resistance to the recently evolved virulent pathotypes such as 3A and 5X. The objective of the present study is to understand the genetic and molecular basis of this resistance to *P. brassicae* pathotype 3H and 3A by using a spring *B. napus* canola population carrying clubroot resistance of ECD 01.

3.2 Materials and Methods

3.2.1 Plant materials

The recombinant inbred line (RIL) population used in this study was generated from an interspecific cross between the clubroot resistant European turnip accession ECD 01 (*Brassica rapa* var. *rapifera*) ($2n = 20$; AA) and clubroot susceptible *B. napus* spring canola line A03-73NA ($2n = 38$; AACC). The details of development of this population have been described elsewhere (Kaur et al. 2022). Briefly, the F_1 plants were backcrossed to the canola parent and the BC_1 plants were self-pollinated for seven generations to obtain the euploid *B. napus* RIL population; this population included the lines either homozygous resistant or susceptible to pathotypes 3H and 3A. Fifty BC_1F_6 homozygous RILs resistant to pathotype 3H, 50 BC_1F_8 homozygous RILs resistant to pathotype 3A, 50 BC_1F_4/BC_1F_5 homozygous RILs susceptible to pathotype 3H and 50 BC_1F_7

homozygous RILs susceptible to pathotype 3A were used for whole genome re-sequencing (WGRS). The 50 RILs resistant to pathotype 3H were derived from nine BC_1F_2 plants, and the 50 RILs resistant to pathotype 3A were derived from 15 BC_1F_5 plants. The 50 RILs susceptible to pathotype 3H and 3A were derived from 14 BC_1F_2 plants.

3.2.2 *Phenotypic evaluation for clubroot resistance*

The *P. brassicae* pathotypes 3H and 3A that we used in this study were obtained from Dr. Stephen Strelkov, Department of Agricultural, Food and Nutritional Science, University of Alberta. The parents and the RILs were tested for resistance to these two pathotypes, and the details, including inoculum preparation, inoculation and scoring the plants can be found elsewhere (Kaur et al. 2022). Briefly, seedlings at an age of 7-10 days after germination were inoculated with 1 ml inoculum at their base on two consecutive days and the soil was kept well saturated with water to maintain favourable conditions for infection. The severity of clubroot disease was evaluated on a 0-3 scale, where 0 = no galls on the roots, 1 = tiny galls on the lateral roots, 2 = small to medium-sized galls on 1/3 to 2/3rd of the lateral roots and 3 = severe galls on the lateral as well as main root. Disease severity index (DSI) of different generation lines was calculated using the following formula:

$$\text{DSI}\ (\%) = \frac{\sum(n\ \times 0 + n \times 1 + n \times 2 + n \times 3)}{N\ \times\ 3} \times 100$$

where, *n* is the number of plants in each disease severity class, and 0, 1, 2 and 3 are the disease symptom severity classes, and *N* is the total number of plants (Kaur et al. 2022). A DSI of 0% was considered resistant (R) and more than 70% was considered susceptible (S).

3.2.3 DNA extraction and whole-genome resequencing (WGRS)

Genomic DNA of the parents and the RILs was extracted using modified SDS DNA extraction protocol (Kotchoni and Gachomo 2009). In order to accomplish this, young leaves of 3-4 weeks

old plants were collected in 2 ml Eppendorf tubes, and the tissue was homogenized to a fine powder using tissue-lyzer (Qiagen, San Diego, USA). The quality and quantity of DNA was determined using a ND-2000 Nanodrop spectrophotometer (Nanodrop Technologies Inc., Wilmington, Del., USA) and Qubit 3.0 fluorometer (ThermoFisher Scientific Inc., USA).

For WGRS, the following six DNA samples were prepared: (i) resistant parent ECD 01 (*B. rapa* var. *rapifera*) (RP) (resistant to pathotypes 3H and 3A), (ii) susceptible parent (*B. napus* canola line A03-73NA) (SP) (susceptible to both pathotypes), (iii) a bulk DNA of the RILs resistant to pathotype 3H (RB-3), (iv) a bulk DNA of the RILs susceptible to pathotype 3H (SB-3), (v) a bulk DNA of the RILs resistant to pathotype 3A (RB-3A), and (vi) a bulk DNA of the RILs susceptible to pathotype 3A (SB-3A). All four bulks (iii to vi) were prepared by mixing equal amount of DNA from 50 RILs; for each RIL seven plants were used. For sequencing, DNA libraries were prepared using the TruSeq DNA Sample Prep Kit (Illumina) protocol, and sequencing was carried out using Illumina HiSeq 4000 platform to generate 150 bp paired end reads. Library preparation, sequencing and data analysis, including alignment of the sequence reads to the reference genome, and single nucleotide polymorphism (SNP) calling were carried out by Genewiz (Genewiz; www.genewiz.com). All six samples were sequenced 30X in order to increase the sequencing depth and breadth and, consequently, more accurate results.

The adapter sequences and nucleotides with poor quality were trimmed off using Trimmomatic v0.36 (Bolger et al. 2014); only the high-quality sequencing reads were used for downstream analysis. Reads shorter than 35 bp were discarded. The trimmed reads of the sample RP were mapped to two reference genomes, the winter *B. napus* cultivar Darmor-*bzh* whole-genome assembly v4.1 (Chalhoub et al. 2014) and the *B. rapa* cv. Chiifu-401 whole-genome assembly v3.0, using Burrows-Wheeler Aligner (BWA) v0.7.12 (Li and Durbin, 2009a). The data

obtained from BWA alignment was processed and the mapped reads at each nucleotide position were extracted using the mpileup function in Samtools v1.3.1 (Li et al. 2009b). The variants between the RP reads and the A genome of the reference genome assemblies were called using VarScan v2.3.9 (Koboldt et al. 2012) with the following settings: minimum coverage 10, minimum number of reads 7, minimum variant frequency 10%, and the highest *p*-value 0.05. Based on this, two sequence assemblies of the RP (referred to as RPA), viz. RPA-Chiifu, which is based on the reference genome *B. rapa* cv. Chiifu-401, and RPA-Darmor, which is based on the reference genome *B. napus* cv. Darmor-*bzh*, were developed. The trimmed reads of the four bulks, RB-3, SB-3, RB-3A and SB-3A, were individually aligned to the two RPAs and SNPs were called for these bulks.

3.2.4 Calculation of SNP index and ΔSNP index

The SNP index is the ratio of SNP variants at each SNP position. This was calculated for the four bulks RB-3, SB-3, RB-3A and SB-3A separately at each SNP position by comparing with the two RPA assemblies using the following formula (Abe et al. 2012):

$$\text{SNP index} = \frac{\text{Counts of alternate (alternate to RPA) nucleotide reads in bulk}}{\text{Total counts of nucleotide reads in bulk}}$$

The following filtering criteria were used for the calculation of SNP index: (i) SNP positions should have the read depth ≥ 10, (ii) the susceptible parent (SP) carries the allele in a homozygous state, (iii) SNP indices should not be ≤ 0.25 for susceptible bulks and ≥ 0.75 for resistant bulks. The ΔSNP index was then calculated by subtracting the SNP index of the resistant bulk from the SNP index of the susceptible bulk. The SNP positions with ΔSNP index = -1 were considered the SNPs associated with the trait of interest. The ΔSNP index of -1 denotes that the allele in resistant

bulk is the same as that of the resistant parent while a different allele is present in the susceptible bulk.

3.2.5 Designing of primers and polymerase chain reaction (PCR)

A total of 359 simple sequence repeat (SSR) primers, and 32 allele-specific primers (AS-primers) from chromosome A02, designed based on ΔSNP index of the resistant and susceptible bulks using the software program BatchPrimer3 (You et al. 2008) and were used to genotype the RIL population carrying resistance to pathotype 3A (Supplementary Table S3.1a and S3.1b). This included the primers for 102 loci from 0-6 and 14-24 Mb region of A02, 138 loci from 0-4, 8-10 and 24-32 Mb region of A03, 13 loci from 2-8 and 10-20 Mb region of A08, and 106 loci from 10-14, 20-34 and 36-46 Mb region of A09. While designing the SSR primers, the universal M13 sequence 5'-CACGACGTTGTAAAACGAC-3' was added to the forward primer. The primers were fluorescently labelled with dyes FAM, VIC, NED and PET (IRD; MWG-Biotech AG, Edensburg, Germany).

Polymerase chain reaction (PCR) was carried out in a total reaction volume of 14 µl, where each reaction mixture contained 50 ng (25 ng/µl × 2 µl) of DNA, 2.5 µl of 1X PCR buffer, 0.125 µl of 20 mM dNTPs, 1.25 µl of 50 mM $MgCl_2$, 0.5 µl each of the forward and reverse primers, 0.25 pmol/L of the labelling dye (FAM, VIC, NED and PET; Applied Biosystem, Foster City, Calif.), 0.125U of GoTaq DNA polymerase and 6.75 µl of nuclease free water. The amplified products were detected using a capillary ABI Sequencer No. 3730 (Applied Biosystems) and were scored and analysed using GeneMapper software v5.0. Alternatively, the amplified products of the AS-primers were stained with SYBR Green1 and separated on 2.5-3% agarose gel and were imaged using Typhoon FLA9500 laser scanner (GE Heathcare).

3.2.6 Identification of polymorphic markers, genotyping, and mapping of the clubroot resistance loci

The above-mentioned SSR and AS markers were tested on the parents ECD 01 and A03-73NA to identify the polymorphic markers, and the polymorphic markers were used to genotype the entire RIL mapping population which included 48 lines carrying resistance to pathotype 3A and 48 susceptible lines. Percent polymorphic markers was calculated using the following formula:

$$\text{Percent polymorphic markers} = \frac{\text{No. polymorphic marker}}{\text{Total no. marker tested} - \text{No. of marker failed to amplify in both parents}} \times 100$$

By using the genotypic data, a genetic map was constructed using the software program QTL Icimapping v4.2 (Meng et al. 2015) for the chromosomes found to be harboring clubroot resistance based on the WGRS results. The genetic distance between the markers was calculated using Kosambi function (Kosambi 1943) with recombination frequency of 0.45 and the logarithm of odds (LOD) threshold of 3.0.

The additive effect and the percent phenotypic variation explained by the QTL were obtained from Inclusive Composite Interval Mapping of ADDitive QTL (ICIM-ADD) analysis (Li et al. 2007) using the software program QTL IciMapping 4.2 (https://www.isbreeding.net; Meng et al. 2015; Hasan et al. 2021b). The criteria set for this analysis were a stepwise regression probability of 0.001, 0.1 cM walking speed, window size of 10 cM, and 5 cM for the control markers (Manichaikul et al. 2009). The method of Churchill and Deorge (1994) was used to confirm the locus ($p < 0.05$) by analysing 1000 permutations of the data.

Single marker analysis was carried out using the genotypic data of the markers and DSI for resistance to pathotype 3A of the RIL population. QTL Icimapping v4.2 was used for this analysis.

The likelihood of marker-trait association was performed following a stepwise regression probability of 0.001. The empirical threshold for LOD (p-value < 0.05) was obtained from 1,000 permutations to declare the marker-trait association.

Genome-wide association study (GWAS) was performed using the weighted mixed linear model in TASSEL v5.2.84 (Bradbury et al. 2007) to identify the genomic regions associated with resistance to clubroot disease. The analyses were performed using the SNP genotype data, the kinship matrix to account for relatedness, the first 3 PCs from principal components as covariates to account for population structure and the disease severity index (DSI) computed per experiment and combined across all experiments. The threshold for declaring significant marker-trait association was set to $p < 1.0 \times 10^{-4}$ or $\log_{10}(1/p)$ value ≥ 4.

3.2.7 Construction of a physical map and identification of putative genes

The forward and reverse sequences of the primers were aligned to the *B. rapa* Chiifu-401 whole genome assembly v3.0 using BLASTn tool in Brassica database (BRAD; https://brassicadb.cn) to determine the physical position of the markers. Following the alignment, the length between the forward and the reverse primer was deduced and if it was similar to the known amplicon size, the position of the marker was considered its physical position in the genome. The sequences of 244 genes related to disease resistance listed in Brassicaceae Database (BRAD) were downloaded (http://brassicadb.cn/#/ResistanceGene/). The 'Blat' function of Genoscope *Brassica napus* Genome Browser (https://www.genoscope.cns.fr/blat-server/cgi-bin/colza/webBlat) was used to identify their homologues in *B. napus*. We filtered the homologous *B. napus* genes located within the QTL regions identified in this study. We used SMART (http://smart.embl-heidelberg.de/) to confirm the putative function of these genes.

3.3 Results

3.3.1 Inheritance of clubroot resistance

A total of 286 plants belonging to 46 BC_1F_6 families were evaluated for resistance to pathotype 3H, where 280 (97.90%) plants were found to be resistant and 6 (2.09%) susceptible. Progeny of 50 resistant plants, i.e., 50 BC_1F_7 RILs, descended from 46 BC_1F_6 families were used to constitute the bulk RB-3 for WGRS; all these 50 BC_1F_7 RILs were confirmed to be homozygous for resistance. Fifty BC_1F_5/BC_1F_6 RILs, descended from 190 BC_1F_4 and 32 BC_1F_5 plants with disease score of 3, were pooled to constitute the SB-3 bulk. A total of 1561 plants belonging to 233 BC_1F_8 families were evaluated for resistance to pathotype 3A, where 1485 (95.13%) plants were found to be resistant and 10 (0.64%) susceptible. To constitute the RB-3A bulk, 50 BC_1F_9 RILs descended from 50 resistant (disease score 0) plants belonging to 110 BC_1F_8 families were used. Fifty BC_1F_8 RILs descended from 90 BC_1F_7 plants with disease score of 3 were pooled to constitute the SB-3A bulk.

3.3.2 Whole genome resequencing analysis

The high-throughput WGRS generated about 81.8 to 192.8 million 150 bp pair-end raw reads from the two contrasting parents (RP and SP) and four bulks (RB-3, SB-3, RB-3A and SB-3A). Among these, 94.6 to 99.3% of the reads could be mapped to the two reference genomes, *B. napus* cv. Darmor-bzh v.4.1 and *B. rapa* cv. Chiifu-401 v3.0 (Supplementary Table S3.2).

The total number of polymorphic SNPs detected from WGRS and by using the two reference genomes (Chiifu-401 v3.0 and Darmor-bzh v.4.1) are presented in Table 3.1. While using the *B. rapa* cv. Chiifu-401 v3.0 reference genome, about 80% greater number of polymorphic SNPs were detected as compared to the number of SNPs detected using the *B. napus* cv. Darmor v4.1 reference genome. Based on ΔSNP index = -1, about 4,000 SNPs were identified to be associated with resistance to pathotype 3H and about 2,500 SNPs to be associated with

resistance to 3A (Table 3.1). The number of SNPs with ΔSNP index = -1 detected by using the two reference genomes gave similar results (Table 3.1).

Table 3.1 Total number of polymorphic SNPs and SNPs with ΔSNP index = -1 detected by using *B. napus* cv. Darmor-bzh v.4.1 and *B. rapa* cv. Chiifu-401 v3.0 genome assemblies.

	Total number of polymorphic SNPs	Total number of SNPs with ΔSNP index = -1
Pathotype 3A resistance		
B. napus cv. Darmor-bzh v4.1	459,777	2,690
B. rapa cv. Chiifu-401 v3.0	821,784	2,319
Pathotype 3H resistance		
B. napus cv. Darmor-bzh v4.1	775,262	4,793
B. rapa cv. Chiifu-401 v3.0	1,416,058	3,270

3.3.3 Distribution of the SNPs associated with clubroot disease resistance on the chromosomes

The distribution of the SNPs with ΔSNP index = -1 associated with resistance to pathotype 3H on the 19 *B. napus* chromosomes and 10 A-genome chromosomes presented in Figure 3.1a and 3.1b. For resistance to pathotype 3H, the greatest number of SNPs were detected from chromosome A03 (33.88 or 40.76%) followed by A08 (9.85 or 12.37%). The SNPs of A03 are largely concentrated in the genomic region of about 16-30 Mb position, and the SNPs of A08 are largely concentrated on 2-14 Mb region (Figure 3.1c and 3.1d). These two genomic regions have been reported to carry clubroot resistance (for review, see Hasan et al. 2021a); therefore, we did not continue further with map construction and detection of QTL for resistance to pathotype 3H.

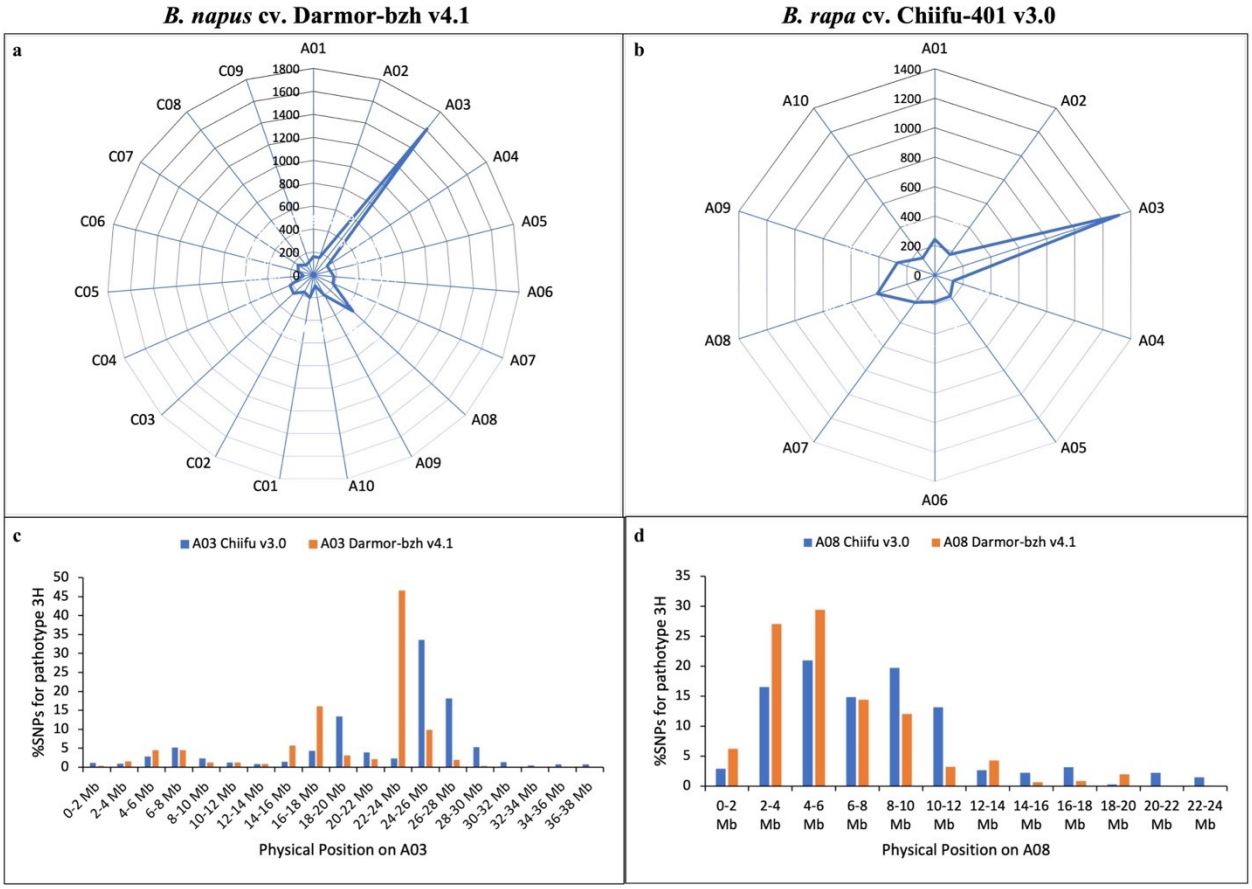


Figure 3.1 (a, b) The distribution of the SNPs with ΔSNP index = -1 associated with resistance to *Plasmodiaphora brassicae* pathotype 3H on *Brassica* chromosomes based on *B. napus* cv. Darmor-bzh v4.1 and *B. rapa* cv. Chiifu-*401* v3.0. (a, b) genome assemblies. (c, d) The distribution of the SNPs with ΔSNP index = -1 from chromosomes A03 and A08.

For resistance to pathotype 3A, the greatest number of SNPs were detected from chromosomes A02 (17.76 or 20.5%) followed by A09 (12.22 or 19.58%); however, a considerable number of SNPs were also detected from A03 (7.16 or 11.0%) (Figure 3.2a and 3.2b). This indicates that these three A genome chromosomes might carry resistance to this pathotype. SNPs in the C genome chromosomes using the Darmor-bzh v.4.1 genome assembly could also be detected; however, their occurrence in most cases was low (less than 5% of the total). Also, the clubroot resistance in the resistant lines was introgressed from the A genome of *B. rapa*; therefore, the C-genome SNPs were ignored. According to Wang et al. (2022), the clubroot resistance locus

of the chromosome A08 plays a role in resistance to multiple pathotypes including 3A; therefore, we continued our investigation on SNPs from A08 chromosome as well. The SNPs of A02 were mostly concentrated in the genomic region of about 0-6 and 14-24 Mb regions (Figure 3.2c), the A03 SNPs were mostly distributed in a wide region with peaks at about 0-4, 8-10 and 24-32 Mb regions (Figure 3.2d), the A08 SNPs were mostly distributed at 2-8 and 10-20 Mb regions (Figure 3.2e), and the A09 SNPs were distributed in a wide region with peaks at about 10-14, 20-24, 30-34 and 36-46 Mb regions (Figure 3.2f) indicating that these genomic regions might be associated with resistance to pathotype 3A. Mapping of these chromosome regions was carried out by genotyping the mapping population with molecular markers from these chromosomes.

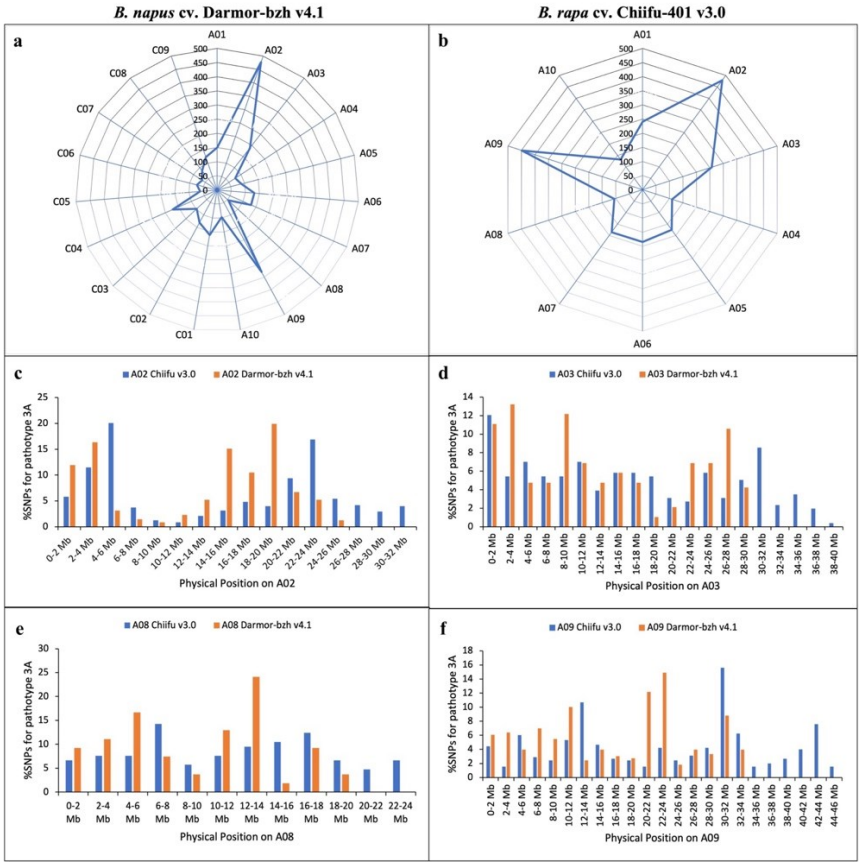


Figure 3.2 (a, b) The distribution of the SNPs with ΔSNP index = -1 associated with resistance to *Plasmodiaphora brassicae* pathotype 3A on *Brassica* chromosomes based on *B. napus* cv. Darmor-bzh v4.1 and *B. rapa* cv. Chiifu-*401* v3.0. (a, b) genome assemblies. (c, d, e, f) The distribution of the SNPs with ΔSNP index = -1 from the chromosome A02, A03, A08 and A09.

3.3.4 Marker development and mapping of the loci conferring resistance to P. brassicae pathotype 3A

A total of 359 SSR markers from chromosome A02, A03, A08 and A09 and 32 SNP allele-specific primers (SNP-AS; two forward primers and one common reverse primer for each SNP) from chromosome A02 were tested for polymorphism between the parents (ECD 01 and A04-73NA) where 70.2% markers were found to be polymorphic (Supplementary Table S3.3).

Among the 102, 138, 13 and 106 SSR markers from chromosome A02, A03, A08 and A09, we observed that 50, 115, 9 and 50 markers produced polymorphic alleles between the parents (Supplementary Table S3.3). Fourteen out of 32 SNP-AS markers from chromosome A02 also exhibited parental polymorphism. These 238 polymorphic markers were used to genotype the mapping population which included 48 resistant and 48 susceptible RILs, and the genotypic data was used for construction of partial genetic linkage map of the four chromosomes (Figure 3.3). The physical map of all four chromosomes was constructed based on the physical positions of the markers from the linkage map (Figure 3.3). The A02 linkage map included 22 markers, A03 included 13 markers, A08 included 9 markers, and A09 included 19 markers.

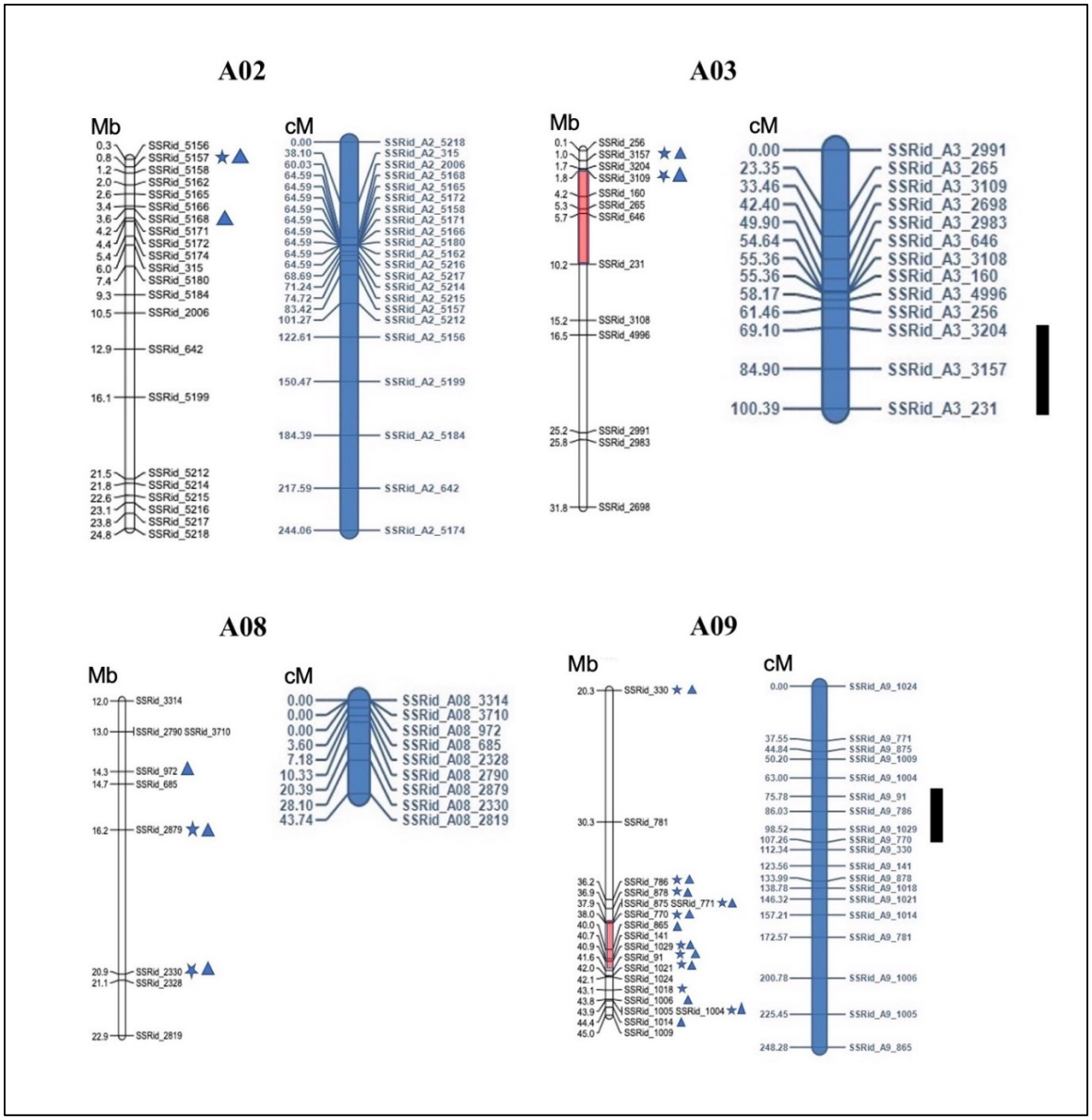


Figure 3.3 The physical (left) and partial genetic linkage maps (right) of the *Brassica* A genome chromosomes A02, A03, A08 and A09. Physical maps were constructed based on position of the markers on *B. rapa* cv. Chiifu-*401* v3.0 genome assembly and presented as million base pairs (Mb), and the genetic map was constructed using software JoinMap 4.1 (Van Ooijen 2011) and presented as centimorgan (cM). The position of the loci associated with resistance to *Plasmodiophora brassicae* pathotype 3A are shown on the linkage maps with black solid lines adjacent to the markers. In case of position on the physical maps, the red color indicates the genomic region associated with resistance identified through QTL mapping; the blue star and delta adjacent to the markers indicates the genomic regions associated with resistance identified through single marker analysis (SMA) and genome-wide association study (GWAS) approaches, respectively.

By using the genotypic and phenotypic (resistance to pathotype 3A) data, quantitative trait loci analysis was performed using QTL IciMapping software. Based on this, one locus on A03 and one on A09 were identified to be associated with resistance to *P. brassicae* pathotype 3A (Table 3.2). These two loci were identified with LOD scores greater than 6.0. The A03 and A09 loci explained about 10% and 33% of the total phenotypic variance and exerted additive effect of -16.9 and -27.3% DSI, respectively; the minus value indicates that the *B. rapa* allele of each of the locus reduced the DSI by 16.9% and 27.3% DSI. The physical positions of these loci were determined based on reference genome of *B. rapa* cv. Chiifu v3.0. The A03 and A09 loci identified through QTL mapping approach could be positioned at 1.65-10.23 Mb and at 38.01-41.5 Mb, respectively (Table 3.2). The SMA and GWAS analysis as well as the physical position of the A03 and A09 markers identified through these approaches further confirmed these two loci as contributing to resistance against pathotype 3A (Table 3.3, 3.4; Figure 3.3). Among the A09 markers, the marker SSRid_A9_330 identified through SMA and GWAS approaches positioned at 20.25 Mb; however, the position of this marker on the linkage map was close to the marker SSRid_A9_770 positioned at 38.01 Mb (Figure 3.3). Therefore, SSRid_A9_330 might be located close to the genomic region of SSRid_A9_770, i.e., in the QTL region identified through construction of linkage map followed by QTL mapping approach.

Table 3.2 Quantitative trait loci (QTL) for resistance to *Plasmodiophora brassicae* pathotype 3A detected using a *Brassica napus* recombinant inbred line (RIL) population (n = 94) derived from (*B. napus* × *B. rapa* turnip) × *B. napus* interspecific cross; analysis carried out using inclusive composite interval mapping (ICIM-Add) method.

Chromo-some	Position cM	Left marker	Right marker	LOD	PVE (%)	Add	Physical position on *B. rapa* Chiifu v3.0
A03	84.9	SSRid_A 3_3204	SSRid_ A3_231	6.1	10.4	-16.9	1.65 – 10.23 Mb

A09	98.5	SSRid_A09_91	SSRid_A09_770	18.5	33.6	-27.3	38.01 – 41.5 Mb

Note. LOD = logarithm of the odd; PVE = phenotypic variation explained by the QTL.
Add = Estimated additive effect of the QTL; negative sign indicates '*B. rapa* turnip' allele reduces disease severity index (DSI), i.e., contribute to resistance.

We also identified the genomic regions on A02 and A08 affecting resistance to pathotype 3A through SMA and GWAS (Table 3.3, 3.4; Figure 3.3), and determined their position on the chromosomes. The physical position of the A02 QTL identified through these two approaches was 0.81 Mb and 0.81-3.63 Mb, respectively; and the position of the A08 QTL identified following these two approaches was 16.21-20.87 Mb and 14.34-20.87 Mb, respectively. However, the A02 and A08 QTL could not be detected through QTL mapping approach.

Table 3.3 Quantitative trait loci (QTL) detected for resistance to *Plasmodiophora brassicae* pathotype 3A by using a *Brassica napus* recombinant inbred line population ($n = 96$) derived from (*B. napus* × *B. rapa* turnip) × *B. napus* interspecific cross following single marker analysis approach.

Chromo-some	Marker name	LOD	R^2 (%)	Add	Physical position (Mb) on *B. rapa* Chiifu v3.0
A02	SSRid_A2_5157	4.3	18.8	-22.3	0.81
A03	SSRid_A3_3109	3.4	15.3	-24.7	1.78
A03	SSRid_A3_3157	6.6	27.5	-27.8	0.98
A08	SSRid_A08_2879	4.8	21.0	-25.6	16.21
A08	SSRid_A08_2330	5.9	25.0	-29.5	20.87
A09	SSRid_A9_771	6.0	25.2	-25.6	37.93
A09	SSRid_A9_91	11.5	42.5	-32.0	41.55
A09	SSRid_A9_786	22.0	65.1	-38.4	36.23
A09	SSRid_A9_1029	8.8	35.0	-28.1	40.94
A09	SSRid_A9_770	7.0	29.0	-26.0	38.01
A09	SSRid_A9_330	5.2	22.3	-22.8	20.25

A09	SSRid_A9_878	6.7	28.0	-30.0	36.80
A09	SSRid_A9_1018	5.3	23.0	-27.2	43.12
A09	SSRid_A9_1021	7.7	31.0	-28.3	42.00
A09	SSRid_A9_1005	5.0	21.0	-21.7	43.86

Note. LOD = logarithm of the odd; R^2 = phenotypic variation explained by the QTL.
Add = Estimated additive effect of the QTL; negative sign indicates '*B. rapa* turnip' allele reduces disease severity index (DSI), i.e., contribute to resistance.

Table 3.4 Quantitative trait loci (QTL) detected for resistance to *Plasmodiophora brassicae* pathotype 3A by using a *Brassica napus* recombinant inbred line population (n = 96) derived from (*B. napus* × *B. rapa* turnip) × *B. napus* interspecific cross following genome-wide association study (GWAS) approach and using software program TASSEL 5.2.84

Chromo-some	Marker	LOD	R^2 (%)	Add	Physical position (Mb) on *B. rapa* cv. Chiifu
A02	SSR_A2_5157	5.61	18.98	-21.6	0.81
A02	SSR_A2_5168	5.52	18.63	-21.1	3.63
A03	SSR_A3_3109	4.55	17.87	-9.9	1.78
A03	SSR_A3_3157	6.92	27.13	-16.9	0.98
A08	SSR_A08_972	5.71	22.56	-13.3	14.34
A08	SSR_A08_2879	6.17	21.15	-24.7	16.21
A08	SSR_A08_2330	7.26	25.23	-31.0	20.87
A09	SSR_A9_866	6.27	24.70	-14.9	38.60
A09	SSR_A9_1006	4.44	17.41	-9.6	43.84
A09	SSR_A9_1005	8.75	33.60	-23.0	43.86
A09	SSR_A9_1021	8.85	33.91	-23.4	42.00
A09	SSR_A9_1014	4.69	18.43	-10.3	44.41
A09	SSR_A9_771	5.98	23.59	-14.0	37.93
A09	SSR_A9_1004	4.42	17.31	-9.5	43.92
A09	SSR_A9_91	14.42	50.15	-45.8	41.55

A09	SSR_A9_786	16.77	55.74	-57.3	36.23
A09	SSR_A9_770	8.89	34.06	-23.5	38.01
A09	SSR_A9_330	5.47	21.59	-12.5	20.25
A09	SSR_A9_1029	5.84	23.04	-13.6	40.94
A09	SSR_A9_802	4.30	16.81	-9.2	31.46
A09	SSR_A9_878	7.19	28.12	-17.8	36.85
A09	SSR_A9_865	4.74	18.66	-10.4	39.99

Note. LOD = logarithm of the odd; R^2 = phenotypic variation explained by the QTL.
Add = Estimated additive effect of the QTL; negative sign indicates '*B. rapa* turnip' allele reduces disease severity index, i.e., contribute to resistance.

3.3.5 Identification of putative candidate genes in the QTL regions

We identified 40 homologous *B. napus* genes to the resistance genes from BRAD. Among these, 15 genes did not have any functional domain, therefore these were removed from further analysis. Of the remaining 25 genes, 9 candidate genes belonged to the QTL region on A02, and 16 candidate genes belonged to the A03 chromosome QTL region. 15/25 genes encoded proteins consisting of TIR/coiled coil/LRR or truncated TIR domain, six genes possessed transmembrane regions displaying transportation roles, remaining four genes were annotated to possess PMEI (pectin methyesterase inhibitor), SANT, F-box and U box domains (Supplementary Table S3.4).

3.4 Discussion

The relationship between the three monogenomic species and three amphidiploid species of the genus Brassica is described in the "Triangle of U" (U 1935) and their cytogenetic relationships are also well documented (Attia and Röbbelen 1986; Chèvre et al. 2018). It is therefore possible for gene transfer to occur between the various Brassica species.

Clubroot resistance (CR) of the primary gene pool of spring *B. napus*, i.e., winter *B. napus* cv. Mendel and rutabaga were used to develop clubroot resistant spring *B. napus* canola (Rahman

et al. 2011b; Hasan and Rahman 2016). The resistance present in winter canola and rutabaga is, in fact, derived from turnips (*B. rapa* ssp. *rapifera*) (Diederichsen and Sacristan 1996; Bradshaw et al. 1997; Frauen 1999). The secondary gene pool i.e., the three diploid species *B. nigra*, *B. oleracea* and *B. rapa* have also been reported to carry CR (Hasan et al. 2012; Peng et al. 2014a). Introgression of resistance from *B. nigra* is, however, challenging because the B genome shows the least homology with the A genome (Busso et al. 1987). Many genetic studies have reported the quantitative nature of CR in *B. oleracea* (Figdore et al. 1993; Voorrips and Kanne 1997; Nagaoka et al. 2010; for review, see Hasan et al. 2021a). On the other hand, CR in *B. rapa* is often controlled by a single dominant gene (Kato et al. 2013; Chu et al. 2014; for review, see Hasan et al. 2021a) which offers a great advantage of using the major CR genes of the A genome instead of the minor genes of the C genome. *B. rapa* has also been reported to carry resistance to a wide range of *P. brassicae* pathotypes (Hasan et al. 2012; Peng et al. 2014a). Among the different types of *B. rapa*, the European fodder turnips have been used extensively in breeding vegetable *B. rapa* (Hirai et al. 2004) as well as *B. napus* (Bradshaw et al. 1997; Yu et al. 2022). The European turnip accession ECD 01, harbors several clubroot resistance loci such as *CRb*, *CRc*, *CRk*, *BraA.CR.a*, *BraA.CR.b*, $Rcr9^{ECD01}$ and $Rcr10^{ECD01}$(Piao et al. 2004; Sakamoto et al. 2008; Hirani et al. 2018; Yu et al. 2022); therefore, the use of this resistance in canola breeding was expected to introduce resistance to the recently evolved virulent pathotypes, such as pathotype 3A. Previously, we reported the prospect of the development of euploid *B. napus* canola lines carrying resistance to multiple *P. brassicae* pathotypes, including pathotype 3A from *B. napus* × *B. rapa* var. *rapifera* (ECD 01) interspecific cross (Kaur et al. 2022). In this study, we used a set of recombinant inbred lines (RILs) derived from this interspecific cross to identify the genomic regions and markers associated

with this resistance for use in molecular breeding, using WGRS followed by QTL mapping approach.

For WGRS, we mapped the sequence reads to two reference genomes to increase the reliability of mapping and identification of markers associated with resistance, and we found similar results demonstrating the reliability of the results. Hasan et al. (2021c) had also used two reference genomes, *B. napus* DH12075 v3.1 and *B. napus* cv. Darmor-*bzh* v4.1, for mapping clubroot resistance using a RIL population developed from *B. napus* × Bilko interspecific cross. Similarly, Wang et al. (2022) also used these two genomes to identify the major loci on A03 and A08 chromosomes conferring resistance to *P. brassicae* pathotypes 3H, 3A and 3D in canola. However, WGRS using single reference genome for mapping of clubroot resistance has also been found to be reliable. For example, Hasan et al. (2021c) used *B. rapa* cv. Chiifu v1.5 and Yu et al. (2022) used *B. napus* cv. Darmor v4.0 for mapping clubroot resistance in populations derived from rutabaga × spring canola and *B. rapa* cv. ECD 01 × canola line DH16516 crosses, respectively.

The plant materials that we used in this study carry resistance to different *P. brassicae* pathotypes, including pathotype 3H and 3A. Among the different *P. brassicae* pathotypes, the pathotype 3H is still predominant in Canada, and several researchers have introgressed resistance to this pathotype and mapped the resistance (e.g., Rahman et al. 2011b; Hasan and Rahman 2016; Fredua-Agyeman and Rahman 2016; Hasan et al. 2021a, 2021b, 2021c). Among the recently evolved virulent pathotypes, pathotype 3A is important as it is currently one of the most prevalent pathotypes in Canada. It overcame clubroot resistance of the available resistant canola cultivars (Hollman et al. 2021). Therefore, we focused on the mapping of these two resistances for use in molecular breeding. The WGRS analysis identified the 16-30 Mb region of A03 and 2-14 MB region of A08 associated with resistance to pathotype 3H (Figure 3.1c and 3.1d), providing support

to previous studies (for review, see Hasan et al. 2021a); therefore, we carried out QTL mapping of resistance to pathotype 3A only. In this case, we identified two major loci, one on chromosome A03 at about 0.98-2.0 Mb and the other on A09 at about 38.01-41.5 Mb of *B. rapa* cv. Chiifu v3.0. Among these, the effect of the A09 locus was 61.5% greater than the effect of the A03 locus.

Previously, QTL or major loci conferring resistance to *P. brassicae* have been reported at 18.50-26.35 Mb region of A02, 15.5-16.3 Mb and 23.8-26.8 Mb region of A03, and 11.34-21.75 Mb region of A08 (for review, see Hasan et al. 2021a; Yu et al. 2017). Werner et al. (2008) carried out QTL mapping of resistance of a *B. napus* population carrying clubroot resistance of *B. rapa* ECD 04 using different *P. brassicae* isolates. In addition to the two major loci on A03 and A08, they also reported a minor QTL on A09 exerting an additive effect of only 4.9% DSI; physical position of this QTL could not be found in this publication. Thus, to our knowledge, the A09 locus conferring resistance to *P. brassicae* pathotype 3A identified in this study can be considered to be a novel locus. Of the 47 resistant lines in the mapping population, most of the lines also carry resistance to pathotype 3H; therefore, it is probable that the A09 locus that we introgressed into canola also plays a role in resistance to pathotype 3H. The A03 locus that we detected in this study is located at about 1-2 Mb region (Figure 3.3; Table 3.3, 3.4). Previously, researchers reported clubroot resistance loci from 15.5-16.3 Mb (Sakamoto et al. 2008; Pang et al. 2018) and 23.8-26.8 Mb region (Yu et al. 2017) of A03. Thus, it is possible that we detected a new locus on chromosome A03.

The A02 locus detected in this study is located at about 1-4 Mb position of the *B. rapa* cv. Chiifu v3.0 reference genome, which could be different from the locus at 18.50-22.10 Mb region reported by Yu et al. (2017). Sakamoto et al. (2008) and Werner et al. (2008) also reported clubroot resistance loci on A02; however, physical position of these loci cannot be found in literature. On

the other hand, the A08 locus that we detected in this study (14-21 Mb) could be the same as the one reported previously by different researchers (Sakamoto et al. 2008; Pang et al. 2018; Zhu et al. 2019b).

3.5 Conclusion

In this study, by using a *B. napus* canola population carrying clubroot resistance from *B. rapa* var. *rapifera* cv. Debra and following WGRS and QTL mapping approaches, we identified the genomic regions of A02, A03, A08 and A09 chromosomes contributing resistance to *P. brassicae* pathotypes 3H and 3A. Among these loci, A02, A03 and A09 appeared to be novel in canola, while the A08 locus corresponded to the previously reported one on this chromosome. The A09 locus explained about 50% greater phenotypic variance as compared to the other loci, and this locus could be detected following all three QTL mapping methods employed in this study. This study also demonstrates the prospects of introgression of multiple clubroot resistance loci from *B. rapa* into *B. napus* canola. A similar approach could be employed on other clubroot resistant *B. rapa* accessions for introgression and pyramiding of multiple resistance genes into canola.

Chapter 4. Comparative transcriptome analysis of *Brassica napus* near-isogenic lines carrying resistance of turnip in response to *P. brassicae*

4.1 Introduction

Plasmodiophora brassicae, a soil-borne obligate protist, is the casual organism of clubroot disease in Brassica. Its life cycle can be divided into primary and secondary infection phases. During the primary infection phase, the resting spores of this pathogen germinate under favourable environmental conditions resulting in motile primary zoospores, which infect the root epidermal cells of the host plant, produce primary plasmodium and eventually zoosporangia. During secondary infection phase, the secondary zoospores, produced in the zoosporangia, infect the cortical cells of the host roots, form multinucleate secondary plasmodium, and ultimately result in the formation of galls (Kageyama and Asano 2009; Liu et al. 2020). The galls in the root hinder the uptake of water and nutrients by the plants resulting in stunted growth and wilting of the plants (Strelkov and Dixon 2014).

Among the different abiotic and biotic stresses affecting canola yield, clubroot disease is one of the most important one, causing about 30% yield loss in canola (Tewari et al. 2005). Management of the clubroot disease is critical for successful crop production. Clubroot disease management strategies in Canada include appropriate crop rotation, sanitizing field equipment, application of soil amendments, and most importantly, the use of clubroot resistant cultivars (for review, see Peng et al. 2014a; for review, see Hasan et al. 2021a). However, the resistance in a cultivar may become ineffective owing to selection pressure on the pathogen populations over a period of time as well as due to the emergence of new pathotypes (Strelkov et al. 2016b, 2018). Therefore, it is important to understand the host-pathogen interactions at the molecular level to gain a better insight into the resistance mechanisms in order to be able to further manipulate the host genes that may be involved in mediating resistance to *P. brassicae* infection.

During pathogen invasion, plants initially defend against the pathogen by using physical barriers, including cell wall, cuticle, waxy layer and xylogen as well as chemical barriers such as phenols, saponins and glucosinolates (for review, see Jones and Dangl 2006). Once the pathogen overcomes these barriers, plants activate a two-tiered immune response known as pathogen-associated molecular pattern (PAMP)-triggered immunity (PTI) and effector-triggered immunity (ETI) (for review, see Jones and Dangl 2006). PTI is activated in the host by pattern recognition receptors (PRRs) such as receptor-like kinases (RLKs) and receptor-like proteins (RLPs) located in the plasma membrane which provides a broad-spectrum resistance against the pathogen (for review, see Boller and Felix 2009). However, the pathogens can overcome this basal level of immunity by secreting various effectors (also called avirulence proteins). The effector molecules are recognized by nucleotide-binding-leucine-rich repeat (NLR) proteins encoded by the resistance (R) genes, and this results in the activation of ETI in the host (for review, see Bent and Mackey 2007). ETI prevents the progression of the pathogen infection by inducing various pathogenesis-related genes such as reactive oxygen species (ROS), calcium and hormone signalling, transcription factors (TFs), cell wall modifications and hypersensitive cell death type resistance against the pathogen (for review, see Ting et al. 2008; Yuan et al. 2021a).

To date, many clubroot resistance (CR) loci/quantitative trait loci (QTL) have been mapped on the A genome chromosomes of *B. rapa*; maximum number being on chromosome A03 followed by A08 (for review, see Hasan et al. 2021a). These include 11 loci namely *Crr2* (Suwabe et al. 2003, 2006), *CRd* (Pang et al. 2018), *Crr3* (Hirai et al. 2004; Saito et al. 2006), *Rcr5* (for review, see Huang et al. 2019), *Rcr1* (Chu et al. 2014), *CRa/CRb* (Piao et al. 2004; Cho et al. 2008; Kato et al. 2012, 2013; Ueno et al. 2012), *Rcr2* (Huang et al. 2017), *CRs* (Laila et al. 2019), *Rcr3* (Karim et al. 2020), *Crr1a* (Suwabe et al. 2003, 2006; Hatakeyama et al. 2013) and *Crr1b* (Hatakeyama

et al. 2013). In addition, many QTL, including *CR6a* and *CR6b* (Lee et al. 2002), *PbBa1.1* (Chen et al. 2013), *CRc* and *CRk* (Sakamoto et al. 2008), *Rcr4* and *Rcr8* (Yu et al. 2017), *CRq* (Yuan et al. 2015), *Bcr1* and *Bcr2* (Zhang et al. 2022), $Rcr9^{ECD01}$ and $Rcr10^{ECD01}$ (Yu et al. 2022), *CrrA5* (Nguyen et al. 2018), *Crr4* (Suwabe et al. 2006), *qBrCR38-1* and *qBrCR38-2* (Zhu et al. 2019b) and *Rcr9* (Yu et al. 2017; Karim et al. 2020) on A01, A02, A03, A05, A06, A07 and A08 chromosomes (for review see, Hasan et al. 2021a) have been identified. These loci/QTL, either individually or in combination, provide race or pathotype specific resistance against clubroot.

In recent years, RNA sequencing (RNA-seq) technique has been employed to study the molecular mechanisms of clubroot resistance in *Brassica* (Chu et al. 2014; Chen et al. 2016; Zhang et al. 2016; Jia et al. 2017; Ciaghi et al. 2019; Fu et al. 2019; Mei et al. 2019; Wang et al. 2019; Summanwar et al. 2020; for review, see Zhou et al. 2020; Wei et al. 2021; Yuan et al. 2021b; Adhikary et al. 2022b). One of the early RNA-seq study (Chu et al. 2014) using clubroot susceptible and resistant *B. rapa* ssp. *chinensis* carrying the locus *Rcr1* showed a significant upregulation of the genes involved in phytohormone [jasmonic acid (JA), ethylene (ET) and auxin] biosynthesis in the resistant plants. Wei et al. (2021) found upregulation of the genes from pathways related to plant-pathogen interaction and plant hormone signal transduction as well as 15 hub genes, including pathogenic type III effector avirulence factor gene (*RIN4*) and auxin-responsive protein (*IAA16*) in clubroot resistant *B. rapa*. Some other studies (Chen et al. 2016; Jia et al. 2017; Yuan et al. 2021b) with *B. rapa* ssp. *pekinensis* reported expression change for the genes carrying TIR-NB-LRR and NBS-LRR domains, as well as the genes involved in pathogenesis-related proteins, cell wall thickening, calcium (Ca^{2+}) influx, glucosinolates biosynthesis and chitin metabolism in the resistant plants. The same group of genes in addition to hormone signaling and generation of ROS has also been reported to be upregulated in clubroot

resistant *B. oleracea* (Zhang et al. 2016; Wang et al. 2019; Fu et al. 2019). The transcriptome profile of the symptomless roots of Kohlrabi (*B. oleracea*) exhibited an increased expression of genes related to cell wall synthesis, cytokinin metabolism and salicyclic acid (SA) defense responses (Ciaghi et al. 2019). Chen et al. (2016) found that among the differentially expressed genes (DEGs) between clubroot resistant and susceptible *B. napus*, genes related to disease resistance (broad-spectrum and clubroot-specific) were upregulated and genes related to cytokinin and myrokinase synthesis were downregulated. Increased production of ROS through the fast initiation of receptor kinases, activation of host intercellular G proteins (guanine nucleotide-binding proteins) and enhanced Ca^{2+} signaling leads to lesser gall formation in resistant *B. napus* (Mei et al. 2019). Similar results have also been presented in a comparative transcriptomic study using resistant and susceptible rutabaga (for review, see Zhou et al. 2020). Galindo-Gonzalez et al. (2020) reported the involvement of the genes from SA, auxin and cytokinin pathway in clubroot resistance in *B. napus*. The suppression of cytokinin and enhancement of SA limits gall formation in clubroot resistant *B. napus* plants, whereas an increased JA content promotes gall formation in the susceptible plants (Prerostova et al. 2018). Using the strand-specific long non-coding RNA-seq (lncRNA-seq) approach, Summanwar et al. (2019, 2020) identified differentially expressed lncRNAs belonging to the plant-pathogen interaction and hormone signaling pathways in canola lines carrying resistance located on chromosome A08 or A03. A comparative transcriptomics study using the above-mentioned canola lines showed the upregulation of genes involved in primary and secondary metabolism (Summanwar et al. 2021). Adhikary et al. (2022b) identified DEGs corresponding to differentially accumulated metabolites, including pyruvate kinase and citrate synthase at early pathogenesis stage in the inoculated canola lines.

As discussed above, based on the results obtained from several RNA-seq studies using disease resistant and susceptible cultivars or lines from a breeding population, the DEGs identified may not reflect only the genes involved in disease resistance but also many other genes involved in the control of other traits. This constraint can be avoided largely by using near-isogenic resistant (NILs) and susceptible lines. By using the resistant NILs for bacterial leaf pustule in soybean and transcriptome profiling, Kim et al. (2011) identified the PAMP and DAMP (damage associated molecular pattern) receptors and the genes induced by these receptors. Chen et al. (2016) carried out transcriptome analysis of Chinese cabbage roots of clubroot resistant and susceptible NILs for the *CRb* locus of A03 chromosome and identified the effector receptors (resistance proteins) and pathogenesis-related genes from SA signaling pathway to be involved in clubroot resistance. By using clubroot resistant and susceptible NILs of *B. napus* for *CRb* locus of A03, Li et al. (2021) identified 18 miRNAs in roots of resistant NILs to be involved in root development, hypersensitive cell death and chloroplast metabolite synthesis. However, very limited studies have been carried out on clubroot resistance in Brassica using NILs to understand the molecular basis of this trait.

Previously, we introgressed clubroot resistance from *B. rapa* var. *rapifera* (turnip) into *B. napus* canola (Kaur et al. 2022) and developed NILs based on this resistance. These NILs have been used in this study to understand the molecular basis of the 'Turnip-resistance' in *B. napus* canola.

4.2 Materials and Methods

4.2.1 Plant materials

In this study, BC_4F_3 generation homozygous resistant and susceptible NILs were used. The NILs were developed by crossing a clubroot resistant *B. napus* line, carrying the resistance from *B. rapa* turnip accession ECD 01 (cv. Debra) (Kaur et al. 2022), with the clubroot susceptible *B. napus*

canola line A04-73NA followed by backcrossing of the hybrids for four times (BC_4) using A04-73NA as the recurrent parent. The clubroot resistant BC_4 plants were self-pollinated for two generations. During the development of the NILs, selection for resistance to pathotype 3H was carried out in each generation. Following this, the three homozygous resistant and three susceptible BC_4F_3 NILs were developed, and were designated as CR-NILs and CS-NILs, respectively.

4.2.2 Inoculation and root sample collection

The NILs were grown in a greenhouse at the University of Alberta in 32-cell (7 cm × 7 cm × 9 cm; L × W × D) trays filled with Sunshine Professional Growing Mix (Sunshine Horticulture, 15831 N.E., WA, USA). The greenhouse conditions were 20-22/15 °C day/night temperature and 16 hours (h) photoperiod with light intensity of 450 μE (mV) m^2 s^{-1} at plant level. The CR- and CS-NILs were inoculated with *P. brassicae* pathotype 3H at 7-10 days after germination following the pipette method; the details of inoculum preparation and inoculation technique can be found in Kaur et al. (2022). The control plants were left un-inoculated.

According to histopathological analysis by Summanwar et al. (2020), the symptoms of primary infection on secondary roots began at 7 days after inoculation (dai) and a greater number of secondary plasmodia were found in the susceptible roots at 14 dai. Based on these observations, the root samples collected at these two time points were used for transcriptome profiling. The plants were removed from the soil at 7- and 14 dai, and their roots were washed with tap water to remove the soil debris, dried with paper towels, collected in falcon tubes (Thermo Fisher Scientific, MA, USA), immediately frozen in liquid nitrogen, and were stored at -80 °C until use. The experimental materials included a total of 24 root samples [2 genotypes (CR-NILs and CS-NILs) × 2 time points (7 dai and 14 dai) × 3 biological replicates × 2 treatments (control and inoculated)].

Each root sample included a pool of three NILs and there were seven plants of each NIL. Thus, the total number of plants used for this experiment was $2 \times 2 \times 3 \times 2 \times 3 \times 7 = 504$ plants.

4.2.3 RNA extraction

The root samples stored at -80°C were homogenized into a fine powder in liquid N_2 using a pestle and a mortar. About 200-250 mg ground root sample was aliquoted in 2 ml Eppendorf tube (Thermo Fisher Scientific, MA, USA) and the RNA was extracted using Trizol reagent (Invitrogen, MA, USA) following the manufacturer's protocol. The RNA was treated with DNase I (Promega, WI, USA) to remove contaminating DNA. The quality and concentration of the RNA was determined using a NanoDrop ND-1000 spectrophotometer (Thermo Scientific, MA, USA). The RNA samples with 260/280 nm ratio of 2.0-2.1 and RNA Integrity Number (RIN) ≥ 8 were used for sequencing.

4.2.4 RNA sequencing and data analysis

For library preparation, the messenger RNA was purified from total RNA using poly-T oligo-attached magnetic beads. After fragmentation, the first strand cDNA was synthesized using random hexamer primers, followed by the second strand cDNA synthesis using dUTP for directional library. Library preparation and sequencing of the 24 RNA samples was carried out by Novogene, U.S.A. (https://en.novogene.com/) using Illumina HiSeq 2500 platform, and 2 × 150-bp paired-end reads were generated. The raw sequence data in Fastq format was processed by Novogene using their in-house perl scripts to obtain clean reads after removal of the adapter or poly-N sequences and low-quality reads. The Q20, Q30 and GC content was determined for the clean reads. The paired-end clean reads were mapped to the *B. napus* cultivar Darmor-bzh v4.1 reference genome (Chalhoub et al. 2014) using Hisat2 v2.0.5 (Kim et al. 2015). The mapped reads were assembled by using Stringtie v1.3.3b (Pertea et al. 2015) in a reference-based approach.

Quantification of the level of gene expression for each sample was carried out based on Fragments Per Kilobase of transcript sequence per Million base pairs sequenced (FPKM). The FPKM value for each gene was calculated based on the length of the fragments and count of the mapped reads using FeatureCounts v1.5.0-p3 (Liao et al. 2014).

4.2.5 Differentially expressed genes (DEGs) and their enrichment analysis

The DEGs between the inoculated (I) and uninoculated or control (C) samples of the resistant (R) and susceptible (S) NILs at the two time points were detected using DESEQ2R package (v1.20.0) (Love et al. 2014) from the following comparisons: RI vs. RC, SI vs. SC, and RI vs. SI. The *p*-values of the DEGs were adjusted using the Benjamini and Hochberg's approach to control the false discovery rate (FDR). The DEGs with $\log_2$FC ($\log_2$-fold change) ≥ 1 and adjusted *p*-value ≤ 0.05 were considered as significant DEGs.

The DEGs were used for Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis to identify the key genes and pathways potentially involved in clubroot resistance. The GO analysis was carried out using the clusterProfiler R package which corrects for any bias in gene length (Yu et al. 2012). The GO terms with corrected *p*-value < 0.05 were considered to be significantly enriched by the DEGs. The KEGG pathway analysis of the DEGs was also carried out using the clusterProfiler R package (Kanehisa and Goto 2000).

4.2.6 Analysis of the RNA-seq data to identify the DEGs responsible for susceptibility to clubroot, PTI and ETI

Brassica-P. brassicae host-pathogen interaction involves the expression of various types of genes, such as the genes induced in response to primary and secondary infections, the genes involved in susceptibility, as well as the genes involved in different stresses during disease development. We

carried out the experiment using CR-NILs and CS-NILs with or without inoculation, and this included the following treatments: resistant inoculated (RI), resistant non-inoculated (control) (RC), susceptible inoculated (SI), and susceptible non-inoculated (control) (SC). By comparing RNA-seq data from RI vs. RC, SI vs. SC, and RI vs. SI comparisons, we identified several DEGs to be involved in disease symptom development, general stress response, clubroot resistance or susceptibility. We further analysed these DEGs using a subtraction approach, as summarized in the following table, to identify the DEGs involved in PTI, ETI and susceptibility (Table 4.1).

Table 4.1 The subtraction approach applied on the differentially expressed genes (DEGs) obtained from different comparisons, including RI vs. RC, SI vs. SC, and RI vs. SI, in *Brassica napus* near isogenic lines (NILs) inoculated or uninoculated with *Plasmodiophora brassicae* pathotype 3H. RI = inoculated resistant NILs; RC = uninoculated (control) resistant NILs; SI = inoculated susceptible NILs; SC = uninoculated (control) susceptible NILs.

Comparisons	DEGs expected to be involved in	Remarks
RI vs. RC	PTI + Stress + ETI (=X)	DEGs involved in PTI and ETI, and the DEGs related to stress induced due to pathogen attack.
SI vs. SC	PTI + Stress + Disease symptom development+ S (=Y)	DEGs involved in PTI, related to stress induced due to pathogen attack, and the DEGs that play a role in disease symptom development and susceptibility (S).
RI vs. SI	PTI + Stress + Disease symptom development + S + ETI (=Z)	
Common DEGs of X and Y	PTI and Stress (=A)	
X - A	ETI (=B)	
Y - A	Disease symptom development + S (=C)	
Z - C	PTI + Stress + ETI (=D)	
Z - B	PTI + Stress + Disease symptom development + S (=E)	
D - B	PTI + Stress (=F)	
E - C	PTI + Stress (=G)	
D - F	ETI (=H)	
Z - D	Disease symptom development + S (=I)	

Z - X	Disease symptom development + S (=J)	
Z - Y	ETI (=K)	
Common DEGs of B, H and K	ETI	DEGs involved in ETI
Common DEGs of A, F and G	PTI + Stress	DEGs involved in PTI + Stress
Common DEGs of C, I and J	Disease symptom development + S	DEGs involved in disease symptom development + S

Note: PTI = pathogen-associated molecular pattern (PAMP)-triggered immunity; ETI = effector-triggered immunity.

4.2.7 Validation of RNA-seq data by qRT-PCR

Validation of the RNA-seq data was carried out through quantitative real time-polymerase chain reaction (qRT-PCR) analysis for 18 DEGs. The cDNAs of the DEGs was synthesized from 2 µg of total RNA using Thermo Scientific's instructions of RevertAid H Minus Reverse Transcriptase (Thermo Fisher Scientific, MA, USA). Gene specific primers were designed using Primer Express v3.0.1 (Applied Biosystems-Life Technologies, ON, Canada). Ubiquitin-Conjugating Enzyme 9 (*UBC9*), a house-keeping gene, was used as an endogenous control. qRT-PCR was performed using ViiA 7 Real-Time PCR system (Applied Biosystems-Life Technologies, MA, USA). Each PCR reaction included 5 µl of FASTSYBR Green Mix (Applied Biosystems-Life Technologies, ON, Canada), 1 µl of cDNA, 2 µl of primer pair, and 2 µl of nuclease free water to make the final reaction volume up to 10 µl. Three biological replicates for each sample and two technical replicates of each biological replicate were analysed. Changes in the relative gene expression were calculated using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001). Statistical analysis of data was carried out using two-way ANOVA.

4.3 Results

4.3.1 Phenotypic evaluation of P. brassicae infection in CR- and CS-NILs

Clubroot disease assessment post-inoculation with *P. brassicae* was based on the development of galls on the roots. No visible disease symptom could be seen in both CR-NILs and CS-NILs at 7 dai. Noticeable galls were seen in the inoculated CS-NILs at 14 dai, however, no symptoms could be seen in the inoculated CR-NILs (hereafter referred as RI) (Supplementary Figure S4.1). The aerial parts of the inoculated CS-NILs also showed visible symptoms such as yellowing of leaves and stunted growth at 14 dai (Supplementary Figure S4.1).

4.3.2 Overview of RNA sequencing data

Differences in gene expression in the CR-NIL and CS-NIL root tissues in response to *P. brassicae* infection at 7- and 14 dai were determined by comparing RNA-seq data from infected samples with their corresponding uninoculated controls. A total of 1.01 billion clean reads were generated by 150 bp paired-end sequencing after removal of the adaptor and low-quality reads (Supplementary Table S4.1). Approximately 88% of the reads were aligned to the *B. napus* reference genome *Darmor-bzh* v4.1 and the majority of the mapped reads (84.03%) aligned to only one position suggesting the uniqueness of the reads. The GC content of the reads was >46% and Q30 values were >94% demonstrating the accuracy and high quality of the sequencing data for downstream analysis.

The Principal Component Analysis (PCA) based on the transcriptome FPKM values showed consistency among the three biological replicates for each sample. The PCA also showed good separation between the control and inoculated samples at both 7- and 14 dai (Supplementary Figure S4.2). In total, 106,682 transcripts were annotated and identified across all the samples and time points. Among these, 20,390 genes were differentially expressed in response to pathogen infection as well as due to genotypic difference. To identify the genes in the CR- and CS-NILs responsive to *P. brassicae* infection, the following comparisons of the RNA-Seq data at 7- and 14

dai were made: CR-NILs inoculated vs. CR-NILs control (RI vs. RC) and CS-NILs inoculated vs. CS-NILs control (SI vs SC). In addition, comparison of inoculated CR-NILs vs. inoculated CS-NILs (RI vs. SI) were made at 7- and 14 dai to identify the genes involved in clubroot resistance. The distribution of DEGs from the above-mentioned comparisons at two time points was visualized using volcano plots (Supplementary Figure S4.3).

For the RI vs. RC data analysis, 519 DEGs (334 upregulated, 185 downregulated) were identified in the CR-NILs at 7 dai, whereas about 16-fold greater number (8,485; 4292 upregulated, 4193 downregulated) of DEGs were identified at 14 dai (Figure 4.1a, 4.1b, 4.1c). Among these 207 DEGs could be detected at both time points (Supplementary Figure S4.4). Based on SI vs. SC data, 424 (293 upregulated, 131 downregulated) and 5,830 DEGs (3543 upregulated, 2287 downregulated) were identified in the CS-NILs at 7- and 14 dai, respectively (Figure 4.1a, 4.1b, 4.1c), where 247 DEGs were common to both time points. Thus, a greater number of genes showed differential expression in the CR-NILs as compared to the CS-NILs. When the inoculated samples of the CR- and CS-NILs (RI vs SI) were compared, 302 DEGs (60 upregulated, 242 downregulated) at 7 dai and 4,830 DEGs (2106 upregulated, 2724 downregulated) at 14 dai were identified (Figure 4.1a, 4.1b, 4.1c). Based on the above-mentioned three comparisons, 7 and 13 genes were found uniquely expressed, respectively, in the CR- and CS-NILs at 7 dai, while 2,977 and 1,465 genes were uniquely expressed, respectively, in the CR- and CS-NILs at 14 dai (Figure 4.1a, 4.1b). Among the 207 common DEGs at 7- and 14 dai from the RI vs RC comparison, 110 DEGs were upregulated and 70 were downregulated at both the time points (Figure 4.2a). Similarly in the SI vs SC comparison, 179 and 40 DEGs out of the total 247 common DEGs were upregulated and downregulated, respectively, at both the time points (Figure 4.2b). The RI vs SI comparison

detected 154 DEGs common at 7- and 14 dai; among them 33 were upregulated and 105 downregulated at both the time points (Figure 4.2c).

We also analysed the SNPs obtained from the resistant inoculated (RI) and susceptible inoculated (SI) samples for their distribution on the 19 *B. napus* chromosomes to identify the genomic regions to be associated with resistance. The distribution of the SNPs from both RI and SI samples was similar (Supplementary Figure S4.5); therefore, no valid conclusion can be drawn based on this SNP data set.

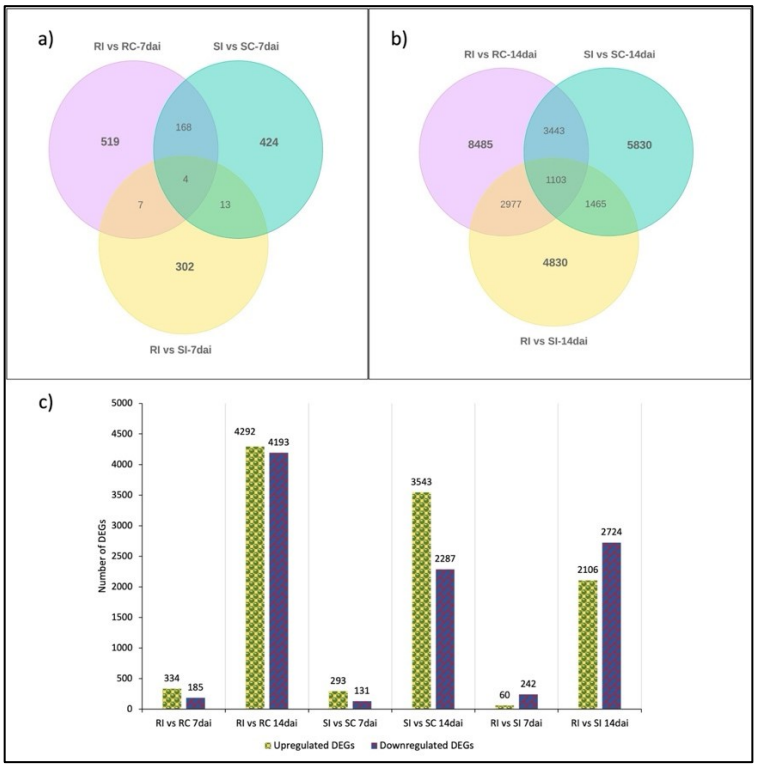


Figure 4.1 Venn diagrams showing differentially expressed genes (DEGs) (*p*-value ≥0.05) in *Brassica napus* near isogenic lines (NILs) due to infection by *Plasmodiophora brassicae* pathotype 3H at **(a)** 7 days after inoculation (dai) and **(b)** 14 dai from three comparisons of the resistant (R) and susceptible (S) near-isogenic lines (NILs) with or without inoculation. The overlapping regions shows the number of DEGs detected from more than one comparison. **(c)** Histogram of the number of upregulated and downregulated DEGs detected at 7- and 14 dai from three comparisons: Resistant inoculated (RI) NILs against its control (RC), susceptible inoculated (SI) NILs against its control (SC), and resistant inoculated (RI) NILs against susceptible inoculated (SI) NILs.

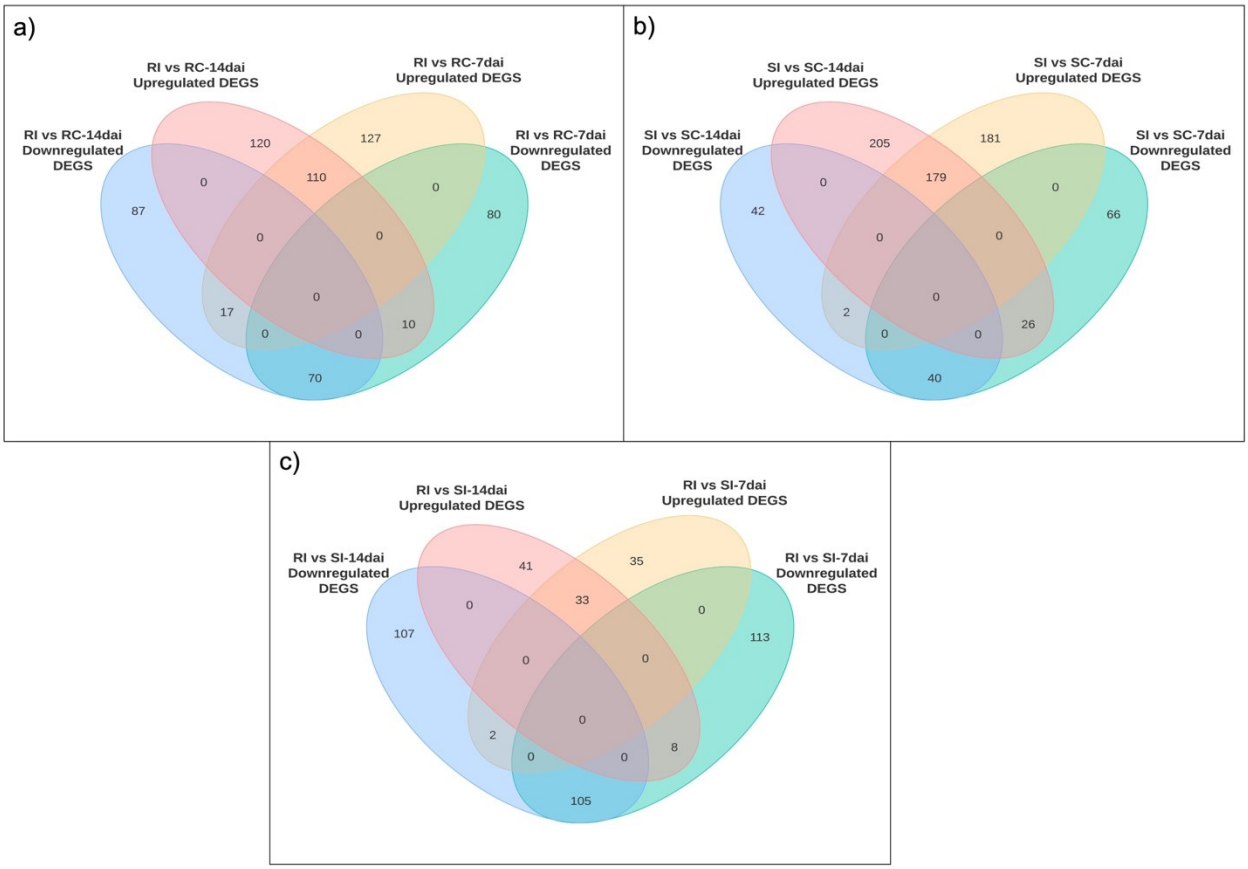


Figure 4.2 Venn diagrams representing the number of upregulated and downregulated differentially expressed genes (DEGs) in *Brassica napus* near isogenic lines (NILs) due to infection by *Plasmodiophora brassicae* pathotype 3H at 7 days after inoculation (dai) and 14 dai from three comparisons of the NILs: **a)** resistant inoculated (RI) NILs against its control (RC), **b)** susceptible inoculated (SI) NILs against its control (SC), and **c)** resistant inoculated (RI) NILs against susceptible inoculated (SI) NILs.

4.3.3 GO and KEGG functional enrichment analysis of the DEGs

To gain a better understanding of the biological mechanisms responsive to *P. brassicae* infection, GO enrichment analysis was performed on the DEGs detected based on RI vs. RC, SI vs. SC and RI vs. SI comparisons. For RI vs. RC comparison, 2,564 DEGs (p value <0.05) were assigned to 108 GO terms belonging to two categories: biological process (BP, 43 terms) and molecular function (MF, 65 terms) (Supplementary Table S4.2a), while 1,576 DEGs (p value <0.05) from SI vs. SC comparison were assigned to 72 GO terms belonging to three categories: BP (33 terms),

cell components (CC, 3 terms) and MF (36 terms) (Supplementary Table S4.2b). A total of 1073 DEGs from RI vs. SI comparison were assigned to 53 GO terms belonging to BP (33 terms), CC (3 terms) and MF (17 terms) (Supplementary Table S4.2c). The most significant 30 GO terms (10 from each category) are shown in Supplementary Figure S4.6.

Between the RI vs. RC and SI vs. SC comparisons, 57 GO terms common between these two comparisons were observed; among which three GO terms belonging to MF (antiporter activity GO:0015297; secondary active transmembrane transporter activity GO:0015291, and phosphoenolpyruvate carboxykinase activity GO:0004611) included greater number of downregulated DEGs in RI vs. RC but upregulated in SI vs. SC (Supplementary Table S4.2d). With respect to BP, the three most significant GO terms were drug membrane transport (GO:0006855), drug transport (GO:0015893) and response to drug (GO:0042493). The three most significant GO terms in MF category, based on *p*-adjusted value, were calmodulin binding (GO:0005516), drug transmembrane transporter activity (GO:0015238) and oxidoreductase activity (GO:0016709).

While analysing data of the RI vs. SI, 51 GO terms were specific to the CR-NIL (RI vs. SI). Based on the total number of DEGs, the top five GO terms of BP category were microtubule-based process (GO:0007017), response to oxidative stress (GO:0006979), movement of cell or subcellular component (GO:0006928), microtubule-based movement (GO:0007018) and defense response (GO:0006952); and the top five of MF category were transferase activity (GO:0016747), antioxidant activity (GO:0016209), peroxidase activity (GO:0004601), oxidoreductase activity (GO:0016684), and cytoskeletal protein binding (GO:0008092) (Supplementary Table S4.2e). On the other hand, 15 GO terms were specific to CS-NIL. Based on the *p*-adjusted value and total number of DEGS, there were three differentially expressed GO terms of BP category involved in

response to external biotic stimulus (GO:0043207), response to other organism (GO:0051707), and defense response to other organism (GO:0098542); two terms of CC category involved in cell wall (GO:0005618) and external encapsulating structure (GO:0030312), and four of the MF category involved in endopeptidase inhibitor and regulator activity (GO:0004866, GO:0061135) and peptidase inhibitor and regulator activity (GO:0030414, GO:0061135) (Supplementary Table S4.2f).

To get a deeper insight into the biological significance of gene function, the DEGs were analyzed by KEGG. A total of 96 pathways were enriched at 7 dai, while 306 pathways were enriched at 14 dai based on RI vs. RC, SI vs. SC, and RI vs. SI comparisons. The top 20 enriched pathways for the DEGs from the three comparisons at two time points are presented in Figure 4.3. Among them, "ath00940: phenylpropanoid biosynthesis", "ath00500: starch and sucrose metabolism", "ath04075: plant hormone signal transduction", "ath01200: carbon metabolism" and "ath04626: plant-pathogen interaction" were the top common categories in response to *P. brassicae* infection. Pathways related to pentose phosphate pathway, glycolysis, and galactose metabolism could be enriched at 7 dai and pathways related to valine, leucine, and isoleucine degradation (ath00280) could be enriched at 14 dai which were exclusively enriched in the resistant NILs in comparison to infected susceptible NILs. It was evident that the number of DEGs increased with time for "ath04075: plant hormone signal transduction" and "ath04626: plant-pathogen interaction" which aid in providing resistance against clubroot.

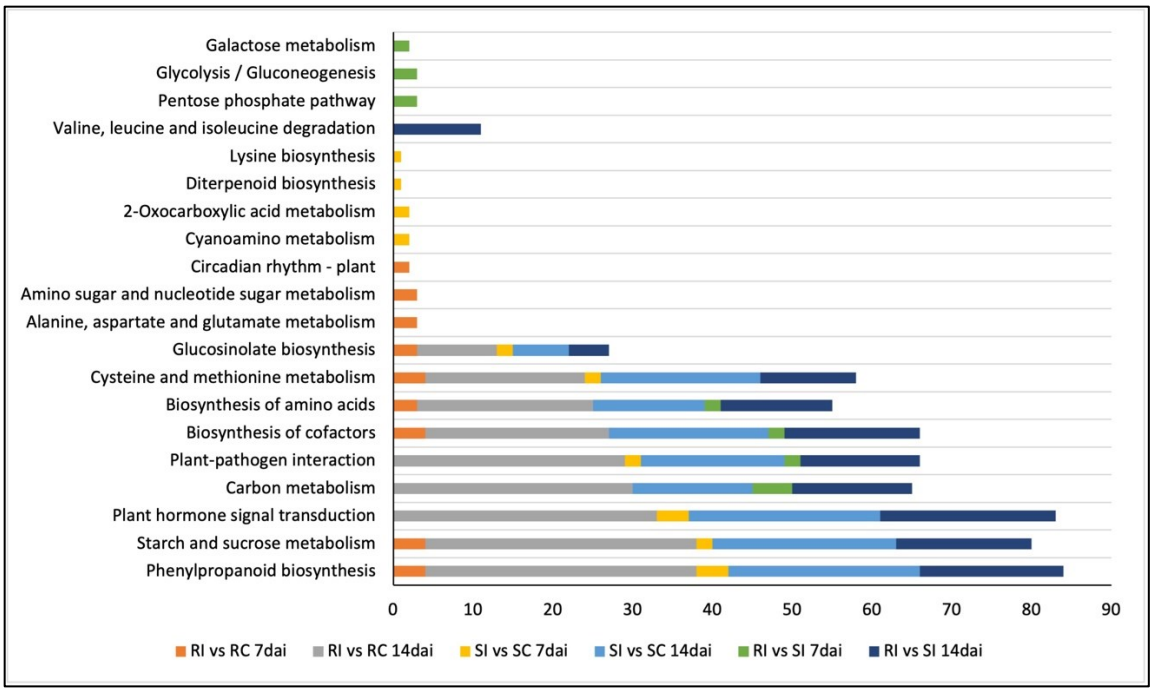


Figure 4.3 The top 20 enriched KEGG pathways of the differentially expressed genes (DEGs) in roots of *Brassica napus* near isogenic lines (NILs) in response to *Plasmodiophora brassicae* pathotype 3H infection at 7- and 14 days after inoculation (dai) from three comparisons of the resistant (R) and susceptible (S) NILs with or without inoculation. The y-axis denotes the enriched KEGG pathways, and the x-axis denotes the gene count. RC: resistant control; RI: resistant inoculated; SC: susceptible control; SI: susceptible inoculated.

4.3.4 Identification of DEGs based on comparison between inoculated and control (RI vs RC and SI vs SC) samples of the CR-NILs and CS-NILs

The pathogen and host plant exhibit a complex interaction where plants employ several defense mechanisms (PTI and ETI response) to combat the pathogen infection and its progression. While comparing the inoculated and control data (RI vs RC and SI vs SC), we found 50 common DEGs from the two comparisons at both time points. These DEGs were related to pathogenesis and stress response (12), transcription factors (10), amino acid biosynthesis and metabolism (seven), cell wall modifications (five) and phytohormone signaling (four), as well as sugar (two), glucosinolates (one) and secondary metabolite synthesis (one) (Figure 4.4, Supplementary Table S4.3).

Pathogenesis and stress response

Among the DEGs related to pathogenesis and stress response, the expression of seven DEGs was upregulated and four downregulated in both CR- and CS-NILs at both time points (Figure 4.4; Supplementary Table S4.3), indicating their possible role in resistance at early stages of infection and disease progression. The upregulated DEGs were homologous to *Chitinase class I, Chitin recognition protein, Hypersensitive-induced response protein 3, AtBBE-like 4* and others, whereas the downregulated DEGs belonged to *Thi4 family, SBP2* and *Cytochrome P450*. One putative gene BnaC04g28450D related to *AtHsfA7a* showed a higher level of expression at 14 dai as compared to 7 dai, in both the resistant and susceptible NILs.

Phytohormone mediated signalling

Phytohormones play an important role in the adaptation of plants to various biotic and abiotic stresses, including clubroot disease (Verma et al. 2016; for review, see Li et al. 2020). The DEGs related to SA (BnaC07g23070D), auxin (BnaC08g18680D) and JA biosynthesis (BnaA03g52720D and BnaC07g44430D) had a consistent expression pattern (upregulated) in the inoculated CR and CS NILs at 7- and 14 dai. However, the two DEGs involved in JA had a higher level of expression in SI over time as compared to RI (Figure 4.4; Supplementary Table S4.3).

Transcription factors (TFs)

Different members of the transcription factor families play a key role in modulating the host immune responses by initiating and regulating transcription of the defense genes (for review, see Eulgem et al. 2007). All the 10 DEGs related to various TFs, including *MYB-DNA binding* (BnaA05g01050D, BnaA05g12950D, BnaA10g17370D), *WRKY DNA-binding domain* and *WRKY54* (BnaA04g23480D, BnaA06g23750D, BnaA07g16850D, BnaA09g24840D, BnaAnng39080D, BnaC06g15910D) and *Dof zinc finger protein DOF5.2* (BnaA04g09490D)

were upregulated in the inoculated CR or CS NILs at 7- and/or 14 dai (Figure 4.4; Supplementary Table S4.3).

Amino acid metabolism

The regulation of amino acid metabolism following *P. brassicae* infection has previously been reported (Wagner et al. 2012, Adhikary et al. 2022b, Sarkar et al. 2023). About 57% (4/7) of the DEGs, viz. *Amino acid kinase family* (BnaA02g09370D)*, Amino-transferase class IV* (BnaA05g36800D)*, Homocysteine S-methyltransferase* (BnaC05g29760D) and *Methionine aminotransferase BCAT4* (BnaC03g41100D) involved in amino acid biosynthesis or metabolism showed an increased expression in the inoculated CS-NIL and decreased expression in the CR-NIL as the infection progressed from 7- to 14 dai; however, the overall expression of these DEGs was low in all samples (Figure 4.4; Supplementary Table S4.3). Only one DEG (BnaA06g05000D) involved in amino acid metabolism was consistently upregulated in the inoculated CR-NIL and CS-NIL, with higher expression at 14 dai in CR-NIL (log_2 fold change = 2.78).

Cell wall modifications

All the five DEGs related to cell wall modifications were consistently upregulated in both the NILs over the course of infection; however, four DEGs (BnaA03g26330D, BnaA10g09640D, BnaAnng29910D, BnaC04g24330D) showed increased expression in SI at 14 dai, as compared to the expression in RI (Figure 4.4; Supplementary Table S4.3).

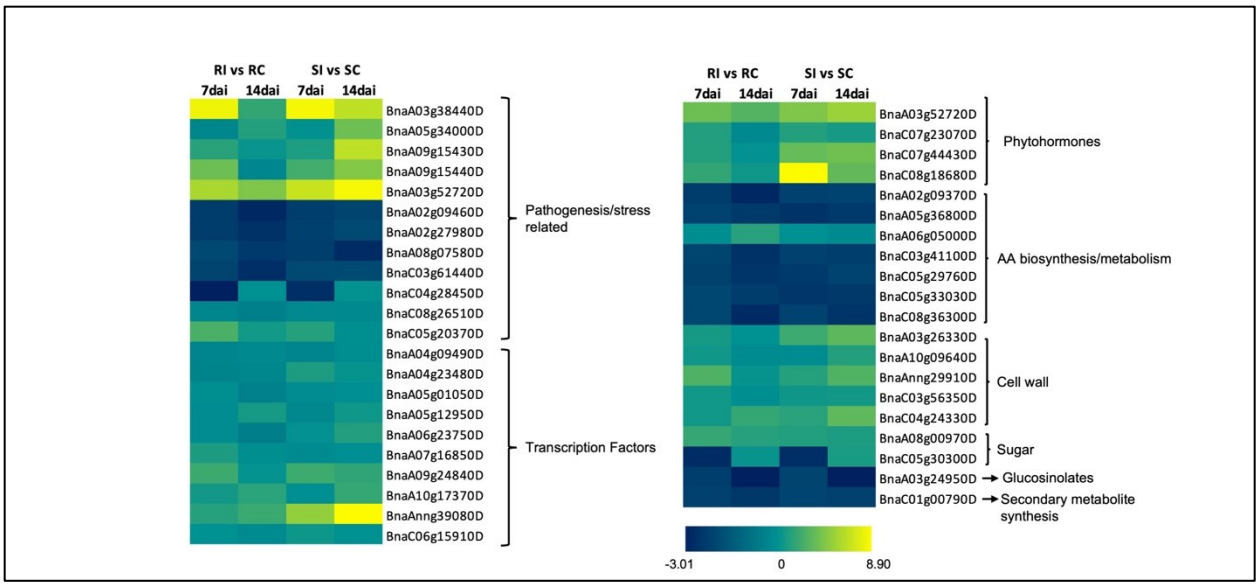


Figure 4.4 Heatmaps representing the expression profiles of the differentially expressed genes (DEGs) related to different biological pathways in the *Brassica napus* near isogenic lines (NILs) inoculated with *Plasmodiophora brassicae* pathotype 3H with respect to their controls. RI = inoculated resistant NILs; RC = uninoculated (control) resistant NILs; SI = inoculated susceptible NILs; SC = uninoculated (control) resistant NILs.

4.3.5 Identification of DEGs based on inoculated samples (RI vs. SI) of the CR-NILs and CS-NILs

Pathogenesis and stress response

While comparing data of the inoculated CR-NIL with the inoculated CS-NIL, we identified 148 DEGs at both time points. Among these, 30 DEGs were associated with stress response, pathogenesis and PRRs, of which about 50% of the DEGs showed an increasing expression as the infection progressed from 7- to 14 dai (Figure 4.5, Supplementary Table S4.4). Ten and 15 putative genes showed a consistent pattern of upregulation and downregulation respectively in the inoculated CR-NIL at both the time points. Five DEGs (BnaC03g57130D, BnaC06g21310D, BnaC06g21340D, BnaC06g21400D, BnaC06g41740D) homologous to *Wound-induced protein* were downregulated at 7 dai while showing upregulation at 14 dai.

Resistance genes (R genes) and TFs

We also detected three R genes that were significantly upregulated in the CR-NILs (Figure 4.5, Supplementary Table S4.4). These three DEGs were homologous to either *NB-ARC* gene family (BnaA08g09210D, BnaA08g10100g) or *Disease resistance protein LAZ5* (BnaA08g09220D) and possess the TIR domain which is known to confer disease resistance. In addition, 12 DEGs related to various TFs (including MYB, WRKY, zinc finger domains) and six DEGs related to phytohormones (mostly ethylene) were also identified. Only two DEGs (BnaAnng24480D which is orthologous to *AtDOF5.5* and a novel gene orthologous to *Oryza sativa* protein *RICESLEEPER2*), were upregulated at both time points. We also identified nine downregulated DEGs, among which six were homologous to *Basic region leucine zipper* (BnaA08g10920D), *Helix loop helix DNA binding domain* (BnaA09g16560D, BnaCnng19670D), *MYB-CC type transfactor* (BnaA08g05130D), *PHD-finger* (BnaA08g10160D) and *transcription regulatory protein SWI3D* (BnaA08g08190D) (Figure 4.5, Supplementary Table S4.4).

Phytohormone mediated signalling

Three ET-related genes annotated as ethylene responsive transcription factors (*APETALA2/ERF018, APETALA2/ERF6, APETALA2/ERF1*) and one JA-related gene, *Allene oxide cyclase*, were identified. Among these, a decreased expression of two ET-related genes (BnaA07g21980D, BnaA08g08300D) was observed as the infection progressed from 7- to 14 dai, while an increased expression of the ET-related gene (BnaA08g08310D) was found at 14 dai (log_2 fold change = -1.43) as compared to 7 dai (log_2 fold change = -2.07) in the CR-NIL. The other JA-related gene (BnaA09g19550D) was upregulated at 14 dai in the CR-NIL (Figure 4.5, Supplementary Table S4.4).

Reactive Oxygen Species (ROS) and calcium mediated signalling

Reactive oxygen species play an important role in host plant response to pathogen attack (for review, see Huang et al. 2019). We detected two DEGs homologous to *Thioredoxin H5* (BnaA08g04320D) and *Aconitase family* (BnaA08g05070D) to be downregulated in the CR-NILs during the progression of *P. brassicae* infection. Calcium plays an important role as a second messenger during plant-pathogen interaction (for review, see Yuan et al. 2017). In this study, two calcium related DEGs were identified showing an increased or decreased expression at 14 dai in response to *P. brassicae* infection in the CR-NILs. The gene homologous to *Calcium uniporter protein 2* (BnaAnng35090D) was upregulated at 7 dai and downregulated at 14 dai and the other gene *Calmodulin-like protein 43* (BnaC02g31060D) was downregulated in the CR-NILs at both time points in response to the pathogen (Figure 4.5, Supplementary Table S4.4).

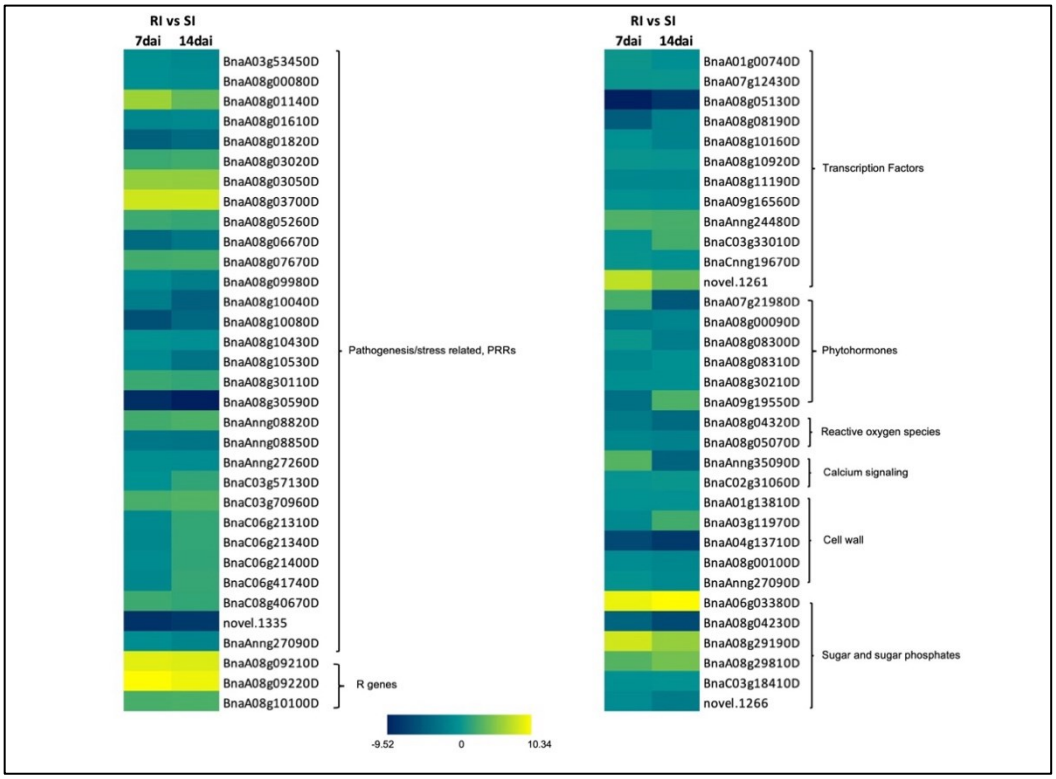


Figure 4.5 Heatmaps representing the expression profiles of the differentially expressed genes (DEGs) related to different biological pathways in the clubroot resistant and susceptible *Brassica napus* near isogenic lines (NILs) (CR-NILs and CS-NILs) inoculated with *Plasmodiophora brassicae* pathotype 3H.

4.3.6 Identification of DEGs involved in signalling mechanisms of PTI and ETI

Based on the formulas described in Table 4.1, we identified three sets of putative genes (Disease symptom + S, PTI + Stress, and ETI, where Disease symptom indicates the DEGs to be playing a role in disease symptom development, Stress indicates the stress related DEGs induced due to pathogen attack, and S indicates the DEGs involved in host susceptibility) to be involved in disease development, PTI and ETI. We identified 541 and 3883 DEGs at 7- and 14 dai, respectively, which are involved in disease symptom development and susceptibility (S); and 454 and 4933 DEGs at 7- and 14 dai, respectively, related to stress induced from pathogen attack or involved in PTI response. Among these, 238 DEGs (109 DEGs with increased expression, 85 DEGs with decreased expression, and the remaining 44 with no change in expression level) were identified at both time points. In addition, 637 and 6533 DEGs were detected, at 7- and 14 dai respectively, which might be playing a role in ETI. From these, 180 DEGs which were detected at both time points; 63 DEGs showed an increased expression and 91 DEGs showed decreased expression (remaining 26 with no change) over the course of infection. Thus, it was evident that the number of DEGs detected at 14 dai was about 7-10-fold greater than the number of DEGs detected at 7 dai. Therefore, we carried out further analysis of the DEGs detected at 14 dai, and categorised these three sets of putative genes (Disease symptom development + S, PTI + stress, and ETI) into different functional categories based on their assigned and known functions to elucidate the role of these genes in susceptibility, PTI and ETI (Figure 4.6).

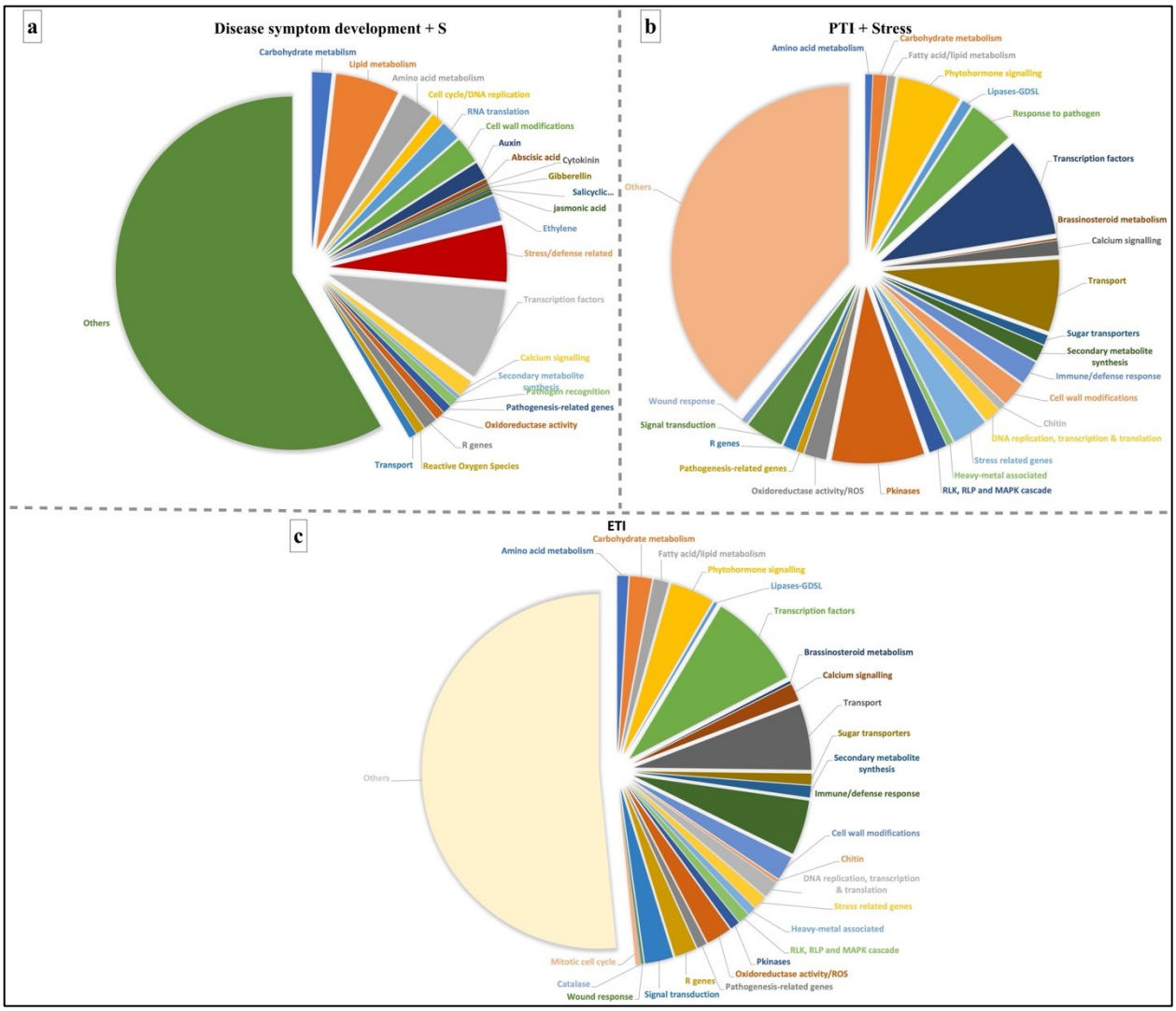


Figure 4.6 Pie charts depicting the categories of differentially expressed genes (DEGs) involved in different biological pathways in the clubroot susceptible (CS-NILs) and resistant (CR-NILs) near isogenic lines (NILs) of *Brassica napus* inoculated with *Plasmodiophora brassicae* pathotype 3H. a) Different categories of genes involved in disease symptom development and susceptibility (S), where n = 2930. b) Different categories of genes involved in pathogen-triggered immunity (PTI) and induced by stress (PTI + Stress), where n = 3895. c) Different categories of genes involved in effector-triggered immunity (ETI), where n = 5004.

Genes responsible for disease symptom development and susceptibility

The upregulated DEGs identified in this study were involved in different pathways in the CS-NILs, including auxin (six DEGs) and cytokinin (five DEGs) production, sugar transport (two DEGs), and in response to water deficit (10 DEGs). The six upregulated DEGs related to auxin biosynthesis

were annotated as *GH3-auxin responsive promoter/protein* (BnaC01g01570D, BnaC03g05810D, BnaC04g09120D, BnaC09g39790D), *indole-3-acetic acid-amido synthetase GH3.17* (BnaC03g05830D) and *SAUR72* (BnaCnng23650D) (Supplementary Table S4.5a). The two DEGs homologous to sugar transporters were annotated as bidirectional sugar transporter *SWEET3* (BnaA02g10400D) and *SUC2* (BnaA09g30430D). We also identified the downregulated DEGs related to antibacterial defense (five DEGs) and R genes (18 DEGs) in the CS-NILs. Another DEG (BnaA03g24410D) annotated as *DMR6-like oxygenase 1*, known to cause susceptibility to downy mildew (Zeilmaker et al. 2015), was upregulated in the CS-NILs. Ten DEGs homologous to *xylanase inhibitor* gene were downregulated in the CS-NILs. Xylanase produced by phytopathogens destroy the plant cell walls; therefore, xylanase inhibitor plays a vital role in plant survival (for review, see Tundo et al. 2022).

PTI: PRRs and MAPK cascade

A total of 25 DEGs encoding PRRs, including RLKs, RLPs and Leucine-rich repeat receptor-like protein kinases (LRR-RKs) were identified. Among these, nine DEGs homologous to the *RLK SUPPRESSOR OF BIR-1* (*SOBIR1*, BnaC03g17800D), *RLP12* (BnaA01g31240D), *RLP30* (two; BnaA06g12200D, BnaC05g13910D), *RLK10* (novel gene), *RLK45* (BnaC03g72920D) and *PERP2* (two novel genes) were found to be upregulated in the CR-NILs (Supplementary Table S4.5b). Four DEGs homologous to mitogen-activated protein kinases viz., *AtMKK4* (BnaA08g01860D), *MAPKKK NPK1* (BnaA10g29440D), *MAPKKK ANP1* (BnaCnng64060D) and *MPK1* (BnaC03g16550D) were upregulated while two DEGs homologous to *MAPKKK YODA* (BnaA01g00740D, BnaC01g01730D) were downregulated in the CR-NILs. Another class of cell surface immune receptors called *Wall-associated receptor kinases* (*WAKs*) was found to be

differentially expressed in four DEGs, of which one DEG homologous to *WAK4* (BnaA07g11030D) was found to be upregulated in the CR-NILs.

PTI: Ca^{2+} influx and signalling mediated resistance

The activation of PRRs leads to the different downstream responses, including calcium burst (for review, see Zipfel 2014). *Calmodulin-gated calcium channels* (*CNGCs*) and *Glutamate Receptor-like* (*GLR*) genes induce the influx of calcium after pathogen recognition. Two DEGs homologous to calcium channels *CNGC15* (BnaA07g13760D) and *CNGC20* (BnaC03g40070D) were upregulated in the CR-NILs. We also identified three DEGs homologous to *GLRs*, among which two DEGs (BnaA08g29720D, BnaCnng22760D) were upregulated in the CR-NILs. Following the calcium influx, the intracellular calcium receptor proteins such as *Ca^{2+}- dependent protein kinase* (*CDPK*) are activated and play an important role in plant defense (for review, see Ren et al. 2021). We found a gene homologous to *CDPK28* (BnaA07g12770D) to be upregulated in the CR-NILs. Three calcium-binding proteins viz., *CBP42* (BnaC07g36400D), *CBP43* (BnaC02g31060D) and *CBP45* (BnaA04g09510D) were also upregulated in the resistant NILs (Supplementary Table 4.5b).

PTI: ROS signalling mediated resistance

Intracellular Ca^{2+} influx activates the *Respiratory Burst Oxidase Homolog* (*RBOH*) proteins, which leads to an increased production of ROS (Stael et al. 2015). In the present study, we found an upregulation of a transcript annotated as *RBOHA* in the CR-NILs. ROS can increase host-resistance by acting as antibacterial effector or can get degraded to elicit the defense responses (Fang 2011). In our analysis, we found 24 DEGs homologous to *peroxidases* (*PODs*), among which 14 DEGs were upregulated in the CR-NILs. Of the 37 DEGs annotated as *Glutathione-S-transferases* (*GSTs*), 33 were upregulated in the CR-NILs. Two (BnaA09g10710D and a novel

gene) out of three DEGs homologous to *Superoxide dismutases* (*SODs*) were upregulated in the resistant NILs. In addition, five DEGs homologous to *catalase* (*CAT*) were identified of which four DEGs (BnaA07g11370D, BnaA08g21730D, BnaC05g15970D, BnaC07g15280D) were upregulated in the CR-NILs (Supplementary Table S4.5b); these genes (*POD*, *GST*, *SOD* and *CAT*) exhibit a role in degradation of ROS.

ETI: NLR-mediated resistance

NLRs are the intracellular receptor proteins which detect the effectors produced by the pathogen and elicit the ETI (for review, see Zhou and Zhang 2020). There are three main categories of NLRs i.e., *coiled-coil-type NLRs* (*CNLs*), *Toll/interleukin-1 receptor/resistance protein (TIR)-type NLRs* (*TNLs*), and *Resistance to Powdery Mildew 8-like domain* (*RPW8*)*-type NLRs* (*RNLs*). In the current study, 47 DEGs were homologous to R genes (*CNLs*, *TNLs* and *RNLs*), among which 38 were upregulated in the CR-NILs. These 38 DEGs included genes homologous to *Arabidopsis* "*RESISTANT TO P. SYRINGAE 4*" (*RPS4*), "*RESISTANT TO P. SYRINGAE 5*" (*RPS5*), "*RESISTANT TO P. SYRINGAE 6*" (*RPS6*), "*RECOGNITION OF PERONOSPORA PARASITICA 1*" (*RPP1*), "*TOLERANCE TO TOBACCO RINGSPOT VIRUS*" (*LAZ5*), "*RESISTANCE TO LEPTOSPHAERIA MACULANS 1A*" (*RML1A*), "*TARGET OF AVRB OPERATION 1*" (*TAO1*), *Arabidopsis* broad-spectrum mildew resistance protein *RPW8,* and many others (Supplementary Table 4.5c). Another R gene called "*non-race specific disease resistance protein 1"* (*NDR1*), which is an integrin-like R protein, was also found to be upregulated in the CR-NILs. We also identified the upregulation of two DEGs (BnaA04g05430D, BnaCnng154590D) homologous to *Phytoalexin Deficit 4* (*PAD4*).

4.3.7 Validation of RNA sequencing data by qRT-PCR

To confirm the expression level of the DEGs identified by RNA-seq, we randomly selected 18 putative genes to evaluate their transcription levels using qRT-PCR (Supplementary Table S4.6). These DEGs were related to resistance, plant-pathogen interaction, phytohormones and transcription factors. Expression change of all the tested genes showed a similar pattern of upregulation or downregulation in both the RNA-seq and qRT-PCR analysis indicating that the RNA-seq results were reliable (Figure 4.7).

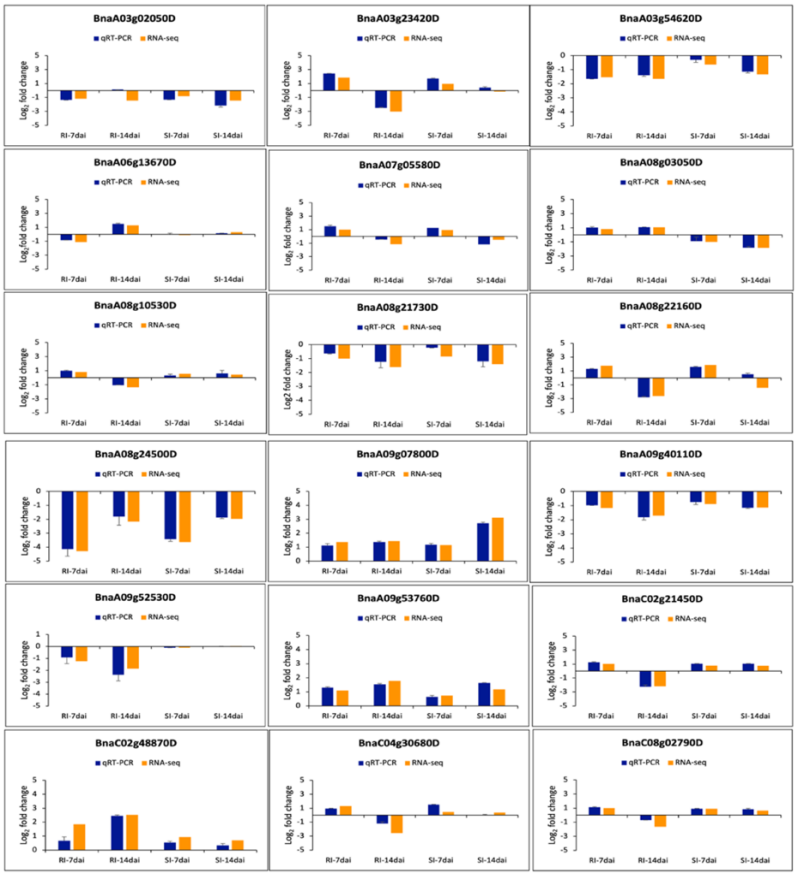


Figure 4.7 qRT-PCR validation of 18 randomly selected differentially expressed genes (DEGs) identified through RNA-seq. The log_2 fold changes were calculated between inoculated resistant (RI) and inoculated susceptible samples (SI) with respect to their controls. Error bars show the standard error of mean for three biological replicates. RI-7dai: resistant inoculated at 7 days after inoculation, RI-14dai: resistant inoculated at 14 days after inoculation, SI-7dai: susceptible inoculated at 7 days after inoculation and SI-14dai: susceptible inoculated at 14 days after inoculation.

4.4 Discussion

Several researchers (Chu et al. 2014; Chen et al. 2016; Jia et al. 2017; Ciaghi et al. 2019; Fu et al. 2019; Mei et al. 2019; Wang et al. 2019; Yuan et al. 2021b, Adhikary et al. 2022b, Sarkar et al. 2023) carried out transcriptome analysis of Brassica root samples following inoculation with *P. brassicae* to understand host-pathogen interaction and to identify the genes involved in clubroot resistance. However, only a few studies (Chen et al. 2016; Li et al. 2021) have utilised NILs to elucidate the pathogen-induced changes in host contributing to clubroot resistance. In the present study, we carried out a comparative transcriptomics analysis using CR-NILs and CS-NILs of *B. napus* carrying a novel resistance introgressed from turnip (*B. rapa* var. *rapifera*) (Kaur et al. 2022). The NILs that we used in this study showed resistance to at least 15 pathotypes, including the recently evolved virulent ones. This provided us not only the opportunity to identify the defense-responsive genes related to different metabolic pathways with enhanced accuracy, but also the genes involved in resistance to a broad spectrum of pathotypes. We witnessed different resistance responses at primary (7 dai) and secondary (14 dai) infection stages. Furthermore, the total number of identified DEGs increased with the progression of disease; this result is consistent with Ning et al. (2019). We found a greater number of DEGs upregulated in the CR-NILs at both the time points as compared to the number of DEGs identified in the CS-NILs. Based on this, it can be speculated that the resistant NILs exhibited an intense defense response and were more sensitive to attack by *P. brassicae*.

The functional enrichment analysis of DEGs showed an activation of different resistance-related pathways, including phytohormone signal transduction, R genes, genes encoding PRRs, MYB and WRKY transcription factors, calcium signaling, and glucosinolates biosynthesis pathway in the resistant NILs. This is in accordance with the already published studies in

Arabidopsis, *B. rapa*, *B. oleracea* and *B. napus* (Chen et al. 2016; Zhang et al. 2016; Jia et al. 2017; Irani et al. 2018; Fu et al. 2019; Ning et al. 2019; Adhikary et al. 2022b, Sarkar et al. 2023). Ludwig-Muller et al. (2009) reported an upregulation of the genes involved in photosynthesis and increased transport of photosynthetic products from leaves to the developing galls in infected *B. rapa*. Recently, Wang et al. (2022) reported the role of sugar transport genes in clubroot susceptibility. Interestingly, the GO annotation analysis in our study showed an upregulation of the gluconeogenesis, hexose biosynthetic process, and monosaccharide biosynthetic process in the inoculated CS-NILs as compared to control. We also found the upregulation of auxin and cytokinin biosynthesis genes as well as the sugar transporter genes in the inoculated CS-NILs, which is in accordance with the already published reports (for review, see Hasan et al. 2021a). The DEGs associated with the above-mentioned pathways that may contribute to clubroot resistance in the CR NILs are discussed below (Supplementary Table S4.7).

Phytohormone signal transduction

Phytohormones such as auxin, cytokinin, SA, JA and ET regulate the biotic and abiotic stress responses (for review, see Han and Kahmann 2019). The accumulation of auxin is essential during clubroot disease progression and gall formation (Ludwig-Muller et al. 2009). The cytochrome P450 family genes play a vital role in the transformation of tryptophan to indole-3-acetaldoxime, a precursor for indole-3-acetic acid (Gonzalez-Lamothe et al. 2009). Six DEGs belonging to this gene family were downregulated in the infected CR-NILs further supporting the role of auxin in clubroot disease development (Supplementary Table S4.7). We also found the upregulation of five DEGs related to auxin in the infected CS-NILs, thus aiding in auxin supply to the gall-forming roots (Supplementary Table S4.5a). The accumulation of JA in the roots of clubroot-susceptible *Arabidopsis* and *B. napus* increases after infection with *P. brassicae* (Gravot et al. 2012; Xu et al.

2018). Jasmonate ZIM domain-containing proteins negatively regulate the biosynthesis of JA. The upregulation of JAZs in the clubroot-resistant *B. oleracea* has been reported by Ning et al. (2019). We also observed the upregulation of *JAZ2* protein (BnaC06g22690D) in the inoculated CR NILs. In contrast, no significant expression of this gene family was found in the susceptible NILs.

NIM1-Interacting 2 (*NIMIN2*) together with *Nonexpressor of PR gene 1* (*NPR1*) play a major role in the regulation of SA-induced host resistance (Koornneef and Pieterse 2008). In this study, we found that *NIMIN2* (BnaA02g27670D) was upregulated in the CR-NILs providing further evidence that SA is involved in the resistance mechanism. The *Ethylene Responsive Factors* (*ERFs*) involved in the ET-mediated signal transduction can be activated or repressed in the host in response to pathogen attack (Müller and Munné-Bosch 2015). Adhikary et al. (2022b) found *ERF4* homolog to be upregulated and *ERF43* to be downregulated in clubroot resistant genotype. In the current study, we found *ERF1A* and *ERF6* to be downregulated at both the time points whereas *ERF018* to be upregulated at 7 dai but downregulated at 14 dai in the CR-NILs. Thus, it is evident that *ERFs* are involved in the *Brassica-P. brassicae* host-pathogen interaction.

Resistance genes

The R genes are the vital component of ETI and act by recognizing the effector molecules released by the pathogen. The NBS-LRR (NLR) is the largest group of R protein genes in plants. The TIR (Toll-Interleukin-1) domain is reported to play a signaling role in the induction of defense response to *P. brassicae* (Swiderski et al. 2009). The complete susceptibility to clubroot in *B. rapa* has been reported due to the lack of about half of the *TIR* domain (Hatakeyama et al. 2013). The upregulation of R genes containing the *TIR* (BnaA08g09210D, BnaA08g09220D, BnaA08g10100D) or LRR domain (BnaA01g31240D, BnaA03g23420D) in the CR-NILs observed in our study is in accordance with the previously published reports (Chu et al. 2014, Chen

et al. 2016, Irani et al. 2018, Ning et al. 2019, Fu et al. 2019) (Supplementary Table 4.7). The R genes belonging to the other two classes i.e., *CNLs* and *RNLs* were also upregulated in the infected CR-NILs (Supplementary Table S4.5c). The downregulation of R genes has also been found in the infected CS-NILs in the current study (Supplementary Table 4.5a).

Pathogenesis-related proteins

The production of pathogenesis-related (PR) proteins is a part of the hypersensitive response or systemic acquired response against phytopathogens (for review, see Linthorst and Van Loon 1991). There are different classes of PR proteins including chitinases (PR-3), thaumatins (PR-5) and lipid transfer proteins. Chitin is the main component of the cell wall of *P. brassicae* (Moxham and Buczacki 1983). It is one of the PAMPs, released in response to infection by most of the pathogens (for review, see Latgé and Beauvais 2014), but its role as a PAMP in *P. brassicae* still needs to be investigated. The enzyme chitinase catalyses the degradation of chitin and acts as a PR protein in resistance against pathogen infection (Schwelm et al. 2015). Three genes were upregulated in the *B. napus* CR-NILs including one *PR-3 chitinase 10* and two *chitinase class I* (*PR-4*) genes. Glazebrok (2005) has reported that *PR-3* is controlled by JA/ET dependent pathway and the induction of *PR-3* along with *PR-4* provides resistance against necrotrophic pathogens. Therefore, the upregulation of *PR-3* and *PR-4* further confirms the role of JA and ET hormones in triggering response against clubroot. The thaumatin family proteins are upregulated in many plants in response to the bacterial and fungal infection (for review, see Liu et al. 2010). We also found the upregulation of four genes belonging to this family, thus providing evidence of the role of this protein in conferring resistance against *P. brassicae*.

Transcription factors

In plants, TFs play a vital role in the initiation and regulation of gene transcription and modulating the defense responses against pathogen invasions. We observed the upregulation of *WRKY62*, *WRKY70* and *zf-dof*, and downregulation of *bHLH55* with disease progression in the resistant NILs. According to Xu et al. (2006), a triple mutant of *AtWRKY18*, *AtWRKY40* and *AtWRKY60* exhibits an increased level of defense response against *P. synringae* and *G. orontii*. These three TFs are homologous to the Arabidopsis *WRKY62* transcription factor. Ulker et al. (2007) demonstrated the involvement of *WRKY70* in senescence, and biotic and abiotic stress responses by positively regulating the accumulation of SA and negatively regulating the JA-mediated signaling pathway. Thus, it is highly likely that an upregulation of *WRKY70* in our study has activated the genes for SA accumulation in the CR-NILs as reported in the previous paragraphs, further confirming the role of SA in resistance against *P. brassicae* in canola.

Calcium influx and ROS

Calcium acts as a second messenger in signal transduction leading to the production of ROS, which regulates the stress responses in plants (Zipfel 2010). Peroxidases cause programmed cell death by catalysing the production of hydrogen peroxide and eventually limiting the spread of disease (for review, see Apel and Hirt 2004). Adhikary et al. (2022b) reported the upregulation of DEGs homologous to EF-hand domain pair and downregulation of *calcium-transporting ATPase 10*, *calmodulin-binding protein-like* and *calcineurin B-like protein 10*. In this study, we found an increased expression of two *mitochondrial calcium uniporter* genes, one *calcineurin-like phosphoesterase* gene and two *peroxidases* in the CR-NILs post inoculation with *P. brassicae*. Beside this, we also found the *calmodulin-like protein 45* (CML-45) was upregulated at 7 dai but downregulated at 14 dai indicating that this gene contributes to the defense response at early stage of infection.

Glucosinolates biosynthesis

Glucosinolates are biologically active secondary metabolites found in the Brassicaceae family that contribute to biotic stress tolerance. In this study, we found a putative gene (BnaA06g12720D) homologous to *sulfotransferase* (*SOT*) consistently upregulated in the CR NILs. The upregulation of four genes homologous to *sulfotransferase* domain in the clubroot resistant genotype of *B. napus* has been reported (Adhikary et al. 2022b). *Glutathione-S-transferases* (*GSTs*) are involved in the secondary metabolism in plants such as the formation of sulphur-containing glucosinolates in *Brassica* species (for review, see Czerniawski and Bednarek 2018). Their accumulation in response to biotic stresses has also been reported (for review, see Chhajed et al. 2020). Four putative genes from the *Phi GST* category (two from *GSTF2* and two from *GSTF3*) were upregulated in the resistant NILs over both 7- and 14-dai. *Adenylyl-sulphate kinase* (*APK*) is required for the metabolism of primary sulphur in the biosynthesis of glucosinolate (Yatusevich et al. 2010). Two candidate *APK* genes (BnaA03g54400D and BnaC07g51290D) were upregulated in the resistant NILs providing evidence of the involvement of these genes in clubroot resistance.

4.5 Model of the molecular response to *P. brassicae* in the resistant *B. napus* NILs

Based on the above discussion, we propose a model of the potential resistance mechanism in the *B. napus*-*P. brassicae* pathosystem (Figure 4.8). Upon the pathogen infection of *B. napus*, the cell surface receptors including RLKs (*RLK10*, *RLK45*), RLPs (*RLP12*, *RLP30*), *SOBIR1*, *PERP2* and WAKs (*WAK4*) detect the extracellular PAMPs/DAMPs released by the pathogen and initiate the first tier of immunity i.e., PTI. The perception of the signals by PRRs results in activation of different downstream molecular events such as activation of MAPK cascade, Ca^{2+} influx, ROS production and biosynthesis of phytohormones (for review, see Peng et al. 2018a). Based on our study, *MPK1* and *AtMKK4* serve as the points of signalling downstream of PRRs. Both these

MAPKs are known to be activated in response to pathogen infection in *Arabidopsis* (Nitta et al. 2014). Activated MAPKs lead to the phosphorylation of TFs to increase their transcriptional activity. For example, the activation of *WRKY46* and *WRKY70* in the CR-NILs explains their role in resistance mechanism. These WRKY TFs positively regulate the SA accumulation. Two Ca^{2+} channels *CNGC15* and *CNGC20* as well GLRs aided in the influx of calcium into the cell's cytoplasm. The Ca^{2+} influx leads to the activation of other transmembrane channels such as influx of H^+ and efflux of K^+, Cl^- and $NO3^-$. The calcium dependent protein kinase *CDPK28* phosphorylates the NADPH oxidase *RBOHA* for ROS production. *CDPK28* has been reported to play a role in ROS signalling and maintaining balance of phytohormones in *Arabidopsis* (Matschi et al. 2015). The ROS produced are degraded by peroxidases, catalases, superoxide dismutases and GSTs resulting in the defense responses against *P. brassicae*. In parallel, the intracellular receptors NLRs (TNLs and RNLs) detect the virulence proteins called effectors leading to a much stronger ETI. *NDR1*, an integrin-like protein, is suggested to have a role in general plant immunity (Knepper et al. 2011). *NDR1* is required for the resistance mediated by CNL genes (*RPS5*). *NDR1* along with *PAD4* helps in SA accumulation via induction of SA/NDR1 mediated pathway. Many other TNL and RNL genes also contribute to clubroot resistance in *B. napus*.

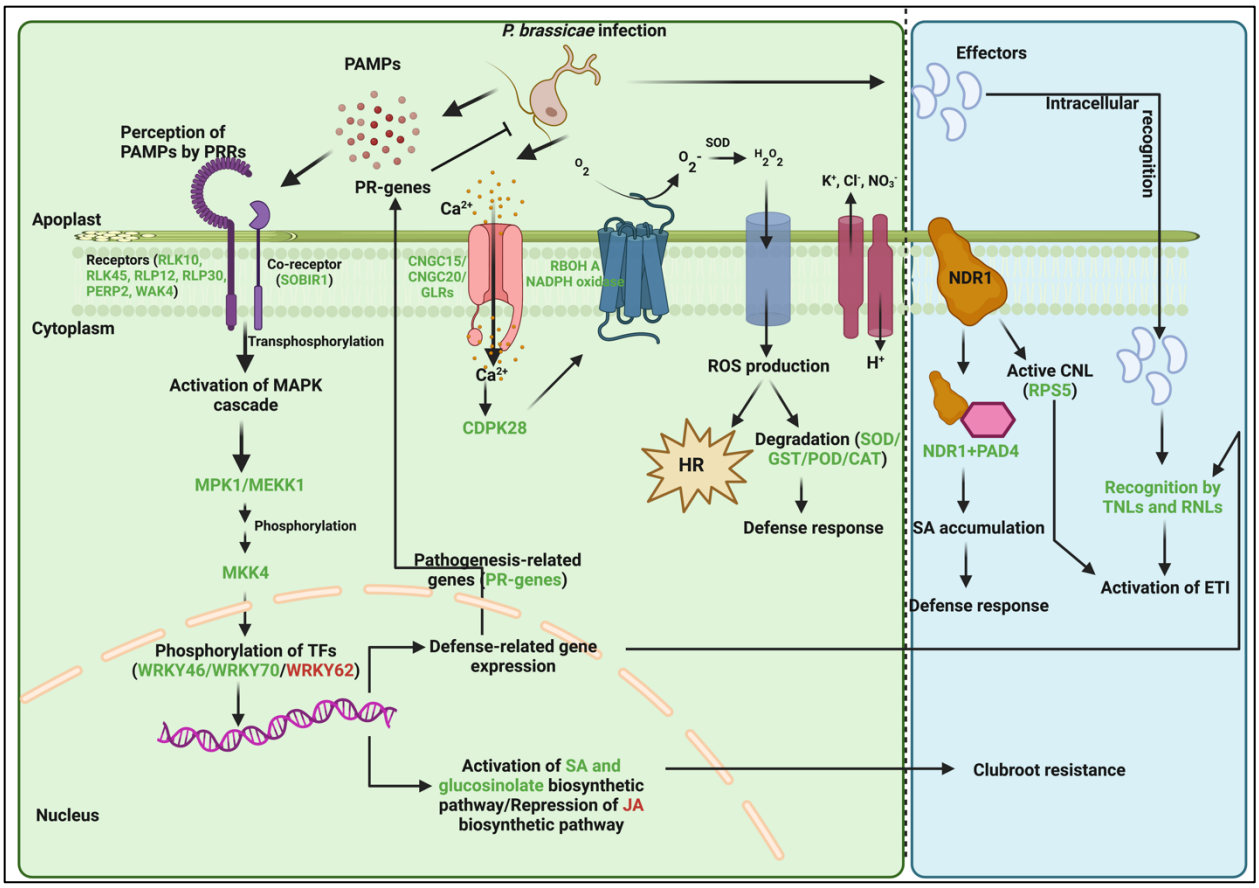


Figure 4.8 Schematic model depicting the potential signalling network in the clubroot resistance mechanism in inoculated resistant near-isogenic lines (CR-NILs) of *Brassica napus*. Genes/pathways in green are upregulated, while the ones in red are downregulated. PAMPs, Pathogen-associated molecular patterns; PRRs, Pathogen Recognition Receptors; RLK10, Receptor-like kinase 10; RLK45, Receptor-like kinase 45; RLP12, Receptor-like protein 12; RLP30, Receptor-like protein 30; PERP2, PEP receptor 2; WAK4, Wall-associated receptor kinase 4; SOBIR1, SUPPRESSOR OF BIR-1; MAPK, mitogen-activated protein kinase; MPK1/MEKK1, mitogen-activated protein kinase 1; MPK4, mitogen-activated protein kinase 4; CNGC15, Calmodulin-gated calcium channel 15; CNGC20, Calmodulin-gated calcium channel 20; GLRs, Glutamate Receptor-like; CDPK28, Calcium-dependent protein kinase 28; RBOHA, Respiratory Burst Oxidase Homolog A; SOD, Superoxide dismutase; GST, Glutathione-S-transferase; POD, Peroxidase; CAT, Catalase; NDR1, non-race specific disease resistance protein 1; PAD4, Phytoalexin-deficient 4; RPS5, RESISTANT TO P. SYRINGAE 5; TNLs, Toll/interleukin-1 receptor/resistance protein (TIR)-type NLRs; RNLs, Resistance to Powdery Mildew 8-like domain (RPW8)-type NLRs; SA, Salicyclic acid; JA, Jasmonic acid; HR, Hypersensitive response; ETI, Effector-triggered immunity. The arrows indicate positive regulation and the arrows with bar (inhibitors) indicate the negative regulation of clubroot resistance. The green shaded area denotes PAMP-triggered immunity (PTI) and the blue area denotes ETI.

4.6 Conclusion

The investigation of global transcriptome profile of roots of *B. napus* CR- and CS-NILs demonstrated a complex and distinct response at primary and secondary infection phases of *P. brassicae*. The upregulation of genes associated with phytohormone signaling, disease resistance, calcium influx and ROS, WRKY TFs and glucosinolates biosynthesis demonstrated their role in conferring resistance against clubroot disease in *B. napus*. Partitioning the DEGs into different sub-categories, including the genes involved in PTI and ETI further extended our knowledge of the *B. napus*-*P. brassicae* interaction. Thus, we provided a comprehensive transcriptomic expression landscape of the DEGs which could be used as the promising candidates for functional validation. This would extend our knowledge on the molecular mechanisms underlying clubroot resistance in *B. napus*.

Chapter 5. Proteome level investigation of host responses to *P. brassicae* infection in *B. napus* near-isogenic lines carrying resistance of turnip.

5.1 Introduction

Canola (Canadian Oil, Low acid), also known as oilseed rape, is the result of classical breeding of *B. napus* to improve the oil quality (less than 2% erucic acid) and meal composition (less than 30 μmol of aliphatic glucosinolates g^{-1} meal). It is the widely cultivated cruciferous oilseed crop with a global production of 86.31 million metric tons (MMT) in 2022-2023 (USDA FAS, 2023). Canola, developed in Canada about 50 years ago, is now at the centre of a burgeoning industry in Canada with an international reach generating $29.9 billion in national economic activities every year (Canola Council of Canada, 2021- Industry Overview). There was 35% increase in the canola's contribution to the Canadian economy in the last 10 years. Unfortunately, canola production is threatened by several biotic stresses, including clubroot disease caused by *Plasmodiophora brassicae* Woronin (Buczacki 1983). This disease affects Brassica crops in more than 60 countries resulting in 10-15% yield loss globally (for review, see Dixon 2009).

P. brassicae, a soil-borne obligate biotrophic protist, can thrive in soil as resting spores for over 17 years (Kageyama and Asano 2009). With favourable environmental conditions (warm and wet soil with pH <6.5), these resting spores produce biflagellate primary zoospores which infect the root hair of susceptible host plants. This is followed by cortical cell infection at the secondary plasmodia phase, which leads to root hormone imbalance resulting in hypertrophy and hyperplasia. This eventually causes the formation of the club-shaped galls in the roots, a typical symptom of clubroot disease (Kageyama and Asano 2009; for review, see Ludwig-Muller et al. 2009). Management of this disease through cultural practices includes sanitation of the field equipment, application of soil amendments like lime, use of fungicides, application of boron, and crop rotation

(for review, see Donald and Porter 2009). However, many of these strategies are not economical or effective in a large cropping system. Therefore, deployment of clubroot resistant cultivar is the most promising and environmental-friendly approach for combating this disease. However, the evolution of new pathotypes with the ability of rendering the available resistances ineffective poses a major challenge in managing this disease (Strelkov et al. 2018; Hollman et al. 2021).

The use of omics technologies, such as genomics, transcriptomics, proteomics and metabolomics, have increased our knowledge of plant-pathogen interactions and helped to identify the QTL/R genes to be associated with disease resistance. To date, several transcriptomics studies have been carried out to understand the *Brassica-P. brassicae* interactions (e.g., Adhikary et al. 2022b; Jia et al. 2017; Fu et al. 2019); however, most of the studies are not necessarily focused on the identification of the proteins corresponding to the transcribed genes. If the transcriptome denotes what a cell can possibly do, the proteome denotes what a cell is supposed to do at a particular point of time. Therefore, study at the proteome level focusing on signaling pathways, secondary metabolic processes and protein-protein interactions is needed for revealing the functional players associated with clubroot resistance mechanisms in plants. The field of proteomics has evolved considerably over time with respect to the techniques used to study the proteomes. Both gel-based and gel-free methods have been used to identify and characterize the proteins (Jorrin-Novo 2020). The two-dimensional polyacrylamide gel electrophoresis (2-DE) has been widely used for resolving the proteins based on molecular mass (MM) and isoelectric point (pI) (for review, see Aslam et al. 2017). With the technological advancements, the fourth-generation mass spectrometry (MS) techniques such as selective reaction monitoring (SRM), multiple reaction monitoring (MRM) and parallel reaction monitoring (PRM) (for review, see Boersema et al. 2015) has allowed the identification of the proteins involved in biotic and abiotic

stresses. iTRAQ (Isobaric Tag for Relative and Absolute Quantitation), a multiple protein labelling technique, has been used for high-throughput quantification of proteins. NMR (Nuclear Magnetic Resonance) spectroscopy, a high-throughput tool, has been used to investigate the 3-dimensional molecular structure of the proteins, which can be used in functional genomics analysis (for review, see Aslam et al. 2017).

Comparative proteomics analysis has been carried out in *B. rapa*, *B. oleracea* and *B. napus* at different time points to get an insight into the intricate plant-pathogen interaction mechanism and to identify the candidate proteins involved in the metabolic or signaling pathways in host defense against *P. brassicae* (Cao et al. 2008; Song et al. 2016; Ji et al. 2018, 2021; Lan et al. 2019; Moon et al. 2020; Adhikary et al. 2022a). Ji et al. (2018, 2021) reported several differentially abundant proteins (DAPs) to be involved in phytohormone regulation at early and late stages of infection by *P. brassicae* in *B. rapa*. Cao et al. (2008) and Lan et al. (2019) reported DAPs associated with the cytokinin pathway in the secondary stage of infection in *B. napus* and *B. rapa* respectively. Su et al. (2018) suggested the role of brassinosteroid (BR) pathways at secondary stage of infection by *P. brassicae* in *B. rapa*. Song et al. (2016) reported a few hundred DAPs acting in calcium-independent manner via a cascade of MAPK involved in mediating clubroot resistance (mediated by gene *Rcr1*) in *B. rapa*. Adhikary et al. (2022a) identified 73 putative proteins orthologous to clubroot resistant proteins and revealed that the proteins involved in the secondary metabolite pathways are related to the resistance mechanism.

To date, most of the proteomics studies on clubroot resistance have been carried out using clubroot resistant and susceptible cultivars, or lines from a segregating population. These types of genetic materials may differ not only for the clubroot resistance genes, but also for several other genes controlling other traits. Due to this, the DAPs identified from a proteomics study using this

type of genetic materials may reflect not only the proteins involved in clubroot resistance but also many other proteins associated with other traits. This constraint can be avoided largely by using near-isogenic clubroot resistant and susceptible lines. In this study, we performed a comparative proteomics analysis of the roots of clubroot resistant (CR) and clubroot susceptible (CS) near-isogenic lines (NILs) of *B. napus* canola at three different time points of infection by *P. brassicae* to identify the proteins playing a role in clubroot resistance. To our knowledge, this is the first proteome-level investigation into clubroot resistance in canola using the NILs.

5.2 Materials and Methods

5.2.1 Plant materials

A set of near isogenic lines (NILs) contrasting for their resistance to *P. brassicae* pathotype 3H were used in this study. The NILs were developed by crossing a clubroot resistant *B. napus* line carrying resistance introgressed from *B. rapa* ECD 01 (cv. Debra) (Kaur et al. 2022) and a clubroot susceptible *B. napus* canola line A04-73NA. The F_1 plants was subjected to recurrent backcrossing for four generations (BC_4) using A04-73NA as the recurrent parent along with selection for resistance to pathotype 3H in each generation. The clubroot resistant BC_4 plants were self-pollinated for two generations to develop the homozygous resistant and homozygous susceptible BC_4F_3 NILs; these NILs hereafter will be referred to as CR-NILs and CS-NILs, respectively. Three resistant and three susceptible BC_4F_3 families derived from a single heterozygous BC_4 plant were used in this study to minimize the genetic background difference of the NILs.

5.2.2 Preparation of inoculum and inoculation technique

The single spore isolates of pathotype 3H, preserved at -80 °C in the galls of clubroot susceptible *B. napus* cv. Hi-Q were used. For preparation of inoculum, 37 g frozen galls were thawed, homogenized in 1000 mL of distilled water, and filtered through multi-layered cheesecloth

(American Fibre and Finishing Inc., Albermarle, NC, US). The density of the resting spores was adjusted to 1×10^7 spores per mL. The NILs were seeded in 32-cell trays filled with Sunshine Professional Growing Mix (Sunshine Horticulture, 15831, N.E., Bellevue, WA, USA), and the plants were grown in a greenhouse at 20-22/15°C (day/night) temperature, 16/8 hours (h) photoperiod and light intensity of 400 µmol/m^2. For inoculation, 1 mL of spore suspension was injected into soil close to the root of the seedlings at 7-10 days after germination. The control plants were treated with sterile water and were maintained in separate trays under same greenhouse conditions. After inoculation, the soil was kept well saturated with water for about two weeks for successful infection. The plants were fertilized using Nitrogen-Phosphorus-Potassium (20-20-20) once a week.

5.2.3 Sample collection for protein extraction

Roots of the CR-NILs and CS-NILs collected at 7-, 14- and 21 days post inoculation (dpi) were used for this study. According to Summanwar et al. (2019), primary infection occurs after 6 dpi, a large number of secondary plasmodia become visible at 14 dpi, and resting spores become visible at about 21 dpi; therefore, we have chosen these three time points. Each root sample included a bulk of three NIL families and seven plants of each NIL family, i.e., a total of 21 plants. The experiment was replicated four times with or without (control) inoculation. Thus, the total number of samples was 2 genotypes (CR-NILs and CS-NILs) × 3 time points × 2 treatments (inoculated and un-inoculated) × 4 biological replicates = 48.

5.2.4 Protein extraction for proteome analysis

Frozen root tissue samples were ground in liquid nitrogen using a mortar and pestle and were aliquoted for protein extraction. Proteins were extracted from 200 mg of ground tissue with a solution of 50 mM HEPES-KOH pH 8.0, 50 mM NaCl, and 4% (w/v) SDS at a 1:2 (w/v) ratio.

The protein extract was reduced with 10 mM dithiothreitol (DTT) and alkylated with 30 mM iodoacetamide (IA). Peptide pools were generated using a KingFisher APEX (ThermoScientific) automated sample preparation device as outlined by Leutert et al. (2019) without deviation. Samples were digested with sequencing grade trypsin (V5113; Promega) and acidified with Formic Acid (FA) to a final concentration of 5% (v/v). Following digestion, samples were acidified with formic acid (A117, Fisher) to a final concentration of 0.5 % (v/v). Peptides were desalted as previously described (Uhrig et al. 2019) using an OT-2 liquid handling robot (Opentrons Labworks Inc.) mounted with Omix C18 pipette tips (A5700310K; Agilent). Desalted peptides were dried and stored at -80 ºC prior to re-suspension in 3.0 % (v/v) ACN / 0.1 % (v/v) FA and before MS injection.

5.2.5 Nanoflow LC-MS/MS analysis

Peptides were analyzed using a Fusion Lumos Tribrid Orbitrap mass spectrometer (Thermo Scientific) in data independent acquisition (DIA) mode, using the previously described BoxCarDIA method (Mehta et al. 2022). One microgram of resuspended peptide was injected using an Easy-nLC 1200 system (LC140; ThermoScientific) and separated on a 25 cm Easy-Spray PepMap C18 Column (ES902; Thermo- Scientific). A 40-minute nonlinear gradient was used as described previously (Mehta et al. 2022) with peptides eluted with a solvent B gradient (0.1% (v/v) FA in 80% (v/v) ACN) consisting of 4−41% B. BoxCarDIA MS^1 analysis was performed by using two multiplexed targeted SIM scans of 10 BoxCar windows each. Detection was performed at 120,000 resolution and normalized Automatic Gain Control (AGC) targets of 100% per BoxCar isolation window. MS^2 analysis was performed at a resolution of 30,000 using twenty-eight 38.5 m/z windows with an overlap of 1 m/z and an AGC target value of 2000%.

5.2.6 Functional annotation and enrichment of differentially abundant proteins

All BoxCarDIA files were processed with Spectronaut v17 (Biognosys AG) using default settings without N-acetyl variable modification. Spectra were searched using the published *B. napus* Westar proteome (Song et al. 2020). Significantly changing differentially abundant proteins were determined and corrected for multiple comparisons (Bonferroni- corrected p-value < 0.05; q-value). To gain a better understanding of the biological roles, the predicted proteins were annotated with gene descriptions from Kyoto Encyclopedia of Genes and Genomes (KEGG; https://www.genome.jp/kegg/), Pfam (Mistry et al. 2020) and SwissProt (https://www. https://www.expasy.org/resources/uniprotkb-swiss-prot) databases. The TAIR (https://www.arabidopsis.org) hit IDs corresponding to the significant *B. napus* proteins (q-value < 0.05) were used for gene ontology analysis using PANTHER Overrepresentation Test in the Gene Ontology database (http://geneontology.org/docs/go-enrichment-analysis/). The parameters used for this analysis were statistical test type FISHER, correction method FDR (False Discovery Rate) and significance level of 0.05. Finally, the TAIR IDs were used for KEGG pathway analysis using the Database for Annotation, Visualization and Integrated Discovery (DAVID) (Jr et al., 2003). The KEGG enriched categories with the FDR-corrected p-value < 0.05 were considered for further analysis after Benjamini-Hochberg correction was applied to multiple testing (Benjamini and Hochberg, 1995).

5.3 Results

5.3.1 Disease symptom assessment post-inoculation with P. brassicae

We monitored the establishment of clubroot disease based on gall formation in roots at 7-, 14- and 21 dpi in both control and inoculated materials. No visible symptoms were observed in the roots of the control and inoculated plants of both CR and CS NILs at 7 dpi. At 14 dpi, swelling of the primary roots in the inoculated CS NILs was visible; however, such symptom was not visible in

the control as well as inoculated CR NILs. At 21 dpi, the roots of the inoculated CS plants were deformed with galls and the lateral roots became fragile. The aerial parts of the plants also showed symptoms of stunting, wilting, and yellowing. The control and CR inoculated plants did not show any sign of clubroot disease (Figure 5.1)

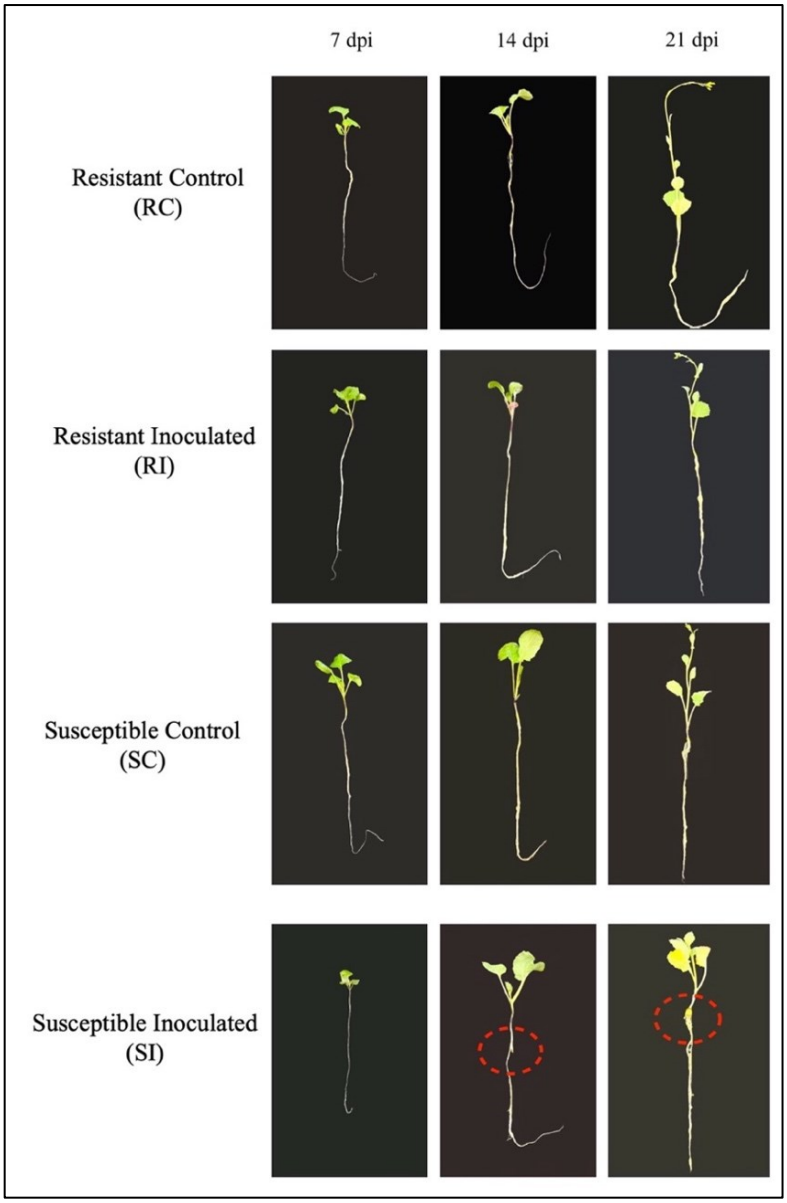


Figure 5.1 Root and shoot phenotypes of the clubroot resistant and susceptible near-isogenic lines (NILs) of *Brassica napus* at 7-, 14- and 21 days post inoculation (dpi) with *Plasmodiophora brassicae* pathotype 3H and control (without inoculation). The dotted red circles show gall formation in the roots.

5.3.2 Overview of the proteomic profile of root samples of B. napus

To study plant responses at the protein level due to pathogen attack, proteome analysis was performed on the infected and uninfected roots at 7-, 14-, and 21 dpi. A total of 10,998 proteins were identified across the genotypes, treatments, and time points (Table 5.1) of which 7,814 proteins were of interest in the context of the present study. Of the 7,814 proteins, 6,757 were significantly ($q < 0.05$) differentially accumulated (DAPs) across the time points and genotypes, where 480 and 2630 proteins were identified in the CS-NILs at 7 dpi and 21 dpi, respectively.

Table 5.1 Total count of peptides and proteins in roots from different comparisons (infected vs. control and resistant vs. susceptible) of *Brassica napus* near isogenic lines (NILs).

Comparison (Group1/Group2)	Total no. of peptides	Total no. of leading proteins
RI vs. RC	4635	2353
SI vs. SC	13,223	4274
RI vs. SI	13,291	4361
Total	31,149	10,988

Note: RI = inoculated resistant NILs; RC = uninoculated (control) resistant NILs; SI = inoculated susceptible NILs; SC = uninoculated (control) susceptible NILs.

5.3.3 Root proteome profile due to difference in genotype and pathogen infection

The root proteomes of the CR and CS NILs following inoculation with *P. brassicae* (RI and SI) and under control condition (RC and SC) were analysed. A total of 2353 proteins were differentially accumulated in the RI (inoculated CR) samples when compared with the RC (control CR) (RI vs. RC) of which 579, 977 and 797 DAPs were observed at 7-, 14- and 21 dpi, respectively. Among these, only 50 DAPs were detected at all three time points (Supplementary Table S5.1a, Figure 5.2a) where 12 proteins showed an increased abundance and four proteins showed a decreased abundance at all three time points (Supplementary Table S5.1a). From the SI vs. SC comparison, almost double the number of proteins (total 4274) were detected as compared to the

RI vs. RC comparison. In this case, the number of DAPs increased with the progression of infection, where the lowest number (480) was detected at 7 dpi and the highest number (2630) at 21 dpi (Figure 5.2b). Sixty-two DAPs were detected at all three time points in the SI-lines; among these, 11 proteins increased in abundance and 15 proteins decreased in abundance (Supplementary Table S5.1b).

Analysis of the proteomes of the infected CR and CS NILs (RI vs. SI) identified 531, 1271 and 2559 DAPs at 7-, 14- and 21 dpi, respectively; an increasing number of DAPs with the progression of pathogen infection was evident in this case (Figure 5.2c). Among these, 96 proteins were detected at all three time points, where 31 showed an increased abundance and 13 showed a decreased abundance (Supplementary Table S5.1c). Frequency distribution of the proteins, either increasing or decreasing in abundance, from all the above-mentioned comparisons at 7-, 14- and 21 dpi is represented in Figure 5.2d.

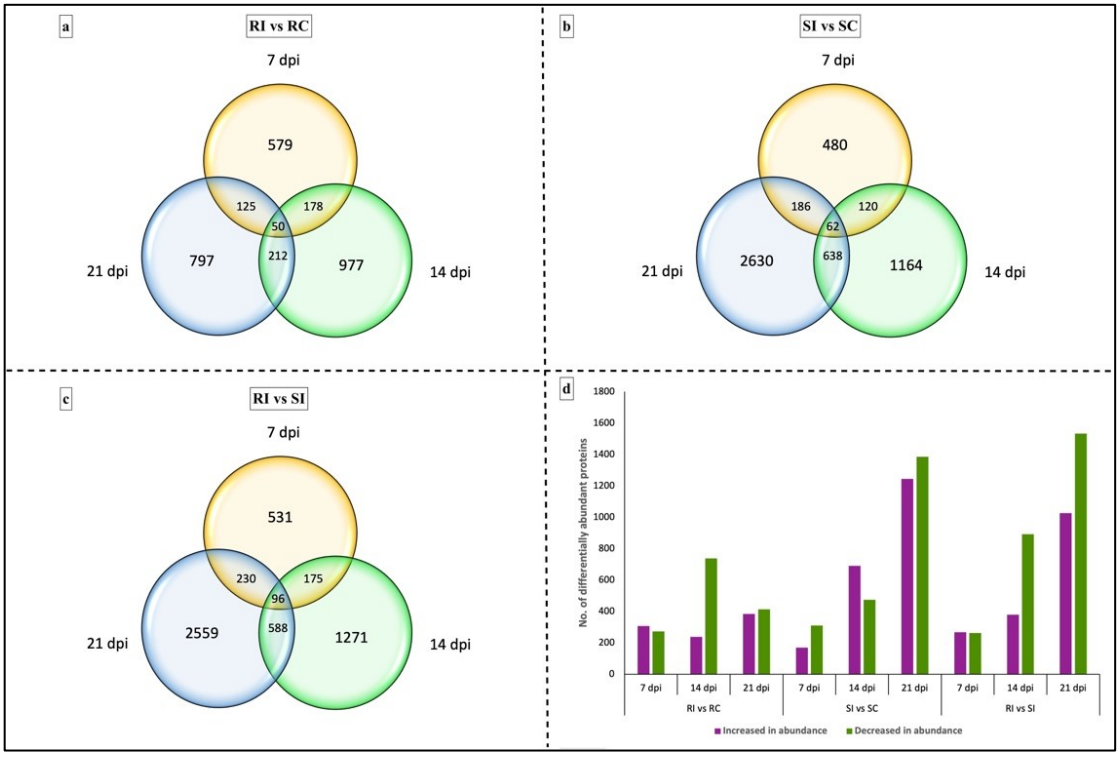


Figure 5.2 Differentially accumulated proteins (DAPs) in roots of the clubroot resistant (CR) and susceptible (CS) *Brassica napus* near-isogenic lines (NILs) inoculated with *Plasmodiophora*

brassicae pathotype 3H. RI = inoculated resistant NILs; RC = uninoculated (control) resistant NILs; SI = inoculated susceptible NILs; SC = uninoculated (control) susceptible NILs; dpi = days post inoculation.

5.3.4 Temporal changes in root proteome of the CR- and CS-NILs in response to pathogen infection (RI vs. RC and SI vs. SC)

Seven days post inoculation

At this stage, 45 and 34 DAPs showed an increased abundance while 39 and 50 DAPs showed a decreased abundance in the CR- and CS-NILs, respectively, in response to pathogen infection (Figure 5.3A). Among these, 23 DAPs showed an increased abundance and 28 DAPs showed decreased abundance in both CR- and CS-NILs (Figure 5.3A); these types of proteins hereafter will be referred to as "shared proteins" (Supplementary Table S5.2a). We also identified the proteins that were showing a differential accumulation pattern in the CR- and CS-NILs, i.e., either decreased in CR but increased in CS NILs, or vice-versa. This type of proteins hereafter will be referred to as "contrasting proteins". A total of 33 (22 + 11) DAPs of this type were identified in the CR- and CS-NILs (Supplementary Table S5.2a).

Among the shared proteins, the receptor-like protein kinase At3g14840 (BnaA01T0264300WE) and WRKY16 (BnaC07T0234800WE) involved in the perception of pathogenesis-related signals showed an increased abundance in both CR- and CS-NILs. On the other hand, proteins homologous to GEM-like protein 5 (BnaA02T0047200WE), Hsp70 protein 5 (BnaC08T0191000WE) and PEP carboxylase 3 (BnaC03T0502800WE) showed decreased abundance in both inoculated CR- and CS-NILs as compared to their controls. Among the contrasting proteins, the DAPs that increased in abundance in CR-NILs but decreased in the CS-NILs included the stress-related proteins like Dehydrin ERD10 (BnaA07T0126200WE), peroxygenase 3, Receptor-like kinase At1g06840 (BnaC05T0048500WE) and proteins involved

in cell wall modifications such as pectin acetylesterase 7 (PAE7, BnaA01T0022800WE). Other proteins such as pectinesterase inhibitor 61 (PME61, BnaA02T0139800WE) and senescence-specific cysteine protease SAG12 (BnaA02T0292500WE) were decreased in CR-NILs but increased in the CS-NILs in response to infection by *P. brassicae* (Supplementary Table S5.2a).

14 days post inoculation

When comparing the proteome profile at 14 dpi, we found 67 and 147 DAPs showing an increased abundance and 180 and 100 DAPs showing a decreased abundance in the CR- and CS-NILs, respectively (Figure 5.3B). Among these, 117 were shared proteins of which 42 showed an increased abundance and 75 showed a decreased abundance in both CR- and CS-NILs. Among the 130 contrasting proteins, 25 showed an increased abundance in the CR-NILs but decreased abundance in the CS-NILs, while 105 DAPs showed a decreased abundance in the CR-NILs but an increased abundance in the CS-NILs (Figure 5.3B; Supplementary Table S5.2b).

Among the shared proteins, the BnaC06T0009500WE, which is homologous to Thioredoxin H5, BnaC03T0109900WE (Basic endochitinase CHB4) and BnaA09T0013100WE (Chitinase 10) increased in abundance in both the CR- and CS-NILs. Among the contrasting proteins, the proteins homologous to ethylene responsive transcription factor 1A (ERF 1A, BnaA01T0006800WE), probable calcium-binding protein CML49 (BnaA01T0296000WE), peroxidase 44 (PER44, BnaA03T0495200WE) and peroxidase 17 (PER17, BnaA04T0138000WE) showed an increased abundance in the CR-NILs but a decreased abundance in the CS-NILs. Contrary to this, the proteins involved in cell wall modifications such as pectate lyase 5 (BnaA02T0167900WE) and xyloglucan-6-xylosyltransferase 2 (XXT2, BnaA02T0256900WE) as well as protein homologous to Respiratory burst oxidase homolog D

(RBOHD, BnaA02T0356200WE) showed a decreased abundance in the CR-NILs, however, an increased abundance in the CS-NILs.

21 days post inoculation

At 21 dpi, we identified 155 and 175 DAPs showing an increased abundance and 152 and 132 DAPs showing a decreased abundance in the CR- and CS-NILs, respectively (Figure 5.3C). Among the shared proteins, 121 DAPs increased, and 98 DAPs decreased in abundance in both CR- and CS-NILs (Figure 5.3C; Supplementary Table S5.2c). Among the 88 contrasting DAPs, 34 increased in abundance in CR-NILs but decreased in the CS-NILs, while 54 decreased in abundance in CR-NILs but increased in the CS-NILs (Figure 5.3C; Supplementary Table S5.2c).

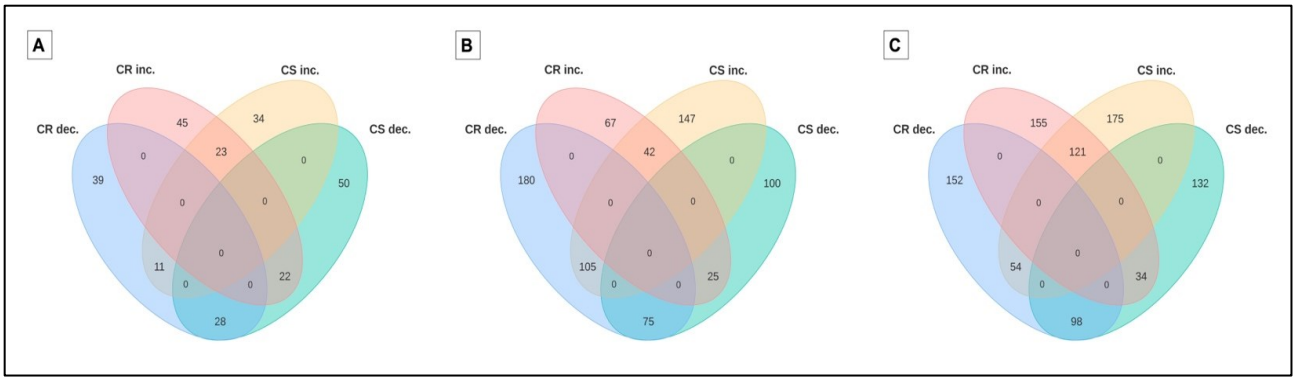


Figure 5.3 Differentially accumulated proteins (DAPs) in roots of the control and inoculated resistant and susceptible *Brassica napus* near-isogenic lines (NILs). The numbers indicate the number of proteins showing significant change in abundance (q value < 0.05). (A) 7dpi, (B) 14dpi, and (C) 21dpi; CR: clubroot resistant NILs, CS: clubroot susceptible NILs; inc. and dec. represent increase and decrease, respectively.

Among the shared proteins, the stress related proteins such as Universal stress protein A-like (BnaC06T0175200WE), Thioredoxin H5 (TRX5, BnaC06T0009500WE) and peroxidase C2 (PRXC2, BnaA09T0252100WE) were increased in abundance in both CR- and CS-NILs. In the pool of contrasting proteins, the stress-related proteins homologous to pathogenesis-related MLP-like protein 34 (BnaC06T0320700WE), cysteine-rich secretory protein 9 (CRRSP9,

BnaC04T0245900WE), chitin recognition protein endochitinase CH25 (BnaC03T0326700WE), receptor like kinase At1g51890 (BnaA06T0020600WE), major intrinsic protein Aquaporin TIP1-1 (BnaA04T0224500WE) as well as the calcium-binding protein CML49 (BnaA01T0296000WE) were increased in the CR-NILs but decreased in the CS-NILs.

5.3.5 Temporal changes in root proteome in the CR- and CS-NILs due to infection by P. brassicae (RI vs. SI)

In the RI vs. SI comparison, 96 proteins were classified as shared proteins, where 31 and 13 proteins showed an increased and decreased abundance respectively in the CR-NILs. The proteins related to cell wall modifications such as Xyloglucan endotransglucosylase/hydrolase protein 22 (BnaA02T0118100WE) and expansin-A15 (BnaA02T0324300WE), putative terpenoid synthase 7 (BnaA08T0101300WE) and defensin-like protein 1 (BnaC02T0415700WE) were increased in abundance in the CR-NILs at 7-, 14- and 21 dpi (Supplementary Table S5.3). Among the 13 proteins which showed a decreased abundance in the CR-NILs at all three time points, a protein homologous to Retinoblastoma-related protein 1 (BnaA05T0375100WE) and Tricyclene synthase (*TPS10*, BnaC03T0219500WE) were found.

In the pool of contrasting 52 proteins (Supplementary Table S5.3), some showed a decreased abundance at 7 dpi, however an increased abundance at 14- and 21 dpi. This included the proteins involved in glucosinolate biosynthesis [probable 2-oxoglutarate-dependent dioxygenase AOP1 (BnaA01T0266700WE)], calcium-mediated signalling [protein RALF-like 22 (BnaA01T0304800WE)], resistance to infection by a pathogen [RPM1-interacting protein 4 (BnaA03T0318600WE)], LRR-RLK (BnaA05T0167600WE), and hypersensitive-induced response protein 3 (BnaC05T0548200WE) (Supplementary Table S5.3).

5.3.6 GO enrichment for differentially accumulated proteins (DAPs)

To identify the putative proteins associated with the resistance mechanism against *P. brassicae*, we performed the GO enrichment analysis on the DAPs from RI vs. RC, SI vs. SC, and RI vs. SI comparisons. In case of RI vs. RC, 404 GO terms were significantly enriched (FDR < 0.05) by 21 dpi; among these 257 terms belonged to Biological Process (BP) category, 43 to Molecular Function (MF) category and 104 to Cellular Components (CC) category (Supplementary Table S5.4a). The top 20 significant GO terms in each of the BP and MF category included response to external stimulus (GO:0009605), biological process involved in interspecies interaction between organisms (GO:0044419), response to other organism (GO:0051707), response to external biotic stimulus (GO:0043207), response to biotic stimulus (GO:0009607), oxidoreductase activity (GO:0016491), antioxidant activity (GO:0016209), peroxidase activity (GO:0004601) and many others.

In the SI vs. SC comparison, 363, 107 and 94 GO terms belonging, respectively, to the BP, MF and CC categories, were significantly (FDR< 0.05) enriched in the CS-NILs by 21 dpi (Supplementary Table S5.4b). The highest level GO terms in the BP category included organic substance metabolic process (GO:0071704), primary metabolic process (GO:0044238), amino acid metabolic process (GO:0006520) and others. The top significant GO terms of the MF category included oxidoreductase activity (GO:0016491), ion binding (GO:0043167), cation binding (GO:0043169), and lyase activity (GO:0016829).

While comparing the DAPs from RI vs. SI, a total of 685 GO terms (419, 134 and 132 related to BP, MF and CC category, respectively) were significantly enriched (FDR < 0.05) as the infection progressed from 7- to 21 dpi (Supplementary Table S5.4c). The top significant (FDR < 0.05) terms of the BP category included the organic substance metabolic process (GO:0071704), primary metabolic process (GO:0044238), amino acid metabolic process (GO:0006520), and of

the MF category included oxidoreductase activity (GO:0016491), ion binding (GO:0043167), cation binding (GO:0043169) and lyase activity (GO:0016829); these GO terms were also enriched in the SI vs. SC. The top 20 significant GO terms from each of the category represented in Figure 5.4.

We also analysed the unique GO terms belonging to the BP and MF categories from all three comparisons and identified 68, 78 and 408 terms from the RI vs. RC, SI vs. SC and RI vs. SI comparisons, respectively which are discussed below (Supplementary Table S5.5a, S5.5b and S5.5c).

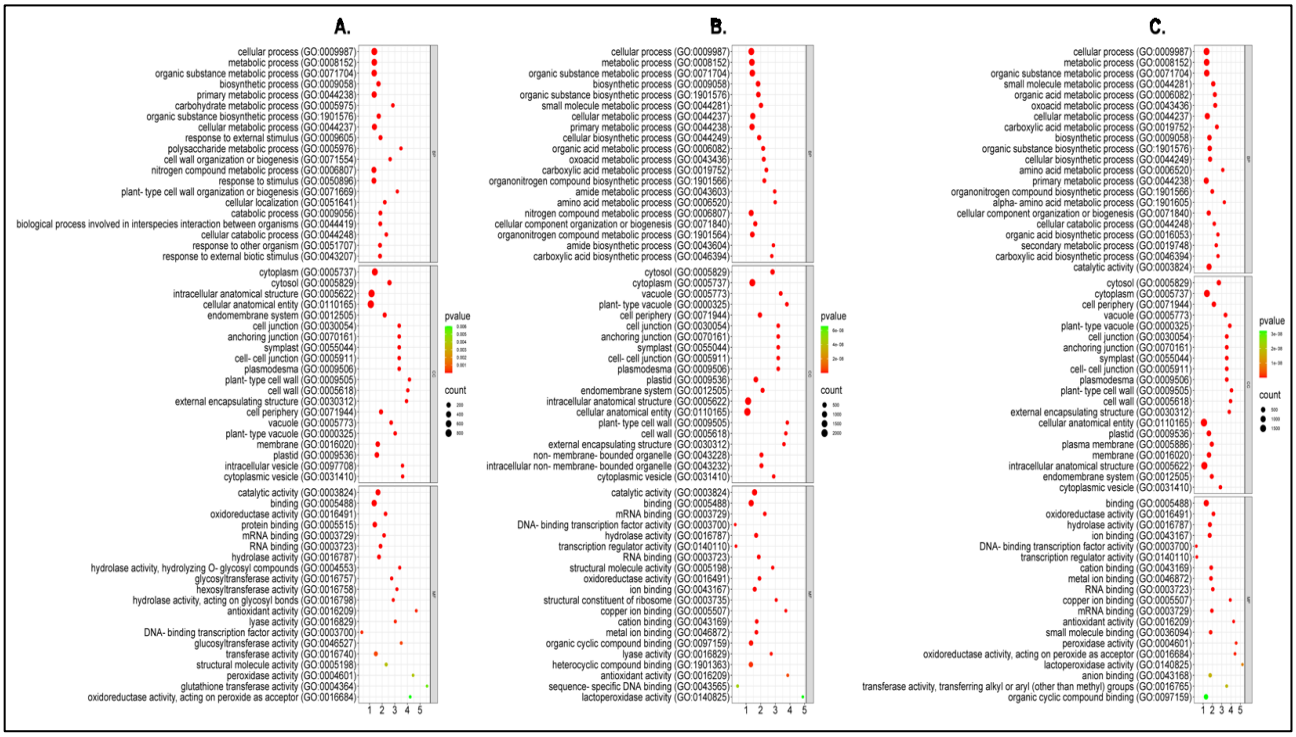


Figure 5.4 The top 20 significantly (FDR<0.05) enriched gene ontology (GO) terms of each of the biological process (BP), molecular function (MF) and cellular components (CC) category for the differentially accumulated proteins (DAPs) identified from RI vs. RC (A), SI vs. SC (B) and RI vs. SI (C) comparisons in roots of *Brassica napus* near isogenic lines (NILs) at 7- , 14-, and 21 days post inoculation (dpi) with *Plasmodiophora brassicae* pathotype 3H. RI = inoculated resistant NILs; RC = uninoculated (control) resistant NILs; SI = inoculated susceptible NILs; SC = uninoculated (control) susceptible NILs.

RI vs. RC

We identified several unique GO terms associated with plant-microbe interactions such as cellular response to stimulus (GO:0051716), response to nematode (GO:0009624), defense response to bacterium (GO:0042742), defense response to fungus (GO:0050832) and response to oomycetes (GO:0002239). The GO terms related to reactive oxygen species (ROS) included cellular response to reactive oxygen species (GO:0034614), cellular response to hypoxia (GO:0071456), oxidoreductase activity (GO:0016701, GO:0016667) were also identified; this indicates the potential involvement of ROS-mediated signalling mechanism in clubroot resistance. Based on RI vs. RC comparison, we also identified GO terms associated with flavonol biosynthetic and metabolic process (GO:0051555, GO:0051554) as well as flavonoid biosynthetic process (GO:0009813) (Supplementary Table S5.5a).

SI vs. SC

Among the 66 GO terms related to BP category, many stress-related GO terms such as cellular response to stress (GO:0033554), response to water deprivation (GO:0009414), cellular response to phosphate starvation (GO:0016036), defense response to symbiont (GO:0140546) were significantly enriched (FDR < 0.05). GO terms associated with auxin synthesis and transport, such as indole-containing compound biosynthetic process (GO:0042435), auxin biosynthetic and metabolic process (GO:0009851, GO:0009850), and tryptophan metabolic process (GO:0006568), were also significantly enriched indicating the possible role of auxin in disease development in the inoculated CS-NILs (Supplementary Table S5.5b). In case of MF category, the GO terms related to salicyclic acid binding (GO:901149) and protein kinase activity (GO:0004672) were significantly enriched (FDR < 0.05).

RI vs. SI

Among the 306 GO terms associated with BP category, several GO terms related to plant-microbe interactions were identified, and this included response to stimulus (GO:0050896), response to stress (GO:0006950), response to external stimulus (GO:0009605), biological process involved in interspecies interaction between organisms (GO:0044419), response to other organism (GO:0051707), response to biotic stimulus (GO:0009607), response to external biotic stimulus (GO:0043207), and response to wounding (GO:0009611). Several GO terms associated with secondary metabolism, such as indole glucosinolate metabolic process (GO:0042343), glucosinolate catabolic process (GO:0019762), phenylpropanoid metabolic process (GO:0009698), glucosinolate metabolic process (GO:0019760), secondary metabolite biosynthetic process (GO:0044550), were also significantly enriched (FDR < 0.05) in the inoculated CR-NILs (Supplementary Table S5.5c). Apart from these, the significant GO terms related to the regulation of hormone levels (GO:0010817) and response to reactive oxygen species (GO:0000302) indicates the potential role of phytohormone-mediated and ROS-mediated signalling in clubroot resistance in the CR-NILs. With respect to the MF category, the GO terms associated with ROS, for instance, oxidoreductase activity (GO:0016491), antioxidant activity (GO:0016209), and peroxidase activity (GO:0004601) were significantly enriched (FDR < 0.05). In addition to this, GO terms related to metal ion binding (GO:0046872), copper ion binding (GO:0005507), calcium ion binding (GO:0005509) were also enriched.

5.3.7 KEGG pathway enrichment of differentially accumulated proteins (DAPs)

The KEGG pathway enrichment analysis of data from the RI vs. RC comparison showed that the significantly enriched (FDR < 0.05) 377 DAPs belonged to three pathways, viz. protein processing in endoplasmic reticulum (44 DAPs), metabolic pathways (304 DAPs), and amino sugar and nucleotide sugar metabolism (29 DAPs), where the amino sugar and nucleotide sugar metabolism

could not be detected from other comparisons (Supplementary Table S5.6a). From SI vs. SC comparison, we identified 1460 significantly enriched (FDR < 0.05) DAPs and they belonged to 16 pathways; among these two pathways viz., ribosome and ribosome biogenesis in eukaryotes could be detected only from this comparison. In RI vs. SI comparison, 1318 DAPs were significantly enriched (FDR < 0.05) and they belonged to 20 pathways; among them, seven pathways, viz. sulfur, glutathione, pyruvate, and butanoate metabolism as well as starch and sucrose metabolism, valine, leucine and isoleucine degradation and citrate cycle were unique to the CR-NILs (Supplementary Table S5.6). The unique significant enriched pathways for the DAPs from the three comparisons at three time points are presented in Figure 5.5.

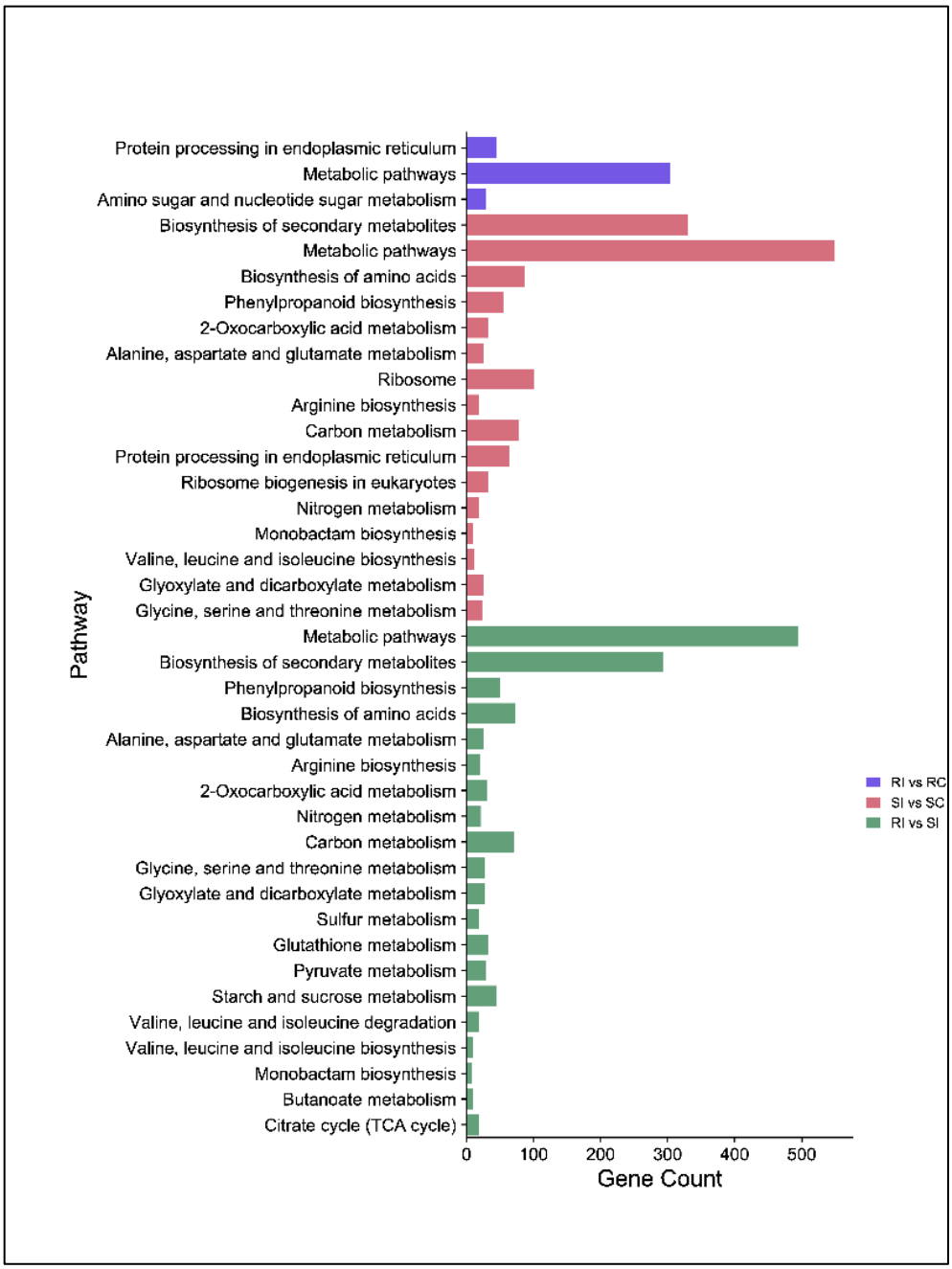


Figure 5.5 The unique significantly (FDR < 0.05) enriched KEGG pathways of the differentially accumulated proteins (DAPs) in roots of *Brassica napus* near isogenic lines (NILs) in response to *Plasmodiophora brassicae* pathotype 3H infection at 7-, 14- and 21 days post inoculation (dpi) from three comparisons of the resistant (R) and susceptible (S) NILs with or without inoculation. The y-axis denotes the enriched KEGG pathways, and the x-axis denotes the gene count. RC: resistant control; RI: resistant inoculated; SC: susceptible control; SI: susceptible inoculated.

5.3.8 Identification of putative differentially accumulated proteins (DAPs) to be involved in clubroot resistance

General stress-related

While comparing the root proteomes of the inoculated CR-NILs and inoculated CS-NILs with their respective controls (RI vs. RC and SI vs. SC), we found 23 significant DAPs associated with general stress responses (Supplementary Table S5.7a). Among these, nine DAPs increased in abundance in both CR- and CS-NILs, however five DAPs increased in abundance in the CR-NILs but decreased in abundance in the CS-NILs. These proteins were related to peroxygenase 3 (PXG3, BnaA04T0204000WE), dehydrin ERD10 (BnaA07T0126200WE), serine/arginine-rich splicing factor SR45a (BnaA08T0160100WE), transcription factor RF2b (BnaC03T0580600WE) and endochitinase (CH25, BnaC03T0326700WE).

In the RI vs. SI comparison, we found three DAPs, viz. patellin-1 (BnaC02T0252500WE), defensin-like protein 1 (BnaC02T0415700WE) and putative vacuolar protein sorting-associated protein 13A (BnaC08T0102600WE), to be associated with general stress responses. All these DAPs showed a consistently increased abundance at all three time points in the inoculated CR-NILs (Supplementary Table S5.7b).

Lipid metabolism

Seventeen proteins related to lipid biosynthesis or degradation were differentially accumulated ($q < 0.05$) in the CR- and CS-NILs upon infection as compared to their respective controls (RI vs. RC and SI vs. SC). Among these, three DAPs related to lipid degradation and homologous to GDSL esterase/lipase (BnaA06T0007800WE, BnaA08T0187900WE, BnaC06T0071100WE) were decreased in abundance in both CR- and CS-NILs (Supplementary Table S5.7a). DAPs showing an increased abundance in both CR- and CS-NILs were also identified; they were related

to lipid biosynthesis and degradation. Two DAPs homologous to acyl-activating enzyme 4 (AEE4, BnaA07T0363700WE) and fatty acid amide hydrolase (FAAH, BnaC03T0458600WE) showed an increased abundance in the inoculated CR-NILs but showed a decreased abundance in the inoculated CS-NILs at 21 dpi.

While comparing the inoculated CR-NILs and CS-NILs (RI vs. SI), we identified two significant DAPs ($q < 0.05$) related to fatty acid biosynthesis [12-oxophytodienoate reductase 1 (OPR1, BnaC02T0289200WE)] and fatty acid degradation [Enoyl-coA delta isomerase 3 (ECI3, BnaC08T0081300WE)]. The former DAP increased in abundance in the CR-NILs at all three time points, while the later one increased in abundance in the CR-NILs at 7 dpi but decreased at 14- and 21 dpi (Supplementary Table S5.7b).

Cell wall modifications

The RI vs. RC and SI vs. SC comparison of the root proteomes disclosed 34 DAPs ($q < 0.05$) significantly associated with cell wall modifications, and these included proteins related to biosynthesis of pectin, lignin and cellulose. Seven of these DAPs homologous to xyloglucan endotransglucosylase/hydrolase protein 22 (XTH22), caffeoylshikimate esterase (CSE), alpha-glucosidase 2 (GCS2), pectin acetylesterase 7 (PAE7), peroxidase C2 (PRXC2), probable galacturonosyltransferase 4 (GAUT4) and caffeic acid 3-O-methyltransferase (COMT1) showed an increased abundance in both inoculated CR- and CS-NILs. Another seven DAPs homologous to CSE, peroxidase 44 (PER44), PER17, shikimate-O-hydroxycinnamoyltransferase (HST), PER2 and endoglucanase 7 (KOR2) showed an increased abundance in the inoculated CR-NILs but decreased abundance in the CS-NILs (Supplementary Table S5.7a).

When comparing the inoculated CR-NILs with the inoculated CS-NILs (RI vs. SI), we found eight significant ($q < 0.05$) DAPs related to cell wall modifications. Five of these, viz.

XTH22, expansin-A15 (EXPA15), putative caffeoyl-coA O-methyltransferase (At1g67980), probable polygalacturonase (At1g80170) and 21kDa protein, showed an increased abundance in the inoculated CR-NILs as the infection progressed from 7- to 21 dpi (Supplementary Table S5.7b).

Phytohormone-mediated signalling

We identified 14 common DAPs related to auxin from the RI vs. RC and SI vs. SC comparisons. Among these, five DAPs showed a decreased abundance in the inoculated CR-NILs but an increased abundance in the inoculated CS-NILs. We also found five and four DAPs showing, respectively, an increasing and decreasing abundance in both CR- and CS-NILs (Supplementary Table S5.7a). Five DAPs associated with salicyclic acid regulation were identified; three of them (BnaA04T020400WE at 7 dpi, BnaA03T0338500WE at 14 dpi, BnaA05T0120800WE at 21 dpi) showed an increased abundance in the CR-NILs but a decreased abundance in the CS-NILs, and two DAPs (BnaA03T0047100WE and Bna05T0120800WE; both at 7 dpi) showed a decreased abundance in both CR- and CS-NILs. In case of jasmonic acid, we identified 12 DAPs; among them, three showed an increased abundance in both the CR- and CS-NILs, six showed a decreased abundance in both NILs, and the remaining three showed a contrasting pattern of accumulation in the CR- and CS-NILs. Two of the three contrasting DAPs for JA were homologous to 3-ketoacyl-coA thiolase 2, peroxisomal (PED1, BnaA05T0111200WE and BnaA04T0202400WE) decreased in the CR-NILs and increased in the CS-NILs. We found five DAPs associated with ethylene regulation. Four of these showed an increased accumulation in both CR- and CS-NILs, either at same level in both NILs or at a higher level in the CS-NILs (Supplementary Table S5.7a).

From the RI vs. SI comparison, we identified two DAPs associated with salicyclic acid and jasmonic acid regulation. The DAP related to SA regulation (BnaA03T0338500WE) increased in

abundance in the CR-NILs as the infection progressed from 7- to 21 dpi. Contrary to this, the JA related DAP homologous to Tricyclene synthase (TPS10, BnaC03T0219500WE) decreased in abundance in the CR-NILs at all three time points (Supplementary Table S5.7b).

Calcium-mediated signalling

Based on RI vs. RC and SI vs. SC comparisons, we identified seven DAPs associated with calcium-mediated signalling. Three of these DAPs homologous to calcium-binding protein CML49 (BnaA01T0296000WE), MLO-like protein 6 (MLO6, BnaA09T0149900WE) and glutamate decarboxylase 2 (GAD2, BnaC06T0348200WE) showed an increased abundance in the CR-NILs but a decreased abundance the CS-NILs (Supplementary Table S5.7a).

We also found three DAPs related to calcium-signalling from the RI vs. SI comparison. These three DAPs are related to vacuolar-sorting receptor 6 (BnaA09T0374500WE), RALF-like 22 (BnaA01T0304800WE) and RALF-like 1 (BnaA08T0299900WE) showed an increased abundance in the CR-NILs as the infection progressed from 7- to 21 dpi (Supplementary Table S5.7b).

Reactive oxygen species-mediated signalling

Based on RI vs. RC and SI vs. SC comparisons, we identified 24 significant DAPs ($q < 0.05$) related to ROS; 11 of them had the same trend of increase in abundance in both CR- and CS-NILs. Five of the 24 DAPs showed a contrasting trend for their abundance in the CR- and CS-NILs; they were homologous to peroxygenase 3 (PXG3, BnaA04T0204000WE), catalase 2 (CAT2, Bnascaffold286T0031400WE), glutathione-S-transferase U12 (GSTU12, BnaA02T0183300WE), peroxidase 2 (PER2, BnaA09T0647200WE) and L-gulonolactone oxidase 2 (GULLO2, BnaC04T0010900WE). We also found one protein homologous to CLPR4 (BnaA01T0100800WE) which consistently increased in abundance with the progression of disease

in the inoculated CR-NILs when compared with the inoculated CS-NILs (Supplementary Table S5.7b).

Glucosinolates

We identified five DAPs involved in glucosinolate biosynthetic pathway from both RI vs. RC and SI vs. SC comparisons at 21 dpi. Two of these had a contrasting pattern (increased in CR-NILs but decreased in CS-NILs) of accumulation and were related to Beta-glucosidase 19 (BGLU19, BnaC05T0118300WE) and nitrile specifier protein 2 (NSP2, BnaC05T0414100WE). The other three DAPs (BnaA04T0259100WE, BnaA09T0385200WE, BnaC08T0085400WE) were decreasing in both CR- and CS-NILs with the progression of infection; however, this decrease was more pronounced in the CS-NILs (Supplementary Table S5.7a).

Based on the RI vs. SI comparison, we identified three DAPs ($q < 0.05$) with increased accumulation in the inoculated CR-NILs at two or three time points. They were related to nitrile specifier protein 5 (NSP5, BnaC02T0465400WE), NSP2 (BnaC05T0414100WE) and 2-oxoglutarate-dependent dioxygenase AOP1 (BnaA01T0266700WE) (Supplementary Table S5.7b).

Resistance

We found six DAPs related to disease resistance protein from both RI vs. RC and SI vs. SC comparisons. Among these, four showed a contrasting pattern of accumulation with increased abundance in the CR-NILs but decreased abundance in the CS-NILs. These DAPs were related to (+)- neomenthol dehydrogenase SDR1 (BnaA10T0002900WE), ENHANCED DISEASE RESISTANCE 2-like (EDR2L, BnaA03T0055600WE) and protein YLS9 (BnaA08T0237900WE, BnaC05T0143800WE). We also found two DAPs homologous to SDR1 (BnaA10T0002900WE)

and RPM1-interacting protein 4 (RIN4, BnaA03T0318600WE) showing an increased abundance in the inoculated CR-NILs at 14 dpi and/or 21 dpi.

5.4 Discussion

Proteomics has been used to elucidate the molecular basis of the plant responses to various biotic stresses in different plant species, including *Brassica* based on functional annotations of the DAPs (Ali et al. 2014; Aroca et al. 2017; Lan et al. 2019; for review, see Liu et al. 2019; Moon et al. 2020). However, only a handful of studies have been conducted at proteome-level to understand the host resistance mechanism in *B. napus*-*P. brassicae* pathosystem (Cao et al. 2008; Adhikary et al. 2022a). To our knowledge, no study has been carried out using NILs to investigate the interaction between *P. brassicae* and *B. napus* at proteome level. In this research, we carried out a comparative proteomics study using homozygous clubroot resistant and susceptible *B. napus* NILs. These NILs carried a novel clubroot resistance introgressed from *B. rapa* turnip cv. Debra (accession ECD 01) (Kaur et al. 2022) and the resistant NILs confirmed carrying resistance to at least 15 *P. brassicae* pathotypes including the ones that recently evolved.

The phenotypic data confirmed that the inoculations with *P. brassicae* were successful as the inoculated CS-NILs developed the disease symptoms while the uninoculated control plants as well as the inoculated CR-NILs did not develop any symptoms. We found that the number of DAPs increased by 1.5-5.5 times as the infection progressed from 7- to 21 dpi and identified a total of 6757 significant ($q < 0.05$) DAPs. These DAPs provided important cues to the resistance response involving several pathways such as calcium-mediated signalling, phytohormone-mediated signalling, and reactive oxygen species. Our proteome analysis results align with the published reports in *B. napus*, *B. oleracea* and *B. rapa* (Song et al. 2016; Moon et al. 2020; Adhikary et al.

2022a). In addition to this, we found DAPs related to lipid and glucosinolate metabolism which may also participate in clubroot resistance in *B. napus*.

Perception of signal

The transmembrane signalling proteins are essential for perception of the pathogen by the host plant leading to the activation of innate immune host-responses (for review, see Couto and Zipfel 2016; for review, see Kobe and Kajava 2001). In the current study, we found increased abundance of four LRR-containing proteins in the inoculated CR-NILs. This included LRR receptor-like serine/threonine-protein kinases At3g14840 (BnaA01T0264300WE), At1g06840 (BnaC05T0048500WE), At1g51890 (BnaA06T0020600WE) and At1g51860 (BnaA05T0167600WE) having the log2 fold change ranging from 0.7 to 1.6 (Supplementary Table 5.7a and 5.7b).

Following the perception of signal, the ubiquitination system and Ca^{2+} ion mediated signalling events elicit the cellular responses, including oxidative burst (for review, see Trujillo and Shirasu 2010; for review, see Zhang et al. 2014). According to Lu et al. (2011) and Gohre et al. (2008), ubiquitination of pattern recognition receptor FLS2 in *Arabidopsis* leads to the activation of signalling pathways. Song et al. (2016) also reported the regulatory role of ubiquitin-dependent proteins in defense response against *P. brassicae* in *B. rapa*. In the present study, we found three ubiquitin-dependent proteins viz., 26S proteasome non-ATPase regulatory subunit 2 homolog B (RPN1B, BnaA01T0006100WE), E3 ubiquitin protein ligase (RGLG2, BnaA07T0290200WE) and ubiquitin carboxyl-terminal hydrolase 12 (UBP12, BnaC02T0487200WE) which increased in abundance in the CR-NILs (Supplementary Table 5.7a and 5.7b). This suggests that the ubiquitination-mediated signalling pathways play a role in the *P. brassicae-B. napus* pathosystem.

Calcium-mediated signalling

With respect to Ca-mediated signalling, we found an increased abundance of CML49 (BnaA01T0296000WE) at 21 dpi (Supplementary Table 5.7a), and an increased accumulation of RALFL22 (BnaA01T0304800WE) and RALF1 (BnaA08T0299900WE) at two out of the three time points (Supplementary Table 5.7b) in the CR-NILs. RALF1 and RALF22 interact with the receptor kinase FERONIA (FER) localized in the plasma membrane (Stegmann et al. 2017; Zhao et al. 2018). The interaction of RALF22 with FER results in increased salt tolerance in *A. thaliana* (Zhao et al. 2018) and the interaction of RALF1 with FER induces the calcium influx into the cell cytoplasm leading to the calcium-mediated signalling (Campos et al. 2018). Another protein IQD1 (BnaA01T0298100WE) was shared between the CR-NILs and CS-NILs; however, it's abundance increased in the CR-NILs and decreased in the CS-NILs (Supplementary Table 5.7a). IQD1 is a calmodulin binding nuclear protein which positively regulates glucosinolate accumulation, thus responding against biotic stress (Levy et al. 2005).

Reactive Oxygen Species (ROS)

The production of ROS is one of the earliest responses from the host cell following invasion by pathogen (for review, see Torres et al. 2006). We found an increased abundance of three peroxidases (BnaA03T0495200WE, BnaA04T0138000WE, BnaA09T0647200WE) in the inoculated CR-NILs but a decreased abundance in the CS-NILs (Supplementary Table 5.7a). It is also known that an excessive production of ROS can damage the host cell. The accumulation of ROS can be balanced by the ROS-scavenging enzymes such as catalase (CAT), superoxide dismutase (SOD) and glutathione-S-transferase (GST) (for review, see Liu et al. 2021). In this study, we also found an increased abundance of two proteins related to catalase-2 (Bnascaffold 286T0031400WE) and GST U12 (BnaA02T0183300WE) in the CR-NILs but a decreased

abundance in the CS-NILs (Supplementary Table 5.7a). Based on this, it is highly likely that these three ROS-scavenging proteins play a role in maintaining ROS homeostasis in the host cells. The Clp protease system is also involved in maintaining ROS homeostasis and thus limiting the ROS-dependent cell damage. They are involved in biotic and abiotic stress tolerance by degradation of the misfolded or damaged proteins within the cell (for review, see Ali and Baek 2020). In this study, we found a consistent increased abundance of ATP-dependent Clp protease CLPR4 (BnaA01T0100800WE) from 7- to 21 dpi in the inoculated CR-NILs as compared to the inoculated CS-NILs (Supplementary Table 5.7b) supporting its possible role in clubroot resistance.

The CLO/peroxygenase genes have been reported to play various roles in plant development, abiotic and biotic stress responses (for review, see Hanano et al. 2023). Partridge and Murphy (2009) found an upregulation of RD20 (*Arabidopsis* isoform AtCLO3) in *Arabidopsis* in response to *L. maculans* infection causing blackleg disease. We also found an increased accumulation of probable peroxygenase 3/CLO3 (BnaA04T0204000WE) in the inoculated CR-NILs suggesting its possible role in clubroot resistance (Supplementary Table 5.7a). We also found an increased abundance of protein related to Thioredoxin H5 (AtTRX5, BnaC06T0009500WE) in both CR- and CS-NILs, however, the level was higher in the CR-NILs (Supplementary Table 5.7a). Laloi et al. (2004) reported an upregulation of this gene in *Arabidopsis* in response to wounding and interactions with the bacterial pathogen *Pseudomonas syringae*.

Phytohormone-mediated signalling

Auxin Response Factors (ARFs) bind to the auxin responsive elements (AuxRe) and positively regulate the auxin-mediated transcriptional response (for review, see Abel et al. 1996). The ARFs target the genes regulating the homeostasis of auxins. ARF7 is reported to regulate the auxin-dependent differential growth in the hypocotyls. We found a decreased abundance of *ARF7*

(BnaA02T0087300WE) in the CS-NILs which might be associated with the reduction of lateral root formation favouring the gall formation (Supplementary Table 5.7a). Similar results have also been reported by Jahn et al. (2013) in *A. thaliana* in response to *P. brassicae* infection. We also found an increased abundance of a protein orthologous to *A. thaliana* ILL2 (BnaA10T0122400WE) in the CR-NILs at 21 dpi (Supplementary Table 5.7a). The ILL2 (IAA-amino acid hydrolase ILR1-like 2) encodes IAA-amino acid hydrolase possessing a broad range of substrate specificity (LeClere et al. 2002) and functions in the inactivation or inhibition of the phytohormone auxin. This suggests that BnaA10T0122400WE played a role in auxin inhibition in the CR-NILs. It is also known that the direction and distribution of auxin in the roots is regulated by Auxin efflux carrier component (PIN) genes (Blilou et al. 2005). Robin et al. (2020) found an increased transcription of BrPIN1 in dividing cells of *B. rapa* roots during cortex infection by *P. brassicae* and gall formation. In the current study, we also found an increased abundance of the protein orthologous to AtPIN1 (BnaA07T0334800WE) in the CS-NILs but decreased abundance in the CR-NILs at 21 dpi (Supplementary Table 5.7a). Based on this, it is likely that this protein plays an essential role in gall formation in the susceptible *B. napus* plants.

We found a contrasting pattern of accumulation of the TGA6 related protein BnaA03T0338500WE in the CR- and CS-NILs. According to Zhou et al (2000), the TGA transcription factor family regulates the transcription of SA-responsive genes through interactions with non-expressor of pathogenesis related 1 (NPR1). Zhang et al. (2003) reported that the TGA6 is required for the induction of the pathogenesis related 1 (PR1) gene in the presence of SA, and also play a role in the positive regulation of the systemic acquired resistance. Therefore, the TGA6 related protein (BnaA03T0338500WE) that we identified in this study exhibiting a contrasting pattern of accumulation, with an increased accumulation in the CR-NILs and a decreased

accumulation in the CS-NILs (Supplementary Table 5.7a), would need further investigation to elucidate its role in clubroot resistance.

In a previous study (Yang et al. 2015), Tricyclene synthase 10 (TPS10) is reported to be involved in the jasmonate defense signalling pathway. Xu et al. (2018) found a higher accumulation of jasmonic acid in susceptible *B. napus* plants at 14- to 28 dpi. In the current study, we also found a decreased abundance of the protein related to TPS10 (BnaC03T0219500WE) in the CR-NILs at all three time points (Supplementary Table 5.7b). The Ethylene Responsive Factors (ERFs) can either be activated or repressed in the host in response to pathogen infection (for review, see Müller and Munné-Bosch, 2015). We found an increased abundance of ERF1A (BnaA01T0006800WE) at 14 dpi in the CR-NILs but a decreased abundance in the CS-NILs as compared to their respective controls (Supplementary Table 5.7a).

Resistance

The ENHANCED DISEASE RESISTANCE 2 (EDR2) limits the initiation of cell death in plant-pathogen interactions (Vorwerk et al. 2007). The hairpin-induced family protein (YLS9) has been reported to play role in defense mechanism against *S. sclerotiorum* in *B. napus* (Zhao et al. 2007). The gene (+)- neomenthol dehydrogenase (SDR1) play a role in basal resistance against pathogens (Choi et al. 2008; Yamauchi et al. 2011). In this study, we identified four putative proteins related to EDR2-like (BnaA03T0055600WE), YLS9 (BnaA08T0237900WE, BnaC05T0143800WE) and SDR1 (BnaA10T0002900WE) showing an increased abundance in the CR-NILs but a decreased abundance in the CS-NILs (Supplementary Table 5.7a) implying that they might be involved in *P. brassicae-B. napus* interaction. We also identified the protein orthologous to RIN4 which showed an increased abundance in the inoculated CR-NILs when compared with the inoculated CS-NILs as the infection progressed from 7- to 21 dpi (Supplementary Table 5.7b). The RIN4 (RPM1-

interacting protein 4) is known to target various bacterial effectors either directly or indirectly as well as activate the R-proteins (Wilton et al. 2010; Selote et al. 2013). Summanwar et al. (2020) also detected 13 lncRNAs regulating RIN4. Based on this, it can be inferred that the RIN4 is involved in innate immunity in *B. napus* against *P. brassicae*.

Glucosinolates

The hydrolysis products of glucosinolates are known to play various roles in plant defense responses against bacterial and fungal pathogens (Brader et al. 2006). The nitrile-specifier protein (NSP) having 30-45% sequence homology with the epithiospecifier proteins (ESPs) activate the myrosinase-catalyzed degradation of glucosinolates to nitriles (Kissen et al. 2009). In the current study, we identified the proteins orthologous to NSP2 (BnaC05T0414100WE) and NSP5 (BnaC02T0465400WE) showing an increased abundance in the inoculated CR-NILs (Supplementary Table 5.7a and 5.7b). We also found an increased abundance of another protein ortholgous to 2-oxoglutarate-dependent dioxygenase AOP1 (BnaA01T0266700WE) in the CR-NILs as the infection progressed from 7 to 21 dpi (Supplementary Table 5.7b). AOP1 is reported to be involved in the glucosinolate biosynthesis pathway (Kawai et al. 2014).

5.5 Conclusion

To our knowledge, this is the first proteome-level investigation into the *P. brassicae-B. napus* pathosystem with the use of NILs. Our results demonstrated the involvement of calcium-signalling pathway, ROS, and a potential crosstalk between them in mediating a response against *P. brassicae*. Evidence for the involvement of phytohormones, glucosinolates as well as other proteins related to resistance mechanisms was also obtained from this study. Thus, the findings from this research provide a deeper insight into the putative proteins to be involved in clubroot resistance, and this lay the foundation for functional validation of the genes and proteins involved

in clubroot resistance; the knowledge can ultimately be used in molecular breeding of clubroot resistant *B. napus* canola cultivars.

Chapter 6. Genome-wide identification and expression analysis of C2H2 zinc finger protein genes in *Brassica napus* reveals their role in regulation of stress tolerance.

6.1 Introduction

The genus *Brassica* includes 39 economically important vegetable and oilseed crop species (for review, see Warwick et al. 2006). Among these, *Brassica napus* (2*n* = 38, AACC) commonly known as canola or rapeseed, is an allopolyploid species that originated from *Brassica rapa* (2*n* = 20, AA) and *Brassica oleracea* (2*n* = 18, CC) through interspecific hybridization (U 1935). Other oilseed crops of the genus Brassica include *Brassica rapa* and *Brassica juncea*. Global production of Brassica oilseeds has been increasing steadily in the last years due to its increasing demand for use as a source of edible oil as well as for the production of biodiesel (Ge et al. 2017). Currently, it is the third most important source of vegetable oil in the world after palm and soybean oils (USDA 2022).

Among the Brassica oilseed crops, *B. napus* canola is the most important with respect to its acreage and production. The growth, development and productivity of this crop is adversely affected by various biotic and abiotic stresses. Among the biotic stresses, clubroot disease caused by *Plasmodiophora brassicae*, stem rot disease caused by *Sclerotinia sclerotiorum*, and blackleg disease caused by *Leptosphaeria maculans* are the major threats to canola production in many countries including Canada (Bradley et al. 2006; for review, see Dixon 2009; Zhang and Fernando 2018). Among the abiotic stresses, high temperatures, drought and salinity result in significant crop damage in canola (Koscielny et al. 2018; Shokri-Gharelo and Noparvar 2018; Wu and Ma 2021).

Throughout the evolution, plants have developed different morphological, physiological, and molecular mechanisms to counter the threats posed by biotic and abiotic stresses (Meng et al.

2018; Bernstein 2019; Kumari et al. 2019). These adaptive mechanisms include the complex regulatory processes involved in the induction and transduction of stress signals, the differential regulation of the functional genes at transcript level mediated by interaction of transcriptional factors (TFs) with the promoter regions, and the post-translational modification of proteins (Seo and Choi 2015; Zhu 2016; Han et al. 2019a). Several TF families have been reported to be involved in biotic and abiotic stress responses in plants among which six are known to be the most important, including basic leucine zipper (bZIP), APETALA2/Ethylene response factor (AP2/ERF), myeloblastosis (MYB)-related transcription factors, basic helix loop helix (bHLH), WRKY and NAC (NAM, ATAF and CUC) (Ng et al. 2018). Several studies have identified TF families involved in mediating abiotic stress responses in *B. napus*, including BnAP2/ERF (Du et al. 2016), R2R3-MYB (Chen et al. 2016), bZIP (Zhao et al. 2016) and NACs induced by *BnNAC2* and *BnNAC5* (Zhong et al. 2012). In case of biotic stress responses, Yang et al. (2009) identified 13 BnWRKYs involved in fungal pathogen defense response, and Wang et al. (2014) demonstrated the overexpression of *BnWRKY33* enhancing resistance to *S. sclerotiorum*. These studies, therefore, highlight the importance of the TFs in mediating plant responses to abiotic and biotic stresses.

Zinc finger proteins (ZFPs) constitute one of the largest and ubiquitous group of TFs in the plant kingdom accounting for about 15% of the total TFs in plant genomes (Kielbowicz-Matuk 2012). They exhibit a 3D 'finger-like' structure harboring 2 antiparallel β-sheets and an α-helix stabilized by a Zn^{2+} ion coordinating cysteine and/or histidine residues (Kielbowicz-Matuk 2012). Based on the number and order of these Zn^{2+} coordinated residues, the ZFPs are divided into various subclasses such as, C2H2 (TFIIIA), C2HC (retroviral nucleocapsid), C2HC5 (LIM domain), C2C2, C3H, C3HC4 (RING finger), C4 (GATA-1), C4HC3 (requiem), C6 (GAL4) and

C8 (Kielbowicz-Matuk et al. 2012; Lyu and Cao 2018; Han et al. 2021). Among these, C2H2-ZFPs, also called Kruppel-like ZFPs, is the largest group and has been of significant interest to researchers. This type of ZFP possesses two conserved cysteine and two histidine residues bound by a Zn^{2+} ion tetrahedrally and form a structure as C-X_{2-4}-C-X_3-F-X_5-L-X_2-H-X_{3-5}-H, where X can be any amino acid (Kubo et al. 1998). In comparison to other eukaryotes, the C2H2-ZFPs in the plant kingdom possess a long and diverse sequence between the tandem zinc fingers (Takatsuji 1998). Furthermore, the α-helix in this type of ZFP contains a conserved sequence, QALGGH, in the zinc-finger (ZF) domain (Kam et al. 2008) and are known as Q-type C2H2-ZFPs (Kubo et al. 1998). The first Q-type C2H2-ZFP gene, *EPF1*, was reported in Petunia (Takatsuji 1992) having the QALGGH conserved sequence; since then, a total of 21 Q-type ZFPs have been reported in this plant (Kubo et al. 1998). Subsequently, several C2H2-ZFPs have been identified in other plant genomes including 176 from *Arabidopsis* (Englbrecht et al. 2004), 189 from rice (Agarwal et al. 2007), 122 from wheat (Faraji et al. 2018), 321 from soybean (Yuan et al. 2018) and 301 from *B. rapa* (Alam et al. 2019).

Functional analysis of C2H2-ZFPs demonstrated their involvement in many biological processes, including plant growth and development, hormone signalling, and resistance to biotic and abiotic stresses (Kim et al. 2004; Luo et al. 2012; Sharma et al. 2021; Yu et al. 2021; Zheng et al. 2021). In *B. napus*, there are only a limited number of studies on C2H2-ZF genes. Tao et al. (2017) reported that the ZFP gene *BnLATE* is involved in silique shattering in *B. napus*, and Lawrence and Novak (2018) identified several Q-type C2H2-ZFPs in *B. rapa*, *B. oleracea* and *B. napus*. The RNA-sequencing data confirmed the expression of 68 of the 74 Q-type C2H2-ZF genes in *B. napus* (Lawrence and Novak 2018); however, their role in growth and development or stress tolerance has not yet been investigated.

In this study, we performed the genome-wide identification of different C2H2-ZF genes in *B. napus* and carried out a comprehensive analysis of 267 C2H2-ZF genes based on phylogenetic relationships, chromosomal location, gene structures, conserved domains, cis-acting elements in the promoter regions, and synteny relationship between the *A. thaliana* and *B. napus* C2H2-ZFs. We also investigated the expression of 20 C2H2-ZF genes in *B. napus* under biotic stresses, such as infection by *P. brassicae* and *S. sclerotiorum*, abiotic stresses, such as cold, salinity and drought, and hormonal treatments in order to explore their putative roles in mediating stress responses.

6.2 Materials and Methods

6.2.1 Analysis of the C2H2-ZF proteins and genes in *B. napus*

6.2.1.1 Genome-wide identification and classification of C2H2-ZF proteins (C2H2-ZFPs)

In order to identify the *B. napus* C2H2-ZF genes, all known *A. thaliana* C2H2 -ZF protein sequences were retrieved from the TAIR database (http://arabidopsis.org/) (Berardini et al. 2015) and used to query the *B. napus* genome with BLASTp (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins) (Altschul et al. 1990). The resulting candidate C2H2-ZFPs were analysed by Simple Modular Architecture Research Tool (SMART) (http://smart.embl-heidelberg.de) (Letunic et al. 2020) for domain analysis using default parameters to confirm the presence of the C2H2 domains. The ProtParam online tool (http://web.expasy.org/protparam/) (Gasteiger et al. 2005) was used to determine the physiochemical properties of these C2H2-ZFPs, including molecular weight, isoelectric point (pI), number of amino acids, instability index, aliphatic index, and total number of atoms.

The C2H2-ZFPs were classified following two approaches which are (1) based on specific amino acid residues in the plant ZF domains and distances between 2~9 C2H2-ZF domains as described by Alam et al. (2019), and (2) based on variation in the plant-specific conserved amino

acid sequence 'QALGGH' and distances between metal ligands (cysteine and histidine) (Zhang et al. 2016; Alam et al. 2019). Following the first approach, the C2H2-ZFPs were classified as Tandem ZFs, when two nearby C2H2-ZF domains were separated by less than 12 amino acids and this subset was designated as Bn-t1-SF, and as dispersed or isolated C2H2-ZFPs, when the C2H2-ZFPs carried a single and/or 2-4 C2H2 domains and every 2 nearby domains are separated by more than 11 amino acid residues. Following the second approach, the C2H2-ZFPs were classified as Q-type, M1-M5 type, Z-type and D-type. The Q-type C2H2-ZFP domain was defined by X_2-C-X_2-C-X_7-QALGGH-X_3-H, where X can be any amino acid. The M1-M5 type C2H2-ZFPs are those containing one to five changed amino acids in the sequence QALGGH and some changes in the spacing between cysteine and histidine residues. The Z-type domains are defined by more than 12 (Z1) and less than 12 (Z2) residues between the first histidine and second cysteine, respectively. The D-type C2H2-ZFP domain does not contain the second histidine in the domain. We further categorised the C2H2-ZFPs carrying a single ZF domain into subsets Bn-1i-Q-SF, Bn-1i-M-SF, Bn-1i-Z-SF and Bn-1i-D-SF. The ZFPs containing two Q-type, two M-type and two Z-type domains were named as Bn-2i-Q-SF, Bn-2i-M-SF and Bn-2i-Z-SF, respectively. Similarly, those containing 3, 4 or 5 ZFs were classified as Bn-3i-SF, Bn-4i-SF and Bn-5i-SF, respectively. Subcellular localization of the identified C2H2-ZFPs was predicted using WoLFPSORT (https://www.genscript.com/wolf-psort.html) (Horton et al. 2007) with default parameters.

6.2.1.2 Phylogenetic analysis of the B. napus C2H2-ZF proteins

The protein sequences of the C2H2-ZF genes were aligned with each other using ClustalW (https://www.genome.jp/tools-bin/clustalw). The aligned sequences were used to generate phylogenetic tree with the software Molecular Evolutionary Genetic Analysis (MEGA) v10.1 (Kumar et al. 2016), using the Neighbor-Joining (NJ) algorithm with the following parameters:

bootstrap analysis with 1000 replicates, substitution-Poisson model and pairwise deletion. The tree was rooted based on distance between the protein sequences (Saitou and Nei 1987).

6.2.1.3 Chromosomal localization of the B. napus C2H2-ZF genes

By using the above-mentioned *B. napus* C2H2-ZF proteins identified through BLASTp (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins) and SMART (http://smart.embl-heidelberg.de) (Letunic et al. 2020) analysis, the genomic and coding sequences of these proteins were retrieved from *Brassica* database (BRAD) (http://Brassicadb.cn/#/). The software program MapChart v.2.3 (Voorrips 2002) was used to map the *B. napus* C2H2-ZF genes on the *B. napus* chromosomes. Two or more *B. napus* C2H2-ZF genes were considered tandemly duplicated if they resided on the same chromosome within a 100 kb region. The segmental duplication event was considered when the sequences were located on the same chromosome as blocks or on different chromosomes, however, showing high sequence similarity.

6.2.1.4 Syntenic relationships between the B. napus and A. thaliana C2H2-ZF genes

To investigate the syntenic relationships of *B. napus* with *A. thaliana*, *B. rapa* and *B. oleracea*, we downloaded their sequences and annotations from JGI Genome portal (https://genome.jgi.doe.gov/) (Nordberg et al. 2014). The homologous genes between the genomes were identified using Blastp with a threshold e-value less than $1e^{-10}$. The co-linear relation was visualized as a synteny plot using the "One Step MCScanX" function in TBtools (https://doi.org/10.1016/j.molp.2020.06.009) (Chen et al. 2020). We used the Search Syntenic Gene tool of BRAD database (Cheng et al. 2012) to find the *A. thaliana* ortholog for the *B. napus* C2H2-ZF genes. For each of the C2H2-ZF genes, the name of the orthologous gene in *A. thaliana* and *B. napus*, their location on the tPCK (translocation Proto-Calepineal Karyotype) chromosomes as well as on the chromosome blocks LF (least fractionated), MF1 (medium fractionated) and MF2

(most fractionated) of the A and C subgenomes (Lysak et al. 2007; Mandakova and Lysak 2008; for review, see Cheng et al. 2012) were retrieved. The Mega v7.0 software was used to calculate the synonymous (K_s) and non-synonymous (K_a) substitution rates and the evolutionary constriction (K_a/K_s) of the C2H2-ZF genes (Kumar et al. 2016).

6.2.1.5 Structure and ontology analysis of the B. napus C2H2-ZF genes

Annotation of the conserved motifs of the C2H2-ZF genes was performed using Multiple Expectation Maximization for Motif Elucidation (MEME) v5.5.1 (https://meme-suite.org/meme/tools/meme) (Timothy et al. 2015) with the following parameters: maximum number of predicted motifs = 10, any number of repetitions, optimum motif width between 10-200 residues, and ε-value for motif retention ≤ 1×10^{-10}. The gene structure, i.e., intron/exon organisation of these genes was elucidated by aligning the genomic and coding sequences in the Gene Structure Display Server (GSDS 2.0, (https://gsds.gao-lab.org) (Hu et al. 2015). The function of the C2H2-ZF genes was predicted using Blast2Go (Conesa and Gotz 2008) and the GO terms were categorised into biological, molecular, and cellular functions.

6.2.1.6 Promoter analysis

Genomic sequences 2 kb upstream of the initiation codon of the *B. napus* C2H2-ZF genes were downloaded from BRAD. The putative promoter regions were analysed for the presence of *cis*-regulatory elements, including the ones for hormone signalling, biotic and abiotic stresses, using the PlantCARE database (http://bioinformatics.psb.ugent.be/webtools/plantcare/
html) (Lescot et al. 2002).

6.2.2 Expression analysis of the *B. napus* C2H2-ZF genes using Quantitative Reverse Transcription-Polymerase Chain Reaction (qRT-PCR)

6.2.2.1 Plant materials

To study the expression of C2H2-ZF genes following inoculation with *P. brassicae* pathotype 3A, we used 12 resistant and 12 susceptible *B. napus* doubled haploid (DH) lines derived from crosses involving clubroot resistant canola lines carrying resistance from rutabaga (*B. napus* var. *napobrassica*) cv. Polycross and clubroot susceptible canola lines. The details of the development of these 24 lines can be found in Wang et al. (2022) and Megha et al. (2022). A spring *B. napus* DH line DH12075 was used to study the expression of the C2H2-ZF genes following infection by *S. sclerotium*, and exposure to salinity, drought and cold stresses as well as following hormonal treatments. Seeds of this line were obtained from Agriculture and Agri-Food Canada, Saskatoon Research Centre, Saskatchewan, Canada. In addition, the winter *B. napus* canola cv. Mendel was used for cold stress; seeds of this cultivar were obtained from Norddeutsche Pflanzenzucht Hans-Georg Lembke KG (NPZ), Germany.

6.2.2.2 Biotic stresses

P. brassicae: The clubroot resistant (R) and susceptible (S) DH lines were grown in a greenhouse at the University of Alberta in 3 cm × 3 cm × 5 cm (L × W × D) cell trays filled with Sunshine® Professional Growing Mix (Sungrow Horticulture, Agawam, MA, USA 01001) at 20-22/15 °C day/night, 16 hours (h) photoperiod and light intensity of 450 µE (mV) m^2s^{-1} at plant level. *P. brassicae* pathotype 3A inoculum used in this study was prepared from preserved galls as described in Kaur et al. (2022). The concentration of the spore suspension was adjusted to 1×10^7 to 1×10^8 resting spores/mL. Resistant and susceptible seedlings (10 days after seeding) were inoculated with either 1 mL of spore suspension or, in the case of control samples, with water. Leaf and root samples were collected from the control and inoculated plants at 7-, 14- and 21 days post inoculation (dpi), frozen in liquid nitrogen and stored at -80 °C until use. The total number of

samples were: 2 genotypes (R and S) × 2 treatments (infected and control) × 3 time points (7, 14 and 21dai) × 3 biological replicates = 36.

S. sclerotiorum: DH12075 plants were initially grown in a greenhouse under the conditions mentioned above, and the 20-day old plants were transferred to a humidity chamber (relative humidity of 90%) 24 h prior to inoculation. The inoculum of *S. sclerotiorum* was prepared by growing its culture on potato dextrose agar medium (PDA; Becton Dickinson, Columbia, MD) for 3 days at 21 ± 2 °C. Following the technique described by Joshi et al. (2016), the true leaves of the plants were inoculated with mycelial plugs (5 mm diameter), while the control plants were inoculated with uninfected sterile PDA plugs. Leaf samples (two from each plant) were collected from both control and infected plants at 12, 24 and 48 h post inoculations (hpi), frozen in liquid nitrogen and stored at -80 °C. The experiment was carried out in three biological replications where each replication included 20 plants at each time point.

6.2.2.3 Abiotic stresses

Cold: 3-week-old plants of DH12075 and cv. Mendel were grown in a greenhouse at 20-22/15 °C day/night and 16 h photoperiod for three weeks and were transferred into a growth chamber at 4 °C and constant light intensity of 45-55 μmol m^{-2} s^{-1} at plant level. The 2nd and 3rd leaf samples were harvested at 1, 2 and 7 days after cold treatment, frozen in liquid nitrogen and stored at -80 °C until use. For each treatment, bulked leaves from five plants were used. The control leaf samples were collected from the plants seeded at the same time, however, kept in the greenhouse (Megha et al. 2018). The experiment was carried out in three replications.

Salinity and drought: DH12075 plants, three weeks after seeding in the above-mentioned greenhouse, were sprayed with either 125 mM NaCl solution (salinity stress) or 20% (w/v) polyethylene glycol (PEG 8000) (drought stress) as described in Megha et al. (2022). The 2nd and

3rd leaves of the treated and untreated (control) plants were harvested after 6, 24 and 48 h of the treatment, frozen in liquid nitrogen and stored at -80 °C until use. The experiment was replicated thrice and the leaf samples from 10 individual plants within each replication were pooled together.

6.2.2.4 Phytohormone treatments

The 3-week-old DH12075 plants, grown in the above-mentioned greenhouse, were sprayed with ABA (Abscisic acid), SA (Salicylic acid) and BAP (6-Benzylaminopurine) solutions following Yang et al. (2009). Briefly, for ABA, a 20 mM stock solution was prepared by dissolving it in absolute ethanol followed by diluting it with 0.1% (v/v) ethanol to a concentration of 50 μM. For SA, a 100 mM stock solution was prepared by dissolving SA in distilled water with a pH of 6.5 and then diluting with water to make 1 mM of working solution. Stock BAP solution (1 mM) was prepared by dissolving it in 1 M NaOH and then diluting with distilled water to 20 μM. The control plants were sprayed with 0.1% (v/v) ethanol for ABA, and with distilled water (pH of 6.5) for SA and BAP treatments. The 2nd and 3rd leaves of 10 plants were harvested as bulk for the treated and control plants at 6, 24 and 48 h after treatment, frozen immediately in liquid nitrogen and stored at -80 °C until use. The experiment was repeated thrice.

6.2.3 qRT-PCR analysis

RNA was extracted from the frozen experimental plant materials using Trizol reagent (Invitrogen, USA) following manufacturer's protocol. The RNA was treated with DNase I (Promega, Madison, WI) to remove the contaminating DNA. The quality and concentration of the RNA was determined using a NanoDrop ND-1000 spectrophotometer (Thermo Scientific, Waltham, MA, USA). The RNA samples with 260/280 nm ratio of 1.8-2.0 and 260/230 nm ratio ≥ 2.0 were used for qRT-PCR analysis. For this, cDNA was synthesized using 2 μg of RNA following the instructions of Thermo Scientific's RevertAid H Minus Reverse Transcriptase kit (Thermo Fisher Scientific,

Waltham, MA, USA). To study the potential functions of *B. napus* C2H2-ZF genes, we randomly selected 2-3 genes from each clade and by using the CDSs of these genes, the gene specific primers were designed using Primer Express v3.0.1 (Life Technologies, ON, Canada); sequence of these primers are listed in Supplementary Table S6.1. qRT-PCR was performed using ViiA 7 Real-Time PCR system (Applied Biosystems – Life Technologies, Waltham, MA, USA) where Ubiquitin-Conjugating Enzyme 9 (*UBC9*), a house keeping gene, was used as an endogenous control. Each qRT-PCR reaction included 5 µl of FASTSYBR Green Mix, 1 µl of cDNA (3:1 dilution), 2 µl of primer pair, and 2 µl of nuclease free water to make the final volume up to 10 µl. Two technical replicates of each biological replicate were analysed. Changes in the relative gene expression were calculated using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen 2001). One-way and two-way ANOVA analysis was carried out using GraphPad Prism version 9.0.0 (www.graphpad.com) and significant difference between relative expression levels of the genes was calculated using Tukey's HSD test with $p < 0.05$.

6.3 Results

6.3.1 Genome-wide identification and classification of the C2H2-ZF proteins (C2H2-ZFPs)

To identify the putative C2H2-ZFPs, we conducted a comprehensive BLASTp search against the *B. napus* genome database using the query sequences of *A. thaliana* and identified a total of 267 C2H2-ZF proteins after manual analysis using SMART database. The sequences of these 267 proteins were retrieved from BRAD and they were named as *BnaZF1* to *BnaZF267* (*B. napus* zinc finger genes) (Supplementary Table S6.2). Of the 267 C2H2-ZF proteins, 27.3% (73/267) were classified as tandem (Bn-tandem-SF) and 72.6% (194/267) as isolated (Bn-isolated-SF) C2H2-ZF proteins. Of the Bn-tandem-SF, 51 (69.9%) C2H2-ZFPs contained three C2H2 domains, 15 (20.5%) had four C2H2 domains, 5 (6.8%) had five C2H2 domains, 1 (1.4%) had six C2H2

domains, and 1 (1.4%) had nine C2H2 domains (Figure 6.1a). The Bn-isolated-SFs could be divided into 11 subsets based on the number and type of domains: Bn-1i-Q-SF (46, 17.2%), Bn-1i-M-SF (44, 16.5%), Bn-1i-Z-SF (7, 2.6%), Bn-1i-D-SF (1, 0.4%), Bn-2i-Q-SF (45, 16.9%), Bn-2i-M-SF (10, 3.7%), Bn-2i-D-SF (1, 0.4%), Bn-3i-SF (10, 3.7%), Bn-2i-Mix-SF (27, 10.1%), Bn-4i-SF (2, 0.7%) and Bn-5i-SF (1, 0.4%) (Figure 6.1b; Supplementary Table S6.3).

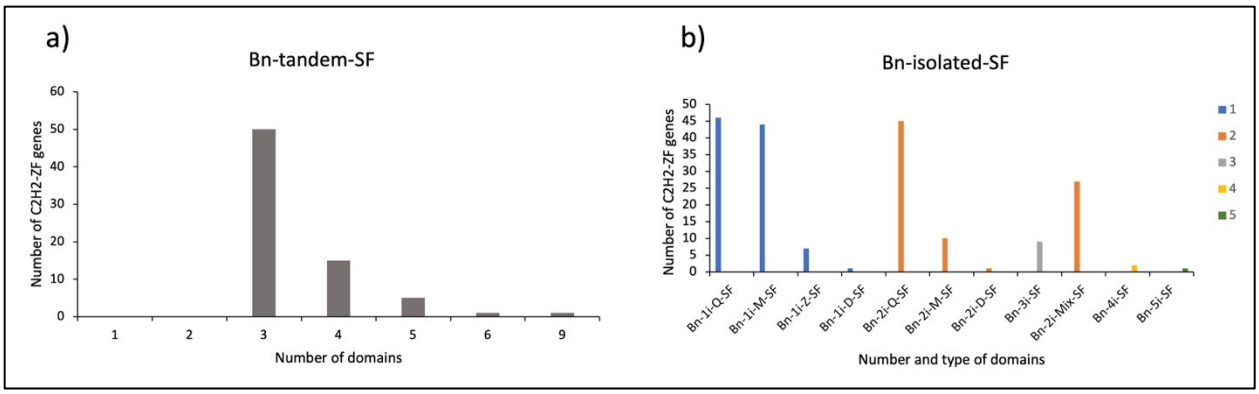


Figure 6.1 Classification of the C2H2-ZFP genes of the *Brassica napus* genome: a) Bn-tandem-ZF type C2H2-ZFs based on number of domains, and b) Bn-isolated-ZF type C2H2-ZFs based on number and types of domains. In case of (b), the different colors indicate the number of domains.

Physicochemical analysis indicated that the length of the proteins encoded by the 267 C2H2-ZF genes varied from 79 amino acids (AAs) (*BnaZF45*) to 1366 AAs (*BnaZF46*) with an average of 346.85 AAs. The theoretical molecular weights (MWs) ranged from 9,024.14 Da (*BnaZF45*) to 153,543.14 Da (*BnaZF46*), and isoelectric points (pI) ranged from 0.46 (*BnaZF176*) to 10.03 (*BnaZF34*) (Supplementary Table S6.2). Subcellular localization analysis using WolFPSORT predicted that 256 proteins were located in the nucleus, 5 in the cytosol, 4 in the mitochondria and 2 in the chloroplast (Supplementary Table S6.2).

6.3.2 Phylogenetic analysis of the *B. napus* C2H2-ZF proteins

Phylogenetic analysis of the 267 C2H2-ZF proteins was performed using Neighbor-Joining method with 1000 bootstrap replications to determine the evolutionary relationship between them. Based on sequence similarities and topology, the 267 C2H2-ZF proteins could be resolved into 5

clades on the phylogenetic tree with significant bootstrap values (Figure 6.2), which are named as Group I, II, III, IV and V with 26, 36, 38, 80 and 87 members, respectively.

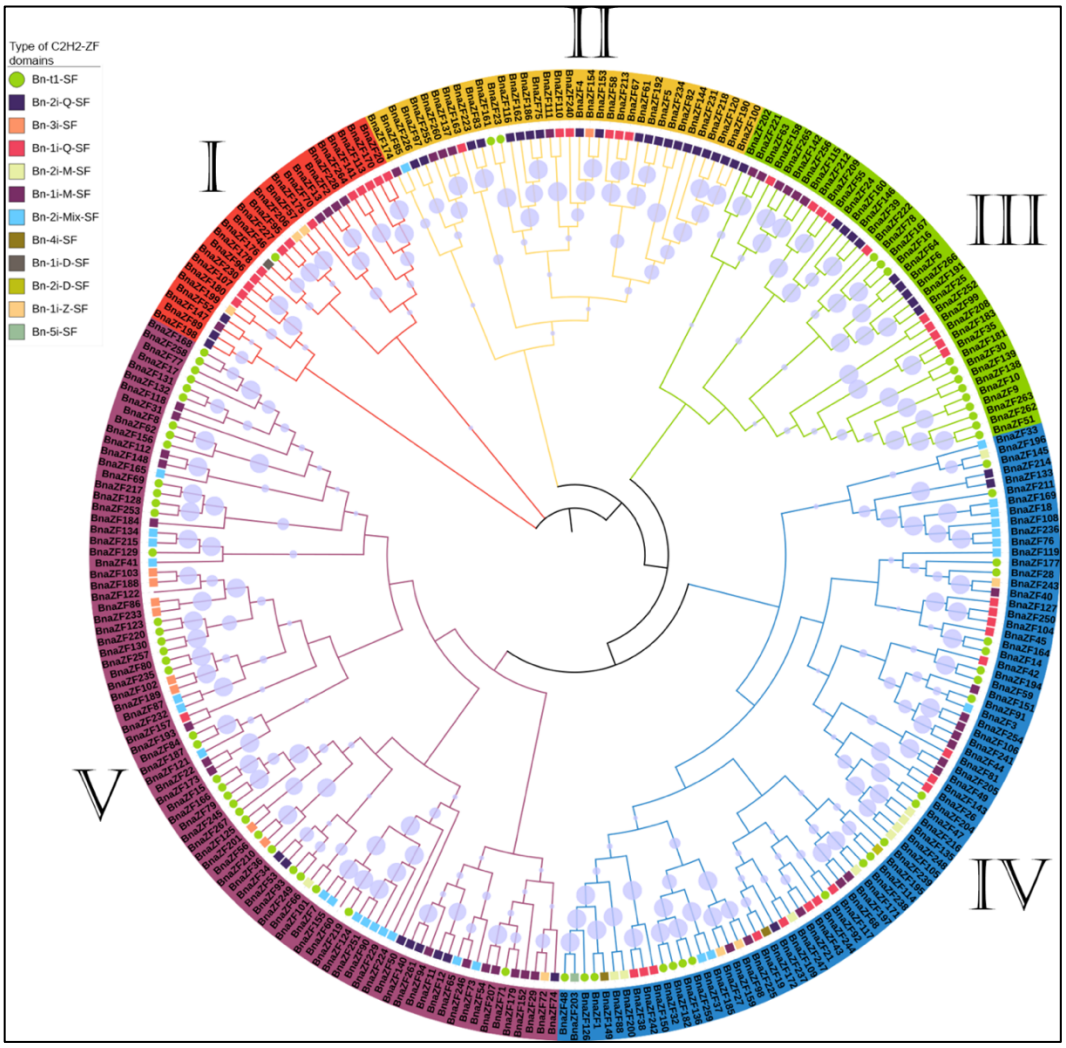


Figure 6.2 Phylogenetic analysis of the 267 *Brassica napus* C2H2-ZF proteins. The phylogenetic tree was generated by using MEGA7.0 with Neighbor-Joining (NJ) algorithm and a bootstrap analysis of 1,000 replicates. The purple circles on the branches indicate the percentage of bootstrap replications; bigger circle indicating the higher percentage of bootstrap values. The five groups (I-V) are indicated with different colors in the outermost circle. The different types of C2H2-ZF protein domains are shown at the tip of the phylogenetic tree.

6.3.3 Chromosomal localization and duplication of the *B. napus* C2H2-ZF genes

The 228 C2H2-ZF genes identified in this study were physically mapped to determine their distribution and location on *B. napus* chromosomes. Although the C2H2-ZF genes were detected

on all 19 chromosomes, their distribution was not uniform. Of the 19 chromosomes, the greatest number (22) of C2H2-ZF genes was located on chromosome A09 followed by 19 on A03, 17 on C04, 16 on C02, 15 on A10, 14 on each of A02, A06, C03 and C07, 12 on A04, 11 on each of A07 and C09, nine on each of C01 and C05, seven on A05, and six on each of A01, A08, C06 and C08 (Figure 6.3, Supplementary Figure S6.1). The exact location of the remaining 39 C2H2-ZF genes could not be determined and, therefore, they were not included in the MapChart dataset.

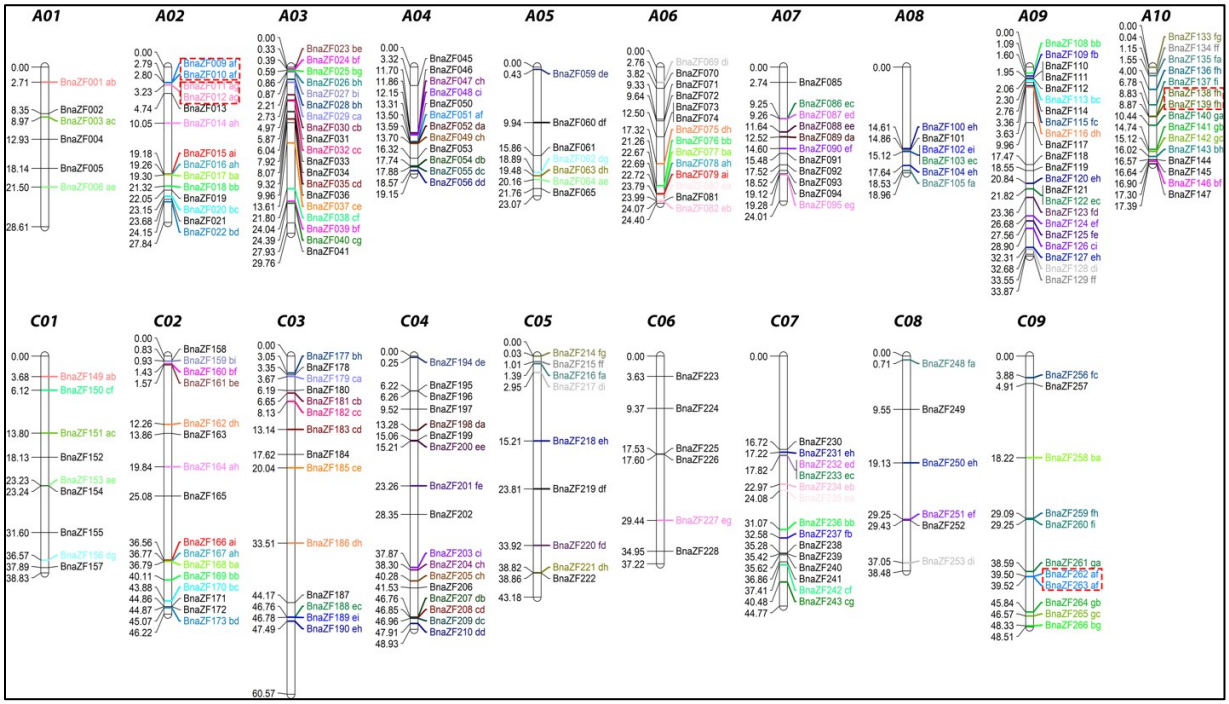


Figure 6.3 Distribution of C2H2-ZF genes in the *Brassica napus* genome. Chromosome numbers (*A01-A10* and *C01-C09*) are indicated above each chromosome. Gene names (*BnaZF1* to *BnaZF267*) and their physical positions (Mb) are indicated at the right and left side of the chromosomes, respectively. Duplicated genes on different chromosomes are represented by different colours and tandemly duplicated genes are outlined with red dashed boxes.

Analysis of the C2H2-ZF genes for gene duplication events was carried out to understand their role in genome expansion and reorganization of the *B. napus* genome. Based on sequence similarities and length in the aligned regions, we identified four tandem duplicated gene pairs

BnaZF9/BnaZF10, *BnaZF11/BnaZF12*, *BnaZF138/BnaZF139* and *BnaZF262/BnaZF263* on chromosomes A02, A02, A10 and C09, respectively (Figure 6.3). In addition, 62 segmental duplication events were detected across different chromosomes. These observations suggest that both inter- and intra-chromosomal duplication events may have contributed to the expansion of this gene family in *B. napus*.

6.3.4 Syntenic relationships between *B. napus* and *A. thaliana* C2H2-ZF genes

To investigate the evolutionary mechanisms of *B. napus* C2H2-ZF genes, we analysed the syntenic relationships of *B. napus* with three other genomes. A total of 1103 C2H2-ZF genes were syntenic with those in *A. thaliana* (236), *B. rapa* (547) and *B. oleracea* (320) (Supplementary Figure S6.2). The syntenic paralogs in the two Brassica sub-genomes A and C as well as their orthologs in *A. thaliana* were identified from BRAD database. We were able to establish the syntenic relationship for 234 (87.64%) of the 267 C2H2-ZF genes and, of the 234 genes, orthologs for 229 (87.64%) genes could be detected in *A. thaliana*. They appeared to have been derived from 20 ancestral blocks, and no orthologs could be detected for the remaining five genes (*BnaZF32*, *BnaZF136*, *BnaZF182*, *BnaZF225* and *BnaZF259* (Supplementary Table S6.4). Of the 234 genes, five genes (*BnaZF21, BnaZ45, BnaZ65, BnaZ91* and *BnaZ144*) were found only in the A sub-genome and two genes (*BnaZF152* and *BnaZ187*) were found only in the C sub-genome (Supplementary Table S6.4); all other genes could be found in both A and C sub-genomes either as a single copy or up to three copies. Thus, a total of 445 and 419 copies of the 234 C2H2-ZF genes were found in the A and C sub-genomes, respectively. The occurrence of the 234 C2H2-ZF genes in the LF, MF-1 and MF-2 chromosome blocks of the A and C sub-genomes were as follows: 167 (62.54 %) on LF, 153 (57.3 %) on MF-1 and 125 (46.8 %) on MF-2 of the A genome, and 171 (64.04 %) on LF, 142 (53.18 %) on MF-1 and 106 (39.7 %) on MF-2 of the C genome.

To understand the selective pressures on these *B. napus* C2H2-ZF genes, the Ka, Ks and Ka/Ks values were calculated for different C2H2-ZF gene pairs between *A. thaliana* and *B. napus* and are summarized in Supplementary Table S6.5. The Ka/Ks ratios varied from 0.05 to 2.93 with an average ratio of 0.68. Sixty-one orthologous gene pairs had Ka/Ks ratios >1 implying that these genes had experienced positive selection (beneficial alleles increasing in prevalence) while the other gene pairs had Ka/Ks ratios <1 suggesting that they had experienced negative selection (detrimental alleles eliminated) during evolution.

6.3.5 Structure and ontology analysis of the *B. napus* C2H2-ZF genes

Information on the organization of introns and exons in the *B. napus* C2H2-ZF genes was gleaned by comparing the full length CDSs and whole gene sequences (Supplementary Figure S6.3). The number of exons in the genes varied from 1 to 21 where 41.57% (111/267) of the genes contained only one exon while only 0.74% (2/267) (*BnaZF137* and *BnaZF260*) of the genes had 21 exons. The length of the genes also varied widely, where *BnaZF81* was the shortest (0.41 kb) and *BnaZF40* was the longest gene (9.2 kb). No significant correlation was found between the length of the genes and the number of exons in the genes. About 66% (176/267) of the *B. napus* C2H2-ZF genes had UTRs either at one or at both ends (Supplementary Figure S6.3)

The *B. napus* C2H2-ZF genes were further analyzed for the presence of conserved motifs using MEME Suite. A total of 10 conserved motifs were identified, of which motifs 1, 2, 3 and 9 represented Q, D, Z and M-type of C2H2-ZF domains, respectively (Supplementary Figure S6.4, Supplementary Table S6.6).

6.3.6 Promoter analysis of the *B. napus* C2H2-ZF genes

In order to gain insights into the transcriptional regulation of the *B. napus* C2H2-ZF genes under biotic and abiotic stresses, the promoter sequences (2 kb upstream region) of the genes were

analyzed using PlantCARE. Twelve *cis*-acting elements related to phytohormone or biotic/abiotic stress responses were identified and included the AuxRR-core, TGA element (responsive to auxins), GARE-motif and TATC-box (GA responsive), ABRE (responsive to abscisic acid), TGACG and CGTCA motif (responsive to MeJA), TCA-element (SA responsive), LTR (responsive to low temperature), MBS (MYB element for drought response), and TC-rich repeats and WUN-motif (defense and stress responsive) (Supplementary Table S6.7). *BnaZF192* possessed the maximum number of *cis* elements (25), including 14 elements responsive to ABA, 6 responsive to MeJA, 2 responsive to GAs, and 1 each for SA responsive, drought inducible and defense and stress responsive elements. In contrast, each of *BnaZF48*, *BnaZF161* and *BnaZF165* possessed only one *cis* element. Moreover, 43.07% (115/267) of the *B. napus* C2H2-ZF genes had TC-rich repeats and 41.19% (110/267) had WUN-motif implying their possible roles in pathogen defense, and 78.27% (209/267) of the *B. napus* C2H2-ZF genes had ABRE *cis*- element in their promoter region implying their possible roles in mediating responses to ABA.

6.3.7 Expression of the *B. napus* C2H2-ZF genes under biotic stresses

P. brassicae infection

To investigate the role of C2H2-ZF genes in biotic stress, we investigated the expression of 20 *B. napus* C2H2-ZF genes representing all the five clades and types of domains in response to *P. brassicae* infection at 7, 14 and 21 dai (days after inoculation). The histopathological analysis of the infected roots reveals the initial symptoms of infection on the secondary roots around 7 dai, presence of greater number of secondary plasmodia in susceptible roots at 14 dai and symptoms of hypertrophy and gall formation in susceptible roots and not in resistant roots at 21dai (Summanwar et al. 2020). This served as the basis for selection of these three time points for expression analysis. qRT-PCR results indicated that the expression of *BnaZF202* increased by

5.16-fold at 14 dai and the expression of *BnaZF1* increased by 4.09-fold at 21 dai, while the expression of *BnaZF94* increased (2.21- and 2.15-fold) at both 14 and 21 dai in the roots of the resistant lines (Figure 6.4). In the case of the susceptible lines, the expression of five genes, i.e., *BnaZF22*, *BnaZF139*, *BnaZF162*, *BnaZF164* and *BnaZF241*, was increased at 7, 14 or 21 dai in the root tissue (Figure 6.4), while the remaining genes showed no change in their expression patterns (Supplementary Figure S6.5). In the leaf tissue of the resistant lines, expression of the gene *BnaZF164* was significantly increased at 21 dai, while in the susceptible lines, expression of three genes (*BnaZF1*, *BnaZF94* and *BnaZF140*) was significantly increased at 7 or 14 dai ($p < 0.05$) (Figure 6.4B).

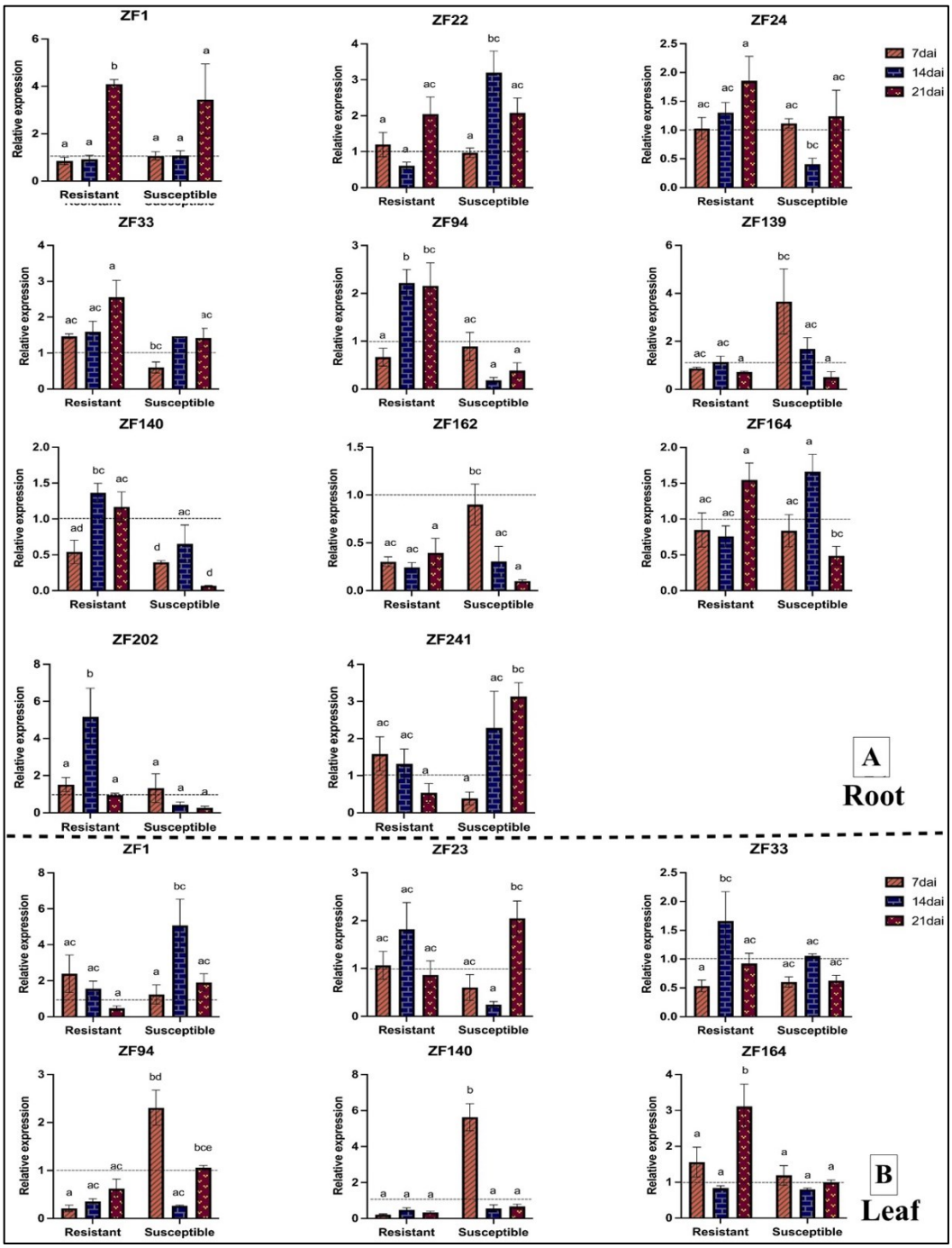


Figure 6.4 Expression of *Brassica napus* C2H2-ZF genes in response to *Plasmodiophora brassicae* infection as determined by qRT-PCR. (A) Expression of 11 genes in the roots, and (B) six genes in the leaves of clubroot resistant and susceptible lines (bulks). Bar and error bar indicates the mean of three biological replicates ± standard error (S.E.). Different letters indicate significant difference between the mean values (p<0.05, Tukey method). Expression level was normalized to the control for each time point and indicated by dotted line at expression level = 1.0.

S. sclerotiorum infection

In the case of *Sclerotinia* infection, the transcript of *BnaZF140* showed significantly (p <0.05) increased accumulation at 24 h and this further increased at 48 h as compared to control and 12 h samples (Figure 6.5). Similarly, *BnaZF170* showed an increased expression at 48 h (14.6-fold). In contrast, a decreased expression of *BnaZF65* and *BnaZF105* was observed with the progression of *Sclerotinia* infection. The expression of *BnaZF241* showed a significant (p <0.05) increase at 12 h (4.91-fold) followed by a decrease at 48 h (2.6-fold) (Figure 6.5). Of the total 20 C2H2 ZF genes analyzed by qRT-PCR, seven (*BnaZF1*, *BnaZF22*, *BnaZF23*, *BnaZF33*, *BnaZF90*, *BnaZF162*, *BnaZF211*) did not show any change in expression patterns following infection with *S. sclerotiorum* (Supplementary Figure S6.6).

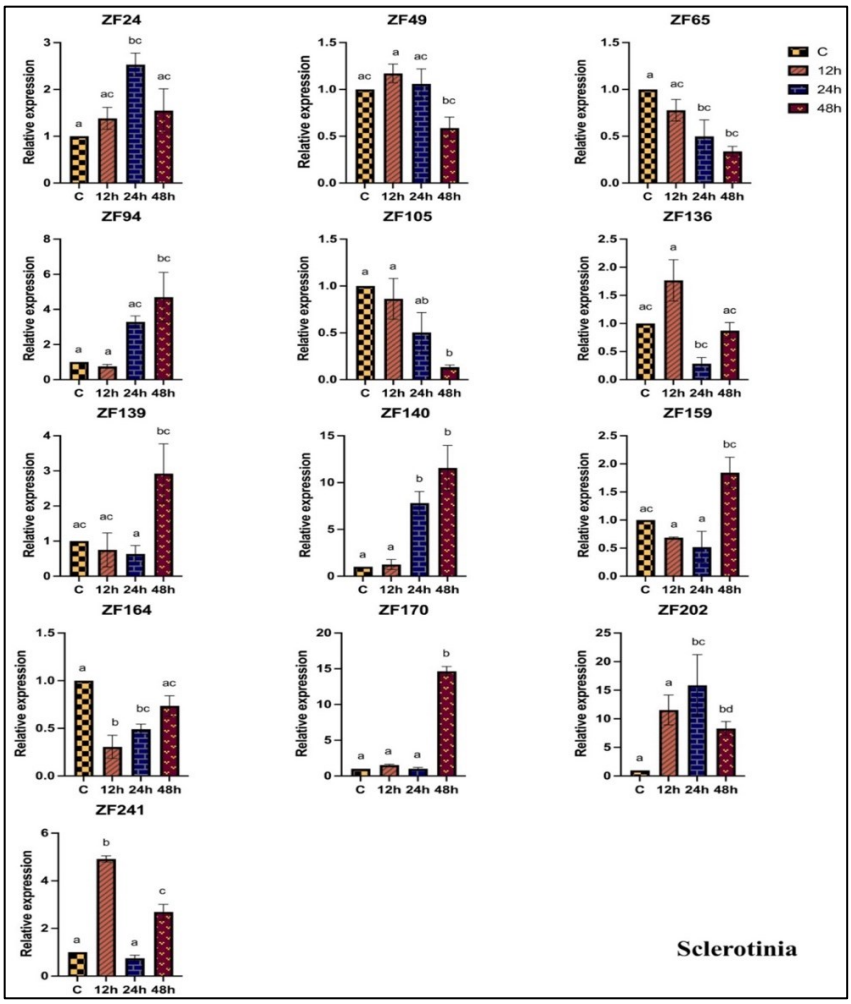


Figure 6.5 Expression profiles of 13 *Brassica napus* C2H2-ZF genes upon *Sclerotinia sclerotiorum* infection as determined by qRT-PCR. Bar and error bar indicates the mean of three biological replicates ± standard error (S.E.). Different letters indicate significant difference between the mean values (p<0.05, Tukey method).

6.3.8 Expression of the *B. napus* C2H2-ZF genes under abiotic stresses

Cold stress

We analysed the expression patterns of the 20 C2H2-ZF genes after exposure of the spring canola line DH12075 and the winter canola cv. Mendel to cold stress for up to 2 days (48 hours) by qRT-PCR. While comparing the expression of the genes between the spring and winter types, a significantly (p <0.05) lower expression in Mendel was observed for *BnaZF24* and *BnaZF90* at both time points (24 h and 48 h of exposure) (Figure 6.6), while for *BnaZF49*, *BnaZF202* and

BnaZF241 the expression was significantly (p <0.05) lower only at 24 h after cold stress (Figure 6.6). While comparing the level of transcript accumulation in Mendel with the progression of cold treatment, *BnaZF140* and *BnaZF170* showed a significantly (p <0.05) lower transcripts levels at 48 h as compared to 24 h of treatment. On the other hand, a significantly (p <0.05) greater expression in Mendel was found for *BnaZF94* at 48 h. A similar trend of reduced transcript accumulation with the progression of cold treatment was observed for *BnaZF170* in both spring and winter canola (Figure 6.6). Eight genes (*BnaZF1, BnaZF22, BnaZF23, BnaZF33, BnaZF65, BnaZF139, BnaZF162* and *BnaZF211*) did not show any change in their expression patterns after exposure to cold (Supplementary Figure S6.7).

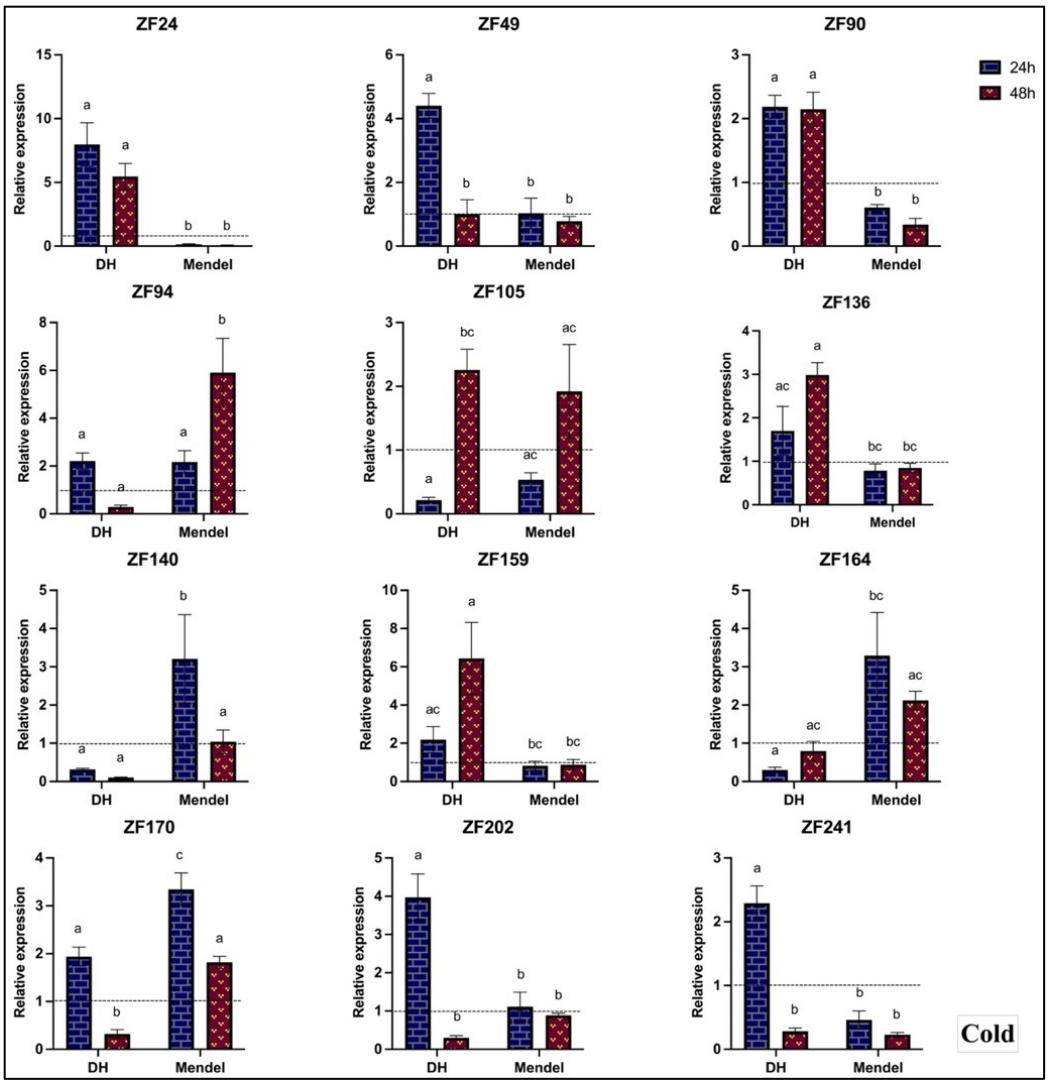


Figure 6.6 Expression profiles of 12 *Brassica napus* C2H2-ZF genes upon cold stress as determined by qRT-PCR. Bar and error bar indicates the mean of three biological replicates ± standard error (S.E.). Different letters indicate significant difference between the mean values (p<0.05, Tukey method). Expression level was normalized to the control for each time point and indicated by dotted line at expression level = 1.0.

Drought and salinity stress

In response to drought stress (20% w/v PEG8000), 13 C2H2-ZF genes did not exhibit differential expression due to PEG treatment (Supplementary Figure S6.8), while expression changes were observed in the case of seven genes at different time points (Figure 6.7). Of the seven genes, *BnaZF90* and *BnaZF94* showed significantly (p <0.05) greater expression at 6 h after the PEG

treatment; however, their expression decreased gradually at 24 h and 48 h after the treatment. The highest expression of *BnaZF23, BnaZF33, BnaZF140* and *BnaZF170* was observed at 24 h after the treatment (Figure 6.7A).

In case of salinity stress, expression of *BnaZF23*, *BnaZF90* and *BnaZF162* was significantly (p <0.05) decreased with the progression of exposure to salinity. Expression of *BnaZF94* did not show significant changes at 6 h and 24 h after exposure, however, it's expression significantly increased at 48 h (3.51-fold increase) (Figure 6.7B). In the case of *BnaZF139*, it's expression significantly (p <0.05) reduced at 6 h and 24 h, however, increased above the control at 48 h after exposure. Four genes (*BnaZF33, BnaZF164, BnaZF170* and *BnaZF202*) showed an increased expression at 6 h or 24 h after exposure (Figure 6.7B), and no significant difference was observed in the expression of the remaining 11 genes following salinity stress (Figure S6.8).

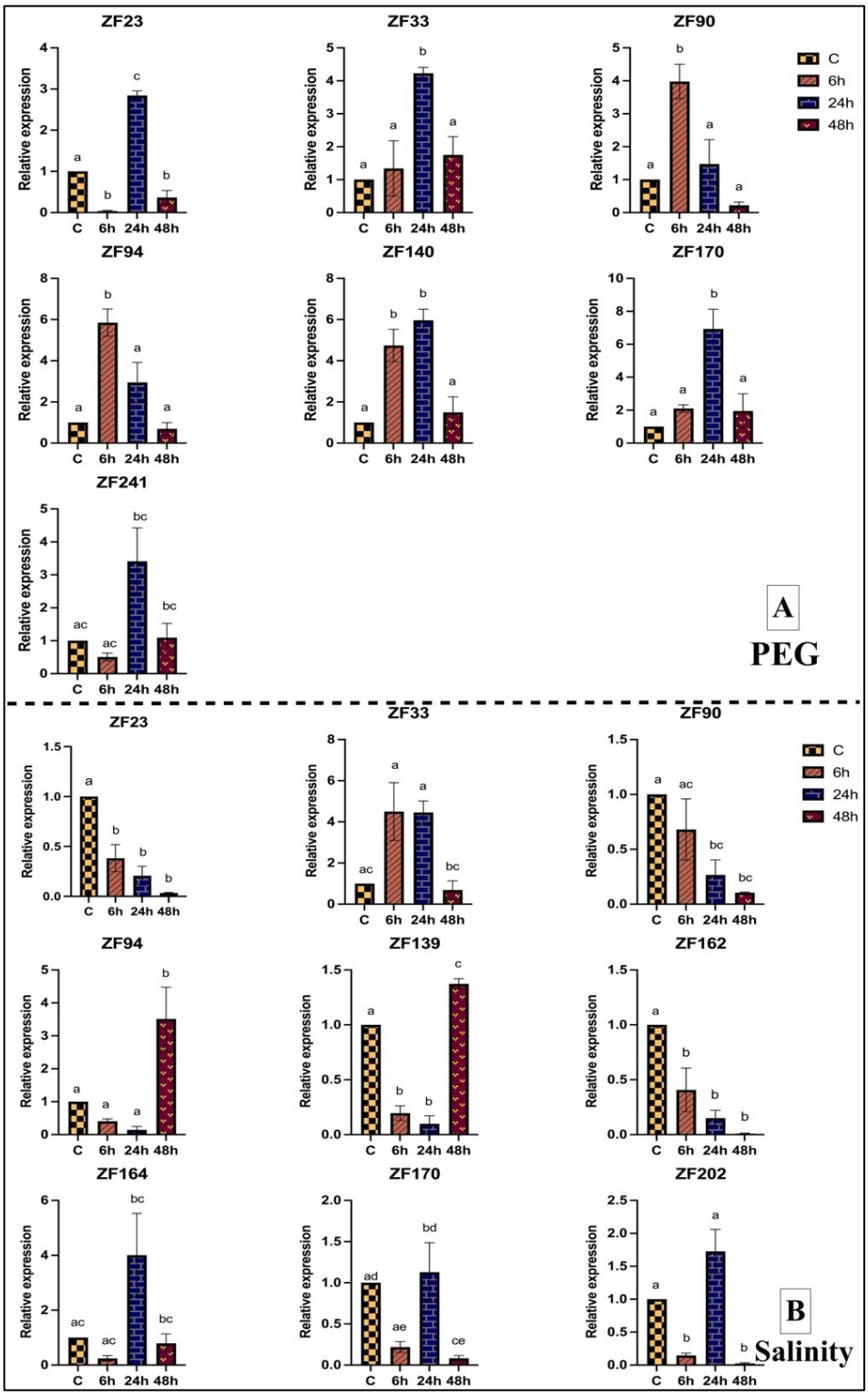


Figure 6.7 Expression profiles of *Brassica napus* C2H2-ZF genes under (A) drought (PEG) and (B) salinity treatments as determined by qRT-PCR. Bar and error bar indicates the mean of three biological replicates ± standard error (S.E.). Different letters indicate significant difference between the mean values (p<0.05, Tukey method).

6.3.9 Expression of the *B. napus* C2H2-ZF genes in response to phytohormone treatments

In the case of ABA treatment, four genes, viz. *BnaZF164, BnaZF170, BnaZF211* and *BnaZF241*, showed a significantly (p <0.05) greater expression at 48 h as compared to 6 h and 24 h as well as the control (Figure 6.8A). *BnaZF24* showed a significantly increased expression at 48 h; however, their expression at 6 h and 24 h was not significantly (p <0.05) different from the control. In contrast, the expression of *BnaZF90* and *BnaZF159* declined significantly (p <0.05) over the time with maximum being detected at 6 h. No significant difference in expression was observed for eleven C2H2-ZF genes (Supplementary Figure S6.9).

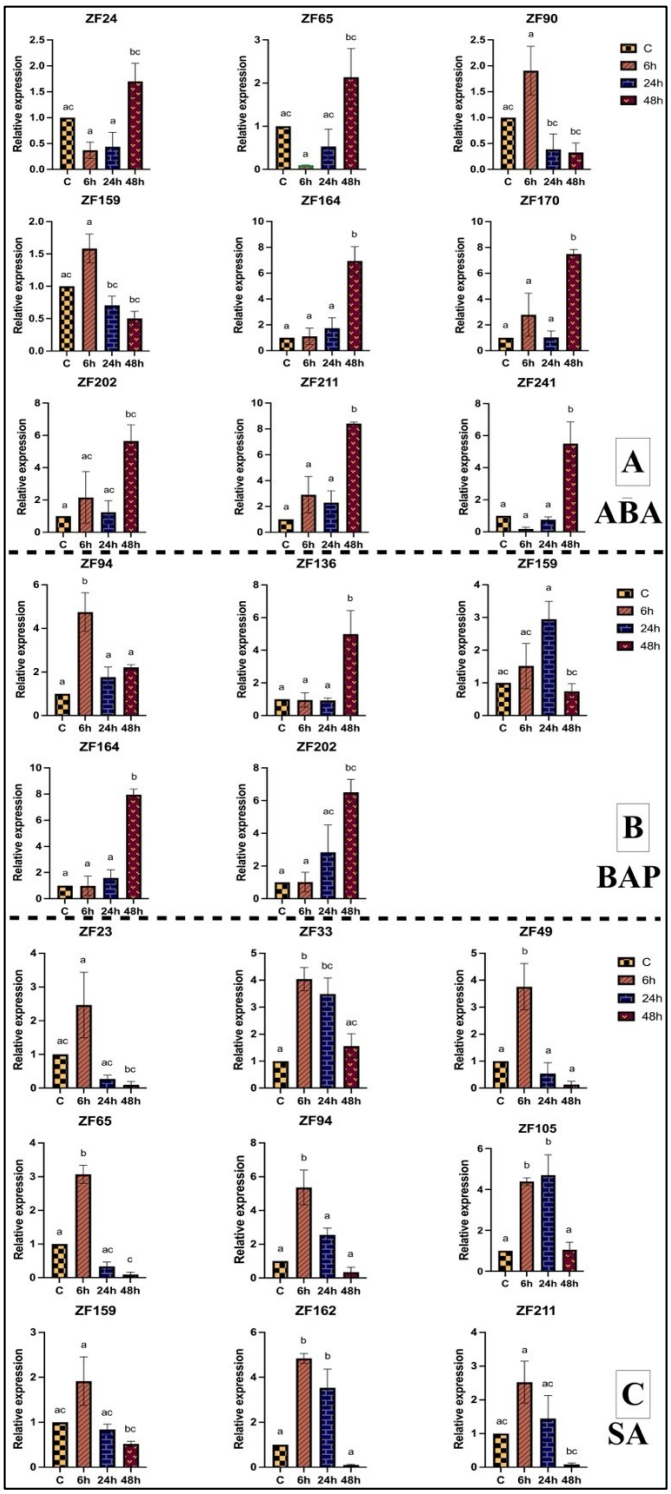


Figure 6.8 Expression profiles of the *Brassica napus* C2H2-ZF genes under treatments with different phytohormones as determined by qRT-PCR: (A) ABA (abscisic acid) treatment, (B) BAP (6-benzylaminopurine) treatment, and (C) SA (salicylic acid) treatment. Bar and error bar indicates the mean of three biological replicates ± standard error (S.E.). Different letters indicate significant difference between the mean values (p<0.05, Tukey method).

BAP treatment changed the expression of *BnaZF94* at 6 h, whereas the expression of *BnaZF136, BnaZF164* and *BnaZF202* increased at 48 h as compared to control as well as their expression at 6 h, where the greatest level of expression change (7.95-fold) was observed in the case of *BnaZF164* (Figure 6.8B). Expression pattern of *BnaZF202* was similar to the expression pattern of this gene in response to ABA treatment (Figure 6.8A, B). No significant difference was observed for expression of 15 C2H2-ZF genes (Supplementary Figure S6.9).

In case of SA treatment, significantly (p <0.05) greater expression was detected at 6 h for *BnaZF49*, *BnaZF65* and *BnaZF94*, but their expression decreased significantly (p <0.05) with the progression of the treatment (Figure 6.8C). The transcript levels of *BnaZF105* and *BnaZF162* stayed high (>4-fold) at 6 h to 24 h and significantly (p <0.05) decreased at 48 h of the SA treatment. Eleven of the 20 C2H2-ZF genes showed no significant change in their expression (Supplementary Figure S6.9).

6.4 Discussion

With the advent of next-generation DNA sequencing technologies, the genomes of many crops including canola have been sequenced. This allows the identification of genes controlling different traits, including stress tolerance, which may permit the development of abiotic and biotic stress tolerant crops. In this study, we identified 267 C2H2-ZF genes in the *B. napus* genome and classified them into 12 classes (Supplementary Table S6.3) based on the plant-specific, conserved amino acid sequence "QALGGH" and the distance between the metal ligands within the C2H2-ZF domains. Among them, only 27.3% possessed tandem ZFs and the majority had isolated ZFs. However, in animal proteins, the "QALGGH" sequence cannot be found and, in addition, the neighbouring fingers are spaced by HC-link composed of 6-8 amino acids (Takatsuji 1999; Kielbowicz-Matuk 2012).

In this study with *B. napus*, we observed that most C2H2-ZF genes contained one or two domains with Q-type motif, which is similar to the reports in *B. rapa* (Alam et al. 2019), cucumber (Chen et al. 2020) and wheat (Faraji et al. 2018). The Q-type C2H2-ZFPs uniquely occur in plants implying their role in plant-specific processes such as growth and development, organogenesis, stress tolerance and defense. The presence of other domains (M1, M2, M3, M4, M5, Z1, Z2, D) denotes the functional diversity of the C2H2-ZFPs. About 96.6% of the C2H2-ZF genes were predicted to be located in the nucleus, confirming their role as transcription factors regulating gene expression (Takatsuji et al. 1992; Morita et al. 2006; Kundu et al. 2018). In contrast, the remaining 3.4% C2H2-ZFPs were located in other cellular organelles (e.g., mitochondria, and chloroplast) or in the cytosol inferring that they may play a role in the regulation of extranuclear genes, such as chloroplastic and mitochondrial genes.

Gene structure analysis carried out in this study demonstrated the presence of a single exon in 41.6% of the C2H2-ZFPs. This may imply that these genes are retained in plants and respond to environmental stresses very rapidly (Hu et al. 2019). Single exon genes are involved in different biological processes, including development and stress responses (Liu et al. 2021). Two of the C2H2-ZF genes contained many introns suggesting the possibility of the formation of different functional protein isoforms because of alternative splicing. It results in the formation of different protein isoforms, leading to changes in the binding properties, intracellular localization, protein stability, enzymatic and signalling activities in plants (Stamm et al. 2005). Also, the differences that we observed in this study for gene length apparently have resulted from gene duplication, integration, or realignment of the genomic regions during the evolution of these genes. It has been reported by Hu et al. (2019) that the genes with a shorter length are involved in response to biotic and abiotic stresses in tomato. In this study, we found that *BnaZF202* and *BnaZF241* had a gene

length of <1 kb and they were involved in multiple stress responses, such as stresses due to *P. brassicae* infection, Sclerotinia infection, cold and salinity.

The variations in protein length, mW and pI values that we observed in this study indicate a high degree of complexity within this gene family. The three Brassica genomes have experienced a whole genome triplication event, which occurred about 15.9 million years ago after the divergence from *A. thaliana* about 17 million years ago (Lysak et al. 2007). Our synteny analysis revealed that, 234 of the 267 C2H2-ZF genes had syntenic relationship either with *B. rapa* or *B. oleracea*. This implies that most of the C2H2-ZF genes in *B. napus* descended from the A and C genomes of the two progenitor species during the evolution of this allopolyploid species. We also found several gene duplications, including four tandemly duplicated gene pairs and 62 segmental duplication events, indicating that the gene duplications have resulted in the expansion of the C2H2-ZF gene family in *B. napus*.

The role of C2H2-ZF gene family in biotic stress tolerance has been reported in many plant species, including *C. annuum* (Kim et al. 2004), potato (Tian et al. 2010), and tobacco (Uehara et al. 2005). A C2H2-ZF gene *CAZFP1* has been found to be significantly induced in response to infection by bacterial pathogen *Xanthomonas campestris* and fungal pathogen *Colletotrichum coccodes* in transgenic *Arabidopsis* (Kim et al. 2004). *CanZF40* and *CanZF46* are predicted to be involved in resistance to anthracnose pathogen in chilli pepper (Sharma et al. 2021). In the case of clubroot, Jiang et al. (2021) reported an increased expression of a C2H2-ZF gene (#16) in a clubroot resistant canola cultivar with the progression of *P. brassicae* infection from 7 to 21 days after inoculation. In the present study, we also found an increased expression of three C2H2-ZF genes in clubroot resistant canola lines after inoculation with *P. brassicae*. This indicates that the C2H2-ZF genes may play a role in response to infection by *P. brassicae*.

Sclerotinia stem rot disease causes significant yield loss in *B. napus* canola. Management of this disease in this crop is a challenging task due to the lack of good genetic resistance (Bradley et al. 2006). Joshi et al. (2016) investigated transcript level changes in canola due to infection by *S. sclerotiorum* and identified 30 TFs belonging to 12 different families including WRKY, NAC, MYB, C3H zinc finger, and C2H2 zinc finger to be induced at early- and late-infection stages. Wang et al. (2014) found that overexpression of *WRKY33* in *B. napus* activates the Pathogenesis Related (*PR1*) and Plant Defensin (*PDF1.2*) (defense responsive) genes and increases resistance to *S. sclerotiorum.* This demonstrates the regulatory roles of the TFs in host plants in response to pathogen infection. In this regard, it is possible that the C2H2-ZF genes that we identified in this study exhibiting significantly increased expression might play an important role in sclerotinia stem rot resistance; however, functional validation of these genes in *B. napus* will be needed.

There are several reports in literature which suggest a role for C2H2-ZF genes in response to various abiotic stresses in plants (Liu et al. 2015; Alam et al. 2019; Hu et al. 2019). For example, the C2H2-ZF gene, *GsZFP1* plays a role in cold and drought tolerance in *A. thaliana* (Luo et al. 2012), while *ZAT18* (*At3g53600*) has been reported to be a positive regulator enhancing drought tolerance in this plant species (Yin et al. 2017). In addition, the constitutive expression of ZFPs has also been reported to increase salt and osmotic stress tolerance in *A. thaliana* (Zhang et al. 2016). The role for ZFPs in ABA signalling during germination and seedling development has also been suggested (Joseph et al. 2014). In this study, we observed the differential expression of several C2H2-ZF genes in response to different abiotic stresses and/or phytohormone treatments at various time points. In addition, we also observed the upregulation of four genes (two in DH12075 and two in Mendel at 24 h and/or 48 h) due to cold stress, while downregulation of two

genes due to salinity stress. Thus, the results from our study further substantiate the importance of the C2H2-ZF genes in abiotic stress tolerance in *B. napus*.

The involvement of C2H2-ZF genes in response to multiple stresses has been reported by Shi et al. (2014) who observed that the transcript for the C2H2-ZF gene *AtZAT6* was upregulated in *A. thaliana* in response to salinity, drought and cold stresses as well as due to infection by *Pseudomonas syringae* pv *tomato*. The over-expression of *lbZFP1*, a C2H2-ZF gene from sweet potato, significantly increased salt and drought tolerance in transgenic Arabidopsis plants (Wang et al. 2016). In this study, we also found differential expression of some of the C2H2-ZF genes, such as *BnaZF90, BnaZF94, BnaZF170, BnaZF202* and *BnaZF241*, in response to multiple stresses, and a few genes, such as *BnaZF94*, exhibited similar expression pattern in response to more than one phytohormone treatment. This provides evidence that some of the *B. napus* C2H2-ZF genes can play a role in response to multiple biotic/abiotic stresses.

To our knowledge, this is the first report on the genome-wide identification and characterization of the C2H2-ZF genes and investigation of their putative roles in oilseed *B. napus* in response to biotic and abiotic stresses. The ZFP genes that we identified in this study may be potential targets for manipulation to improve abiotic and biotic stress tolerance in canola. However, functional characterization of these genes, for example, through transformation of *A. thaliana*, will be needed to confirm this as well as to get a better understanding of these genes in stress tolerance.

Chapter 7. General discussion, conclusion and future prospects

7.1 General discussion and conclusion

Canola, scientifically known as *B. napus*, is a significant oilseed crop in Canada, both in terms of agricultural production and economic significance (USDA 2022). It is primarily grown for edible oil; however, this oil is also used for industrial purposes, such as production of biodiesel. The canola oil is composed of low saturated fatty acids, high levels of monounsaturated and polyunsaturated fatty acids and almost no erucic acid (Przybylski et al. 2005). Canada, being one of the largest producers of canola worldwide, contributes significantly to its global production. The crop is mainly grown in the Western provinces including Alberta, Saskatchewan, and Manitoba (Canola Council of Canada 2020). During each stage of its growth, canola is exposed to various abiotic and biotic stresses; among these clubroot disease has been one of the longstanding threats. This disease is caused by *Plasmodiophora brassicae* Woronin and has been reported worldwide with major yield losses in Europe, Canada, and China (for review, see Dixon et al. 2009; Chai et al. 2014).

It is crucial to remember that managing the clubroot disease calls for a proactive, multifaceted strategy. Owing to the long life of the *P. brassicae* resting spores in the soil (Wallenhammer 1996), the cultural practices and chemical control strategies alone were found to be ineffective in controlling this disease (for review, see Hwang et al. 2014; Peng et al. 2014a; Cao et al. 2020). Therefore, growing of clubroot resistant cultivars is the effective approach to minimize disease severity and yield losses (for review, see Hwang et al. 2014). Plant breeders have developed clubroot resistant canola cultivars with varying levels of resistance to specific pathotypes. To date, about 24 major clubroot resistance (CR) loci and more than 60 quantitative trait loci (QTL) have been identified in different Brassica species (for review, see Hasan et al.

2021a). These CR loci/QTL confer resistance against specific *P. brassicae* pathotypes. The pathotype population in the soil changes over time due to the emergence of new virulent pathotypes rendering the available resistances ineffective (Strelkov et al. 2016b; Askarian et al. 2021; Hollman et al. 2021). This necessitates the introgression of additional resistance genes from different Brassica genetic resources to broaden the genetic base of CR genes in *B. napus* canola as well as pyramiding of multiple resistance genes in a cultivar for resistance to multiple pathotypes.

The CR loci from the A genome of *B. rapa* and amphidiploid rutabaga (*B. napus* var. *napobrassica*) have been introgressed into *B. napus* canola (Karling 1969; Johnston 1970; Ayers and Lelacheur 1972; Buczacki et al. 1975; Diederichsen and Sacristan 1996; Frauen 1999; Rahman et al. 2011, 2014b; Hasan et al. 2012; Peng et al. 2014b; Hirani et al. 2016; Fredua-Agyeman et al. 2019). Several accessions of European fodder turnips (*B. rapa* ssp. *rapifera*), including ECD 01 carry CR loci conferring resistance against different *P. brassicae* races or pathotypes (Hirai et al. 2004; Suwabe et al. 2003, 2006; Piao et al. 2004; Sakamoto et al. 2008; Chen et al. 2013; Hirani et al. 2018; Fredua-Agyeman et al. 2020; Zhan et al. 2020; Yu et al. 2022). Therefore, the utilization of this resistance in canola breeding was expected to introduce new CR loci and resistance to the recently evolved virulent pathotypes including pathotype 3A. Clubroot resistance from *B. rapa* into canola has been done by different researchers and following different approaches, including the *B. napus* × *B. rapa* interspecific cross (Chu et al. 2013; Hirani et al. 2016; for review, see Hasan et al. 2021a); however, details of the behaviour of the progeny of this cross in regards to clubroot resistance has not been described in literature. This thesis research focused on the introgression of clubroot resistance from the *B. rapa* European fodder turnip cv. Debra (ECD 01) into spring *B. napus* canola through *B. napus* × *B. rapa* interspecific cross. The F_1 plants were backcrossed to canola and the BC_1 population was subjected to pedigree breeding

with selection for clubroot resistance and canola-quality traits. The selection for resistance to pathotype 3H was performed in the early generations and for resistance to pathotype 3A in BC_1F_5 generation. The advanced generation lines (BC_1F_8) were evaluated for resistance to pathotype 3H, 3A, 2B and 5X as well as for agronomic and seed quality traits. Following this approach, it was possible to develop BC_1F_8 canola lines with nuclear DNA content similar to *B. napus* as well as resistant to the above-mentioned pathotypes. Thus, results from this thesis research demonstrate that resistance to multiple pathotypes can be introgressed from a turnip into canola even when selection for resistance performed to a single pathotype in early generations. The PCA and correlation analysis of these advanced generation lines suggested the involvement of a similar genetic control of resistance to pathotypes 3A, 2B and 5X; however, a different and simpler genetic control is involved for resistance to pathotype 3H.

The traditional plant breeding approaches, which is solely based on the phenotypic selection, can limit the genetic gains for the desirable traits (for review, see Langridge et al. 2011). This demands a deeper understanding of information flow from the gene to protein via transcript for the improvement of complex traits, like disease resistance. The decrease in the costs of next-generation sequencing techniques has fastened the discovery of disease resistance genes and QTL in different crops to harness precision breeding (Varshney et al. 2009). This thesis research also focused on understanding the genetic and molecular basis of the clubroot resistance of ECD 01 introgressed into canola with the aid of WGRS. By using the *B. napus* recombinant inbred line (RIL) population derived from the Canola × Turnip ECD 01 interspecific cross, I identified two major CR loci located on chromosome A03 (0.98-2.0 Mb) and A09 (38.01- 41.5 Mb) conferring resistance to pathotype 3H and 3A. I also found a genomic region on chromosome A02 (1-4 Mb) and on A08 (14-21 Mb) contributing resistance to pathotype 3H and 3A. Previous studies have

identified 18.50-26.35 Mb region of A02, 15.5-16.3 Mb and 23.8-26.8 Mb region of A03, and 11.34-21.75 Mb region of A08 conferring resistance to *P. brassicae* (Yu et al. 2017; for review, see Hasan et al. 2021a). According to Yu et al. (2022), the loci $Rcr10^{ECD01}$ and $Rcr9^{ECD01}$ of the chromosomes A03 and A08 respectively, confer resistance to pathotypes 3H, 3A and 3D while $Rcr9^{ECD01}$ contributes to the resistance to pathotype 5X. By using a *B. napus* population carrying a resynthesized *B. napus* (*B. oleracea* × *B. rapa* ECD 04) in their pedigree, Werner et al. (2008) reported a minor effect QTL on A09 with an additive effect of only 4.9%; however, physical position of this QTL could not be found. Based on available literature, it is evident that I detected three novel loci (A02, A03 and A09) conferring resistance to *P. brassicae* pathotypes 3H and 3A; among these, the A09 locus explained more than 50% of the total phenotypic variance. These three loci can be deployed in pyramiding of genes in *B. napus* canola for imparting resistance to multiple pathotypes.

A good understanding of host-pathogen interactions and identification of the candidate genes involved in host defense responses in *P. brassicae-B. napus* pathosystem is needed for the development of gene-based molecular markers for use in molecular breeding. RNA-seq is a powerful tool for the analysis of differentially expressed genes (DEGs). This technique has been employed by several researchers to understand the molecular mechanisms of clubroot resistance in Brassica (Chu et al. 2014; Chen et al. 2016a; Zhang et al. 2016; Jia et al. 2017; Ciaghi et al. 2019; Fu et al. 2019; Mei et al. 2019; Wang et al. 2019; Wei et al. 2021; Summanwar et al. 2020; Zhou et al. 2020; Yuan et al. 2021b; Adhikary et al. 2022b). Most of these studies have been carried out using clubroot resistant and susceptible cultivars or lines with wide difference in their background genetics. The use of these type of materials in the RNA-seq study may not only disclose the DEGs involved in clubroot resistance but also several other DEGs involved in the

control of other traits. This constraint can be overcome to a great extent by using near-isogenic lines (NILs). Only few studies have been conducted on clubroot resistance in Brassica using NILs (Kim et al. 2011; Chen et al. 2016a; Li et al. 2021); however, none of these used the clubroot resistance of the ECD 01. I used clubroot-resistant and clubroot-susceptible *B. napus* NILs carrying resistance introgressed from turnip (*B. rapa* var. *rapifera* ECD 01) for comparative transcriptome analysis of clubroot resistance. These NILs were resistant to at least 15 *P. brassicae* pathotypes, including the ones that evolved most recently. I witnessed an intense defense response in the resistant NILs as compared to the susceptible NILs. The results from this thesis research provide information on the role of genes associated with phytohormone signaling, disease resistance, calcium influx and ROS, WRKY TFs, glucosinolates biosynthesis in clubroot resistance in *B. napus*. These genes can be used as the candidates for functional validation to extend our knowledge on the molecular mechanisms underlying clubroot resistance in *B. napus* as well as to develop gene-based molecular markers.

To date, several transcriptomics studies have been carried out to understand the *Brassica-P. brassicae* interactions (Jia et al. 2017; Fu et al. 2019; Adhikary et al. 2022b; the present thesis research); however, our knowledge of the resistance at protein level is limited. The dynamic nature of the proteins, i.e., the products of the transcribed genes, present in a cell at a given time point necessitates the investigation of the *B. napus-P. brassicae* pathosystem at proteome-level. In this thesis research, I performed a comparative proteome analysis using the afore-mentioned homozygous clubroot resistant and susceptible *B. napus* NILs and found the involvement of putative proteins related to calcium-signalling pathway, ROS, phytohormones, glucosinolates as well as other proteins related to resistance mechanisms to be involved in the 'Turnip-resistance' in spring *B. napus* canola.

The transcriptomics study from this thesis research as well as from the previous studies have highlighted the importance of transcription factors (TFs) in mediating host responses against different biotic stresses including clubroot disease. Among all the TFs, the C2H2- zinc finger proteins (ZFPs) have been of significant interest to the researchers; however, they lacked characterization in *B. napus*. In this thesis research, I identified 267 C2H2-ZF genes in *B. napus* distributed on the 19 chromosomes; these genes carry stress responsive *cis*-acting elements in the promoter regions. Most of the *B. napus* C2H2-ZF genes descended from the A and C genomes of the two progenitor species *B. rapa* and *B. oleracea* during the evolution of this allopolyploid species. I also found an increased expression of three C2H2-ZF genes in clubroot resistant canola lines after inoculation with *P. brassicae* indicating that these genes may play a role in *B. napus* in response to infection by *P. brassicae*.

In summary, this thesis research not only used the traditional breeding approaches for introgression of clubroot resistance from turnip into canola, but also used different omics approaches, such as genomics, transcriptomics, and proteomics, to provide insights into the molecular mechanisms of clubroot disease resistance in *B. napus* canola. The knowledge gained from this research can be used for efficient introgression of clubroot resistance from other *B. rapa* into canola through *B. napus* × *B. rapa* interspecific cross, as well as to develop gene-based molecular markers for use in canola breeding.

7.2 Future prospects

- The clubroot resistance of the *B. rapa* European fodder turnip accession ECD 01 (cv. Debra) can be used as an excellent source of resistance to *P. brassicae* pathotypes 3H, 3A, 2B and 5X; this resistance can be used to broaden the genetic base of *B. napus* canola.

- Fine map the CR loci and develop genetic markers for use in marker-assisted selection as well as map-based cloning of the clubroot resistance genes.
- Functional validation of the putative genes/proteins involved in clubroot resistance in *B. napus*.
- Functional characterization of C2H2-ZF genes through transformation of *A. thaliana* to gain better understanding of these genes in stress tolerance in *B. napus*.

Literature cited

Abe, A., Kosugi, S., Yoshida, K., Natsume, S., Takagi, H., Kanzaki, H., Matsumura, H., Yoshida, K., Mitsuoka, C., Tamiru, M., Innan, H., Cano, L., Kamoun, S., and Terauchi, R. **2012**. Genome sequencing reveals agronomically important loci in rice using MutMap. Nature Biotech. **30**:174-178. https://doi.org/10.1038/nbt.2095.

Abel, S., Ballas, N., Wong, L.M., and Theologis, A. **1996**. DNA elements responsive to auxin. Bioessays. **18:**647-654. https://doi.org/10.1002/bies.950180808.

Adhikary, D., Kisiala, A., Sarkar, A., Basu, U., Rahman, H., Emery, N., and Nat, N.V.K. **2022b**. Early-stage responses to *Plasmodiophora brassicae* at the transcriptome and metabolome levels in clubroot resistant and susceptible oilseed *Brassica napus*. Mol. Omics. **18**:991-1014. https://doi.org/10.1039/D2MO00251E.

Adhikary, D., Mehta, D., Uhrig, R. G., Rahman, H., and Kav, N. N. V. **2022a**. A proteome-level investigation into *Plasmodiophora brassicae* resistance in *Brassica napus* canola. Front. Plant Sci. **13**:860393. https://doi.org/10.3389/fpls.2022.860393.

Agarwal, P., Arora, R., and Ray, S. et al. **2007**. Genome-wide identification of C_2H_2 zinc-finger gene family in rice and their phylogeny and expression analysis. Plant Mol. Biol. **65**:467-485. https://doi. org/10.1007/s11103-007-9199-y.

Ahmed, H.U., Hwang, S.F., Strelkov, S.E., Gossen, B.D., Peng, G., Howard, R.J., and Turnbull, G.D. **2011**. Assessment of bait crops to reduce inoculum of clubroot (*Plasmodiophora brassicae*) of canola. Can. J. Plant Sci. **91**(3):545-551. https://doi.org/10.4141/cjps10200.

Aigu, Y., Cao, T., Strelkov, I.S., Manolii, V.P., Lemoine, J., Manzanares-Dauleux, M.J., ... and Gravot, A. **2020**. Identification of winter and spring *Brassica napus* genotypes with partial

resistance to Canadian isolates of *Plasmodiophora brassicae*. Can. J. Plant Pathol. **42**(4): 538-546. https://doi.org/10.1080/07060661.2020.1723870.

Alam, I., Batool, K., Cui, D.L., Yang, Y.Q., and Lu, Y.H. **2019**. Comprehensive genomic survey, structural classification, and expression analysis of C2H2 zinc finger protein gene family in *Brassica rapa* L. PLoS One. **14**(5):p.e0216071. https://doi.org/10.1371/journal.pone.0216071.

Ali, A., Alexandersson, E., Sandin, M., Resjo, S., Lenman, M., Hedley, P., Levander, F., and Andreasson, E. **2014**. Quantitative proteomics and transcriptomics of potato in response to *Phytophthora infestans* in compatible and incompatible interactions. BMC Genomics. **15**:1-18. https://doi.org/10.1186/1471-2164-15-497.

Ali, M.S., and Baek, K.H. **2020**. Protective roles of cytosolic and plastidal proteasomes on abiotic stress and pathogen invasion. Plants (Basel). **9**(7):832. https://doi.org/ 10.3390/plants9070832.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. 1990. Basic local alignment search tool. J. Mol. Biol. 215:403-410. https://doi.org/10.1016/S0022-2836(05)80360-2.

Apel, K. and Hirt, H. **2004**. Reactive oxygen species: metabolism, oxidative stress, and signal transduction. Annu. Rev. Plant Biol. **55**:373-399. https://doi.org/10.1146/annurev.arplant.55.031903.141701.

Arie, T., Kobayashi, Y., Kono, Y., Gen, O., and Yamaguchi, I. **1999**. Control of clubroot of crucifers by *Phoma glomerata* and its product epoxydon. Pestic. Sci. **55**:602-604. https://doi.org/10.1002/(SICI)1096-9063(199905)55:5<602::AID-PS964>3.0.CO;2-C.

Arie, T., Kobayashi, Y., Okada, G., Kono, Y., and Yamaguchi, I., **1998**. Control of soilborne clubroot disease of cruciferous plants by epoxydon from *Phoma glomerata*. Plant Pathol. **47**: 743-748. https://doi.org/10.1046/j.1365-3059.1998.00298.x.

Aroca, A., Benito, J.M., Gotor, C., and Romero, L.C. **2017**. Persulfidation proteome reveals the regulation of protein function by hydrogen sulfide in diverse biological processes in *Arabidopsis*. J. Exp. Bot. **68**(17):4915-4927. https://doi.org/10.1093/jxb/erx294.

Askarian, H., Akhavan, A., Manolii, V.P., Cao, T., Hwang, S.F., and Strelkov, S.E. **2021**. Virulence spectrum of single-spore and field isolates of *Plasmodiophora brassicae* able to overcome resistance in canola (*Brassica napus*). Plant Dis. **105**:43-52. https://doi.org/10.1094/PDIS-03-20-0471-RE.

Aslam, B., Basit, M., Nisar, M.A., Khurshid, M., and Rasool, M.H. **2017**. Proteomics: Technologies and their applications. J. Chromatogr. Sci. **55**(2):182-196. https://doi.org/10.1093/chromsci/bmw167.

Attia, T., and Röbbelen, G. **1986**. Cytogenetic relationship within cultivated Brassica analyzed in amphihaploids from the three diploid ancestors. Can. J. Genet. Cytol. **28**(3):323-329. https://doi.org/10.1139/g86-048.

Attri, R., and Rahman, H. **2018**. Introgression of allelic diversity from genetically distinct variants of *Brassica rapa* into *Brassica napus* canola and inheritance of the *B. rapa* alleles. Crop Pasture Sci. **69**:94-106. https://doi.org/10.1071/CP17193.

Ayers, G.W., and Lelacheur, K.E. **1972**. Genetics of resistance in rutabaga to two races of *Plasmodiophora brassicae*. Can. J. Plant Sci. **52**:897-900. https://doi.org/10.4141/cjps72-155.

Bayer, P.E., Hurgobin, B., Golicz, A.A., Chan, C.K.K., Yuan, Y., Lee, H., Renton, M., Meng, J., Li, R. and Long, Y. **2017**. Assembly and comparison of two closely related *Brassica napus* genomes. Plant Biotechnol. **15**(12):1602-1610. https://doi.org/10.1111/pbi.12742.

Bent, A.F., and Mackey, D. **2007**. Elicitors, effectors, and R genes: the new paradigm and a lifetime supply of questions. Annu. Rev. Phytopathol. **45**:399-436. https://doi.org/10.1146/annurev.phyto.45.062806.094427.

Berardini, T.Z., Reiser, L., Li, D., Mezheritsky, Y., Muller, R., Strait, E., and Huala, E. 2015. The Arabidopsis Information Resource: making and mining the "gold standard" annotated reference plant genome. Genesis. 53:474-485. https://doi.org/10.1002/dvg.22877.

Bernstein, N. **2019**. Plants and salt: Plant response and adaptations to salinity. In *Model Ecosystems in Extreme Environments* 101-112. Academic Press. https://doi.org/10.1016/B978-0-12-812742-1.00005-2.

Bhattacharya, I., Dutta, S., Mondal, S., and Mondal, B. **2014**. Clubroot disease on *Brassica* crops in India. Can. J. Plant Pathol. **36**: 154-160. https://doi.org/10.1080/07060661.2013.875064.

Blilou, I., Xu, J., Wildwater, M., Willemsen, V., Paponov, I., Friml, J., Heidstra, R., Aida, M., Palme, K., and Scheres, B. **2005**. The PIN auxin efflux facilitator network controls growth and patterning in Arabidopsis roots. Nature. **433**:39-44. https://doi.org/10.1038/nature03184.

Boersema, P.J., Kahraman, A., and Picotti, P. **2015**. Proteomics beyond large-scale protein expression analysis. *Curr. Opin. Biotechnol.* **34**:162-170. https://doi.org/10.1016/j.copbio.2015.01.005.

Böhm, S., Frishman, D., and Mewes, H.W. **1997**. Variations of the C2H2 zinc finger motif in the yeast genome and classification of yeast zinc finger proteins. Nucleic acids res. **25**(12): 2464-2469. https://doi.org/10.1093/nar/25.12.2464.

Bolger, A.M., Lohse, M., and Usadel, B. **2014**. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics. **30**:2114-2120. https://doi.org/10.1093/bioinformatics/btu170.

Boller, T., and Felix, G. **2009**. A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. Annu. Rev.Plant Biol. **60**(1): 379-406. https://doi.org/10.1146/annurev.arplant.57.032905.105346.

Bradbury, P.J., Zhang, Z., Kroon, D.E., Casstevens, T.M., Ramdoss, Y., and Buckler, E.S. 2007. TASSEL: Software for association mapping of complex traits in diverse samples. Bioinformatics. 23:2633-2635. https://doi.org/10.1093/bioinformatics/btm308.

Brader, G., Mikkelsen, M. D., Halkier, B. A., and Tapio Palva, E. **2006**. Altering glucosinolate profiles modulates disease resistance in plants. The Plant Journal. **46**(5):758-767. https://doi.org/10.1111/j.1365-313X.2006.02743.x.

Bradley, C.A., Henson, R.A., Porter, P.M., LeGare, D.G., del Río, L.E., and Khot, S.D. **2006**. Response of canola cultivars to *Sclerotinia sclerotiorum* in controlled and field environments. Plant Dis. **90**:215-219. https://doi.org/10.1094/PD-90-0215.

Bradshaw, J., Gemmell, D., and Wilson, R. **1997**. Transfer of resistance to clubroot (*Plasmodiophora brassicae*) to swedes (*Brassica napus* L. var. *napobrassica* Peterm) from *B. rapa*. Ann. Appl. Biol. **130**:337-348. https://doi.org/10.1111/j.1744-7348. 1997.tb06837.x.

Buczacki, S. T. **1983**. "Plasmodiophora. an inter-relationship between biological and practical problems," in Zoosporic Plant Pathogens: A Modern Perspective, ed. S. T. Buczacki (London: Academic Press Inc.), 161-191.

Buczacki, S.T., and Ockendon, J.G. **1979**. Preliminary observations on variation in susceptibility to clubroot among collections of some wild crucifers. Ann. Appl. Biol. **92**:113-118. https://doi.org/10.1111/j.1744-7348.1979.tb02963.x.

Buczacki, S., Toxopeus, H., Mattusch, P., Johnston, T., Dixon, G., and Hobolth, L. **1975**. Study of physiologic specialization in *Plasmodiophora brassicae*: Proposals for attempted rationalization through an international approach. Trans. Br. Mycol. Soc. **65**(2):295-303. British Mycological Society. https://doi.org/10.1016/S0007-1536(75)80013-1.

Busso, C., Attia, T., and Röbbelen, G. **1987**. Trigenomic combinations for the analysis of meiotic control in the cultivated Brassica species. Genome. **29**(2):331-333. https://doi.org/10.1139/g87-054.

Cai, C., Wang, X., Liu, B., Wu, J., Liang, J., Cui, Y., Cheng, F. and Wang, X. **2017**. *Brassica rapa* genome 2.0: a reference upgrade through sequence re-assembly and gene re-annotation. Mol. Plant. **10**(4): 649-651. http://dx.doi.org/10.1016/j.molp.2016.11.008.

Campos, W.F., Dressano, K., Ceciliato, P.H.O., Guerrero-Abad, J.C., Silva, A.L., Fiori, C.S., Do Canto, A.M., Bergonci, T., Claus, L.A.N., Silva-Filho, M.C., Moura, D.S. **2018**. *Arabidopsis thaliana* rapid alkalinization factor 1-mediated root growth inhibition is dependent on calmodulin-like protein 38. J. Biol. Chem. **293**(6):2159-2171. https://doi.org/10.1074/jbc.M117.808881.

Canola Council of Canada. **2020**: Economic impact. https://www.canolacouncil.org/about-canola/economic-impact/.

Canola Council of Canada. **2021**. Industry Overview. Available online at: https://www.canolacouncil.org/about-canola/industry/ (accessed March 27, 2023).

Canola Encyclopedia. **2021**. Clubroot Disease. Available from https://www.canolacouncil.org/canola-encyclopedia/diseases/clubroot/#manage-clubrootdisease (accessed 4 April 2023).

Cao, T., Manolii, V.P., Zhou, Q., Hwang, S.F., and Strelkov, S.E. **2020**. Effect of canola (*Brassica napus*) cultivar rotation on *Plasmodiophora brassicae* pathotype composition. Can. J. Plant Sci. **100**:218-225. https://doi.org/10.1139/cjps-2019-0126.

Cao, T., Manolii, V.P., Strelkov, S.E., Hwang, S.F., and Howard, R.J. **2009**. Virulence and spread of *Plasmodiophora brassicae* [clubroot] in Alberta, Canada. Can. J. Plant Pathol. **31**(3):321-329. https://doi.org/10.1080/07060660909507606.

Cao, T., Srivastava, S., Rahman, M. H., Kav, N. N. V., Hotte, N., Deyholos, M. K., et al. **2008**. Proteome-level changes in the roots of *Brassica napus* as a result of *Plasmodiophora brassicae* infection. Plant Sci. **174**:97-115. https://doi.org/10.1016/j. plantsci.2007.10.002.

Casséus, L. **2009**. Canola: a Canadian success story. Statistics Canada. http://www.statcan.gc.ca/pub/96-325-x/2007000/article/10778-eng.pdf. (Retrived on January 04, 2014).

Chai, A.L., Xie, X.W., Shi, Y.X., and Li, B.J. **2014**. Research status of clubroot (*Plasmodiophora brassicae*) on cruciferous crops in China. Can. J. Plant Pathol. **36**:142-153. https://doi.org/10.1080/07060661.2013.868829.

Chalhoub, B., Denoeud, F., Liu, S., Parkin, I. A. P., Tang, H., Wang, X., Chiquet, J., Belcram, H., Tong, C., Samans, B., Correa, M., Da Silva, C., Just, J., Falentin, C., Koh, C. S., Le Clainche, I., Bernard, M., Bento, P., Noel, B., ... and Wincker, P. **2014**. Early allopolyploid evolution

in the post-Neolithic *Brassica napus* oilseed genome. Science. **345**:950–953. https://doi.org/10.1126/science.1253435.

Chang, A., Lamara, M., Wei, Y., Hu, H., Parkin, I.A.P., Gossen, B.D., et al. **2019**. Clubroot resistance gene Rcr6 in *Brassica nigra* resides in a genomic region homologous to chromosome A08 in *B. rapa.* BMC Plant Biol. **19**(1):224. https://doi.org/10.1186/s12870-019-1844-5. PMID:31142280.

Chen, B., Niu, F., Liu, W.Z., et al. **2016b**. Identification, cloning and characterization of R2R3-MYB gene family in canola (*Brassica napus* L.) identify a novel member modulating ROS accumulation and hypersensitive-like cell death. DNA Res. **23**:101-114. https://doi.org/10.1093/dnares/dsv040.

Chen, C., Chen, H., Zhang, Y., Thomas, H.R., Frank, M.H., He, Y., and Xia, R. **2020**. TBtools: An integrative toolkit developed for interactive analyses of big biological data. Mol Plant. **13(8)**:1194-1202. https://doi.org/10.1016/j.molp.2020.06.009.

Chen, J., Jing, J., Zhan, Z., Zhang, T., Zhang, C., and Piao, Z. **2013**. Identification of novel QTLs for isolate-specific partial resistance to *Plasmodiophora brassicae* in *Brassica rapa*. PLoS One. **8**:e85307. https://doi.org/10.1371/journal.pone.0085307.

Chen, J., Pang, W., Chen, B., Zhang C, and Piao Z. **2016a**. Transcriptome analysis of *Brassica rapa* near-isogenic lines carrying clubroot-resistant and susceptible alleles in response to *Plasmodiophora brassicae* during early infection. Front Plant Sci. **6**:1183. https://doi.org/10.3389/fpls.2015.01183.

Chen, Y., Wang, G., Pan, J., Wen, H., Du, H., Sun, J., Zhang, K., Lv, D., He, H., Cai, R., and Pan, J. **2020**. Comprehensive genomic analysis and expression profiling of the C2H2 Zinc Finger

protein family under abiotic stresses in cucumber (*Cucumis sativus* L.). Genes. **11**:171. https://doi.org/10.3390/genes11020171.

Cheng, F., Wu, J., Fang, L., and Wang, X. **2012**. Syntenic gene analysis between *Brassica rapa* and other Brassicaceae species. Front. Plant Sci. **3**:198. https://doi.org/10.3389/fpls.2012.00198.

Chèvre, A.M., Mason, A.S., Coriton, O., Grandont, L., Jenczewski, E., and Lysak, M.A. **2018**. Cytogenetics, a science linking genomics and breeding: *The Brassica Model*. In: Liu, S., Snowdon, R., Chalhoub, B. (eds) The *Brassica napus* Genome. In: Liu, S., Snowdon, R., Chalhoub, B. (eds). https://doi.org/10.1007/978-3-319-43694-4_2.

Chhajed, S., Mostafa, I., He, Y., Abou-Hashem, M., El-Domiaty, M., and Chen, S. **2020**. Glucosinolate biosynthesis and the glucosinolate–myrosinase system in plant defense. Agronomy. **10**(11):1786. https://doi.org/10.3390/agronomy10111786.

Chiang, M. S., and Crete, R. **1970**. Inheritance of clubroot resistance in cabbage (*Brassica oleracea* L. var. *capitata* L.). Can. J. Genet. Cytol. **12**:253-256. https://doi.org/10.1139/g70-036.

Cho, K., Hong, S., Han, Y., Yoon, B., Ryu, S., and Woo, J. **2008**. Resistance in Chinese cabbage using single-spore isolates of *Plasmodiophora brassicae* and development of RAPD markers linked to its resistance gene. J. Crop Sci. Biotechnol. **11**(2):101-106.

Choi, H.W., Lee, B.G., Kim, N.H., Park, Y., Lim, C.W., Song, H.K., and Hwang, B.K. **2008**. A role for a menthone reductase in resistance against microbial pathogens in plants. *Plant Physiology*. **148**(1):383-401. https://doi.org/10.1104/pp.108.119461.

Choi, S.R., Oh, S.H., Chhapekar, S.S., Dhandapani, V., Lee, C.Y., Rameneni, J.J., and et al. **2020**. Quantitative trait locus mapping of clubroot resistance and *Plasmodiophora brassicae*

pathotype Banglim-specific marker development in *Brassica rapa*. Int. J. Mol. Sci. **21**(11):4157. https://doi.org/10.3390/ijms21114157.

Chu, M., Song, T., Falk, K.C., Zhang, X., Liu, X., Chang, A., Lahlali, R., McGregor, L., Gosse, B.D., and Yu, F., et al. **2014**. Fine mapping of *Rcr1* and analyses of its effect on transcriptome patterns during infection by *Plasmodiophora brassicae*. BMC Genomics. **15**(1):1-20. https://doi.org/10.1186/1471-2164-15-1166.

Chu, M., Yu, F., Falk, K.C., Liu, X., Zhang, X., Chang, A., and Peng, G. **2013**. Identification of the clubroot resistance gene *Rpb1* and introgression of the resistance gene into canola breeding lines using a marker-assisted selection approach. Acta Hortic. **1005**:599-605. https://doi.org/10.17660/ActaHortic.2013.1005.74.

Churchill, G.A., and Doerge, R.W. **1994**. Empirical threshold values for quantitative trait mapping. Genetics. **138**:963-371. https://doi.org/10. 1093/genetics/138.3.963.

Ciaghi, S., Schwelm, A., and Neuhauser, S. **2019**. Transcriptomic response in symptomless roots of clubroot infected kohlrabi (*Brassica oleracea* var. *gongylodes*) mirrors resistant plants. BMC Plant Biol. **19**:288. https://doi.org/10.1186/s12870-019-1902-z.

Ciftci-Yilmaz, S., and Mittler, R. **2008**. The zinc finger network of plants. Cell. Mol. Life Sci. **65**: 1150-1160. https://doi.org/10.1007/s00018-007-7473-4.

Cloonan, N., and Grimmond, S.M. **2008**. Transcriptome content and dynamics at single-nucleotide resolution. Genome Biol. **9**:234. https://doi.org/10.1186/gb-2008-9-9-234.

Conesa, S., and Götz, E. **2008**. Blast2GO: a comprehensive suite for functional analysis in plant genomics. Int. J. Plant Genom. Article ID:619832. https://doi.org/10.1155/2008/619832.

Couto, D., and Zipfel, C. **2016**. Regulation of pattern recognition receptor signalling in plants. Nat. Rev. Immunol. **16**:537-552. https://doi.org/10.1038/nri.2016.77.

Crisp, P., Crute, I.R., and Sutherland, R.A. **1989**. The exploitation of genetic resources of *Brassica oleracea* in breeding for resistance to clubroot (*Plasmodiophora brassicae*). Euphytica. **42**: 215-226. https://doi.org/10.1007/BF00034457.

Czerniawski, P. and Bednarek, P. **2018**. Glutathione S-transferases in the biosynthesis of sulfur-containing secondary metabolites in Brassicaceae plants. Front. Plant Sci. **9**:1639. https://doi.org/10.3389/fpls.2018.01639.

Dakouri, A., Lamara, M., Karim, M.M., Wang, J., Chen, Q., Gossen, B.D., et al. **2021**. Identification of resistance loci against new pathotypes of *Plasmodiophora brassicae* in *Brassica napus* based on genome-wide association mapping. Sci. Rep. **11**:6599. https://doi.org/10.1038/s41598-021-85836-9.

Dakouri, A., Zhang, X., Peng, G., Falk, K.C., Gossen, B.D., Strelkov, S.E., and Yu, F. **2018**. Analysis of genome-wide variants through bulked segregant RNA sequencing reveals a major gene for resistance to *Plasmodiophora brassicae* in *Brassica oleracea*. Sci. Rep. **8**(1): 17657. https://doi.org10.1038/ s41598-017-17765-5. PMID:29311619.

Delourme, R., Laperche, A., Bouchet, A.S., Jubault, M., Paillard, S., Manzanares-Dauleux, M.J., and Nesi, N. **2018**. Genes and quantitative trait loci mapping for major agronomic traits in *Brassica napus* L. The *Brassica napus* genome pp:41-85. https://doi.org/10.1007/978-3-319-43694-4_3.

Deora, A., Gossen, B.D., and McDonald M.R. **2012**. Infection and development of *Plasmodiophora brassicae* in resistant and susceptible canola cultivars. Can. J. Plant Sci. **34**(2):239-247. https://doi.org/10.1080/07060661.2012.681071.

Devos, S., Laukens, K., Deckers, P., Van Der Straeten, D., Beeckman, T., Inzé, D., Van Onckelen, H., Witters, E., and Prinsen, E. **2006**. A hormone and proteome approach to picturing the

initial metabolic events during *Plasmodiophora brassicae* infection on *Arabidopsis*. Mol. Plant Microbe Interact. **19**:1431-1443. https://doi.org/10.1094/MPMI-19-1431.

Diederichsen, E., and Sacristan, M. **1996**. Disease response of resynthesized *Brassica napus* L. lines carrying different combinations of resistance to *Plasmodiophora brassicae* Wor. Plant Breed. **115**:5-10. https://doi.org/10.1111/j.1439-0523. 1996.tb00862.x.

Diederichsen, E., Beckmann, J., Schondelmeier, J., and Dreyer, F. **2006**. Genetics of clubroot resistance in *Brassica napus* 'Mendel'. Acta Hortic. **706**:307-312. https://doi.org/10.17660/ActaHortic.2006.706.35.

Diederichsen, E., Frauen, M., and Ludwig-Müller, J. **2014**. Clubroot disease management challenges from a German perspective. Can. J. Plant Pathol. **36**:85-98. https://doi.org/10.1080/07060661.2013.861871.

Diederichsen, E., Frauen, M., Linders, E. G., Hatakeyama, K., and Hirai, M. **2009**. Status and perspectives of clubroot resistance breeding in crucifer crops. J. Plant Growth Regul. **28**: 265-281. https://doi.org/10.1007/s00344-009-9100-0.

Dixon, G. R. **2009**. The Occurrence and Economic Impact of *Plasmodiophora brassicae* and Clubroot Disease. *J. Plant Growth Regul.* **28**:194-202. https://doi.org/10.1007/s00344-009-9090-y.

Dixon, G. R. **2017**. Managing clubroot disease (caused by *Plasmodiophora brassicae* Wor.) by exploiting the interactions between calcium cyanamide fertilizer and soil microorganisms. J. Agr. Sci. **155**(4):527-543. https://doi.org/10.1017/S0021859616000800.

Dixon, G.R. **2012a**. Calcium cyanamide - a synoptic review of an environmentally benign fertiliser which enhances soil health. Acta Hortic. **938**:211-217. https://doi.org/10.17660/ActaHortic.2012.938.27.

Dixon, R.A. **2001**. Natural products and plant disease resistance. Nature. **411**:843-847. https://doi.org/10.1038/35081178.

Donald, C., and Porter, I. **2009**. Integrated control of clubroot. J. Plant Growth Regul. **28**(3):289-303. https://doi.org10.1007/ s00344-009-9094-7.

Donald, E.C., and Porter, I.J. **2014**. Clubroot in Australia: The history and impact of *Plasmodiophora brassicae* in *Brassica* crops and research efforts directed towards its control. Can. J. Plant Pathol. **36**:66-84. https://doi.org/10.1080/07060661.2013.873482.

Du, C., Hu, K., Xian, S., Liu, C., Fan, J., Tu, J., and Fu, T. **2016**. Dynamic transcriptome analysis reveals AP2/ERF transcription factors responsible for cold stress in rapeseed (*Brassica napus* L.). Mol. Genet. Genomics, **291**(3):1053-1067. https://doi.org/10.1007/s00438-015-1161-0.

Du, J., Wang, S., He, C., Zhou, B., Ruan, Y., and Shou, H. **2017**. Identification of regulatory networks and hub genes controlling soybean seed set and size using RNA sequencing analysis. *J. Exp. Bot.* **68**(8):1955-1972. https://doi.org/10.1093/jxb/erw460.

Englbrecht, C.C., Schoof, H., and Böhm, S. **2004**. Conservation, diversification and expansion of C_2H_2 zinc finger proteins in the *Arabidopsis thaliana* genome. BMC Genomics. **5**:39. https://doi. org/10.1186/1471-2164-5-39.

Engqvist, L.G. **1994**. Distribution of clubroot (*Plasmodiophora brassicae* Wor) in Sweden and the effect of infection on oil content of oilseed rape (*Brassica napus* L.). J. Swedish Seed Assoc. **104**(2):82-86.

Ernst, T.W., Kher, S., Stanton, D., Rennie, D.C., Hwang, S.F., and Strelkov, S.E. **2019**. *Plasmodiophora brassicae* resting spore dynamics in clubroot resistant canola (*Brassica napus*) cropping systems. Plant Pathol. **68**(2):399-408. https://doi.org/10.1111/ppa.12949.

Eulgem, T., and Somssich, I.E. **2007**. Networks of WRKY transcription factors in defense signaling. Curr. Opin. Plant Biol. **10**:366-371. https://doi.org/10.1016/j.pbi.2007.04.020.

Fang, F.C. **2011**. Antimicrobial actions of reactive oxygen species. MBio. **2**(5):e00141-11. https://doi.org/10.1128/mbio.00141-11**.**

Faraji, S., Rasouli, S.H., and Kazemitabar, S.K. **2018**. Genome-wide exploration of C2H2 zinc finger family in durum wheat (*Triticum turgidum* ssp. Durum): insights into the roles in biological processes especially stress response. Biometals, **10**:1007. https://doi.org/10.1007/s10534-018-0146-y.

Farid, M., Yang, R.C., Kebede, B., and Rahman, H. **2020**. Evaluation of *Brassica oleracea* accessions for resistance to *Plasmodiophora brassicae* and identification of genomic regions associated with resistance. Genome. **63**(2):91-101. https://doi.org/10.1139/gen-2019-0098.

Feurtado, J.A., Huang, D., Wicki-Stordeur, L., Hemstock, L.E., Potentier, M.S., Tsang, E.W., and Cutler, A.J. **2011**. The *Arabidopsis* C2H2 zinc finger INDETERMINATE DOMAIN1/ENHYDROUS promotes the transition to germination by regulating light and hormonal signaling during seed maturation. The Plant Cell. **23**(5):1772-1794. https://doi.org/10.1105/tpc.111.085134.

Feussner, I., and Polle, A. **2015**. What the transcriptome does not tell—proteomics and metabolomics are closer to the plants' patho-phenotype. Curr. Opin. Plant Biol. **26**:26-31. https://doi.org/10.1016/j.pbi.2015.05.023.

Figdore, S.S., Ferreira, M.E., Slocum, M.K., and Williams, P.H. **1993**. Association of RFLP markers with trait loci affecting clubroot resistance and morphological characters in *Brassica oleracea* L. Euphytica. **69**:33-44. https://doi.org/10.1007/BF00021723.

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

Fredua-Agyeman, R., and Rahman, H. **2016**. Mapping of the clubroot disease resistance in spring *Brassica napus* canola introgressed from European winter canola cv. "Mendel." Euphytica. **211**:201-213. https://doi.org/10.1007/s10681-016-1730-2.

Fredua-Agyeman, R., Hwang, S.F., Strelkov, S.E., Zhou, Q., Manolii, V.P., and Feindel, D. **2019**. Identification of Brassica accessions resistant to 'old' and 'new' pathotypes of *Plasmodiophora brassicae* from Canada. Plant Pathol. **68**(4):708-718. https://doi.org/10.1111/ppa.12980.

Fredua-Agyeman, R., Hwang, S.F., Strelkov, S.E., Zhou, Q., and Feindel, D. **2018**. Potential loss of clubroot resistance genes from donor parent *Brassica rapa* subsp. *rapifera* (ECD 04) during doubled haploid production. Plant Pathol. **67**(4):892-901. https://doi.org/10.1111/ppa.12816.

Fredua-Agyeman, R., Jiang, J., Hwang, S.F., and Strelkov, S.E. **2020**. QTL mapping and inheritance of clubroot resistance genes derived from *Brassica rapa* subsp. *rapifera* (ECD 02) reveals resistance loci and distorted segregation ratios in two F_2 populations of different crosses. Front. Plant Sci. **11**:899. https://doi.org/10.3389/fpls.2020.00899.

Friberg, H., Lagerlöf, J., and Rämert, B. **2005**. Germination of *Plasmodiophora brassicae* resting spores stimulated by a non-host plant. Eur. J. Plant Pathol. **113**(3):275-281. https://doi.org/10.1007/s10658-005-2797-0.

Friberg, H., Lagerlof, J., and Ramort, B. **2006**. Usefulness of non-host plants in managing *Plasmodiophora brassicae*. Plant Pathol. **55**:690-695. https://doi.org/10.1111/j.1365-3059.2006.01408.x.

Friedt, W., Tu, J., and Fu, T. **2018**. Academic and economic importance of *Brassica napus* rapeseed. The Brassica napus genome. Springer.

Fu, P., Piao, Y., Zhan, Z., Zhao, Y., Pang, W., Li, X., and Piao, Z. **2019**. Transcriptome arofile of *Brassica rapa* L. reveals the involvement of jasmonic acid, ethylene, and brassinosteroid signaling pathways in clubroot resistance. Agronomy. **9**(10):589. https://doi.org/10.3390/agronomy9100589.

Galindo-González, L., Manolii, V., Hwang, S.F., and Strelkov, S.E. **2020**. Response of *Brassica napus* to *Plasmodiophora brassicae* involves salicylic acid-mediated immunity: An RNA-seq-based study. Front. Plant Sci. **11**:1025. https://doi.org/10.3389/fpls.2020.01025.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., and Bairoch, A. 2005. Protein Identification and Analysis Tools on the Expasy Server. (In) John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press. Pp. 571-607.

Ge, J.C., Yoon, S.K., and Choi, N.J. **2017**. Using canola oil biodiesel as an alternative fuel in diesel engines: A review. Appl. Sci. **7**:881. https://doi.org/10.3390/app7090881.

Giri, J., Vij, S., Dansana, P. K., and Tyagi, A.K. **2011**. Rice A20/AN1 zinc-finger containing stress-associated proteins (SAP1/11) and a receptor-like cytoplasmic kinase (OsRLCK253) interact via A20 zinc-finger and confer abiotic stress tolerance in transgenic *Arabidopsis* plants. New Phytol. **191**(3):721-732. https://doi.org/10.1111/j.1469-8137.2011.03740.x.

Glazebrook, J. **2005**. Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. Annu. Rev. Phytopathol. **43**:205-227. https://doi.org/10.1146/annurev.phyto.43.040204.135923.

Gludovacz, T.V., Deora, A., McDonald, M.R., and Gossen B.D. **2014**. Cortical colonization by *Plasmodiophora brassicae* in susceptible and resistant cabbage cultivars. Eur. J. Plant Pathol. **140**:859-862. https://doi.org/10.1007/s10658-014-0492-8.

Göhre, V., Spallek, T., Häweker, H., Mersmann, S., Mentzel, T., Boller, T., et al. **2008**. Plant pattern-recognition receptor FLS2 is directed for degradation by the bacterial ubiquitin ligase AvrPtoB. Curr. Biol. **18**:1824-1832. https://doi.org/10.1016/j.cub.2008.10.063.

González-Lamothe, R., El-Oirdi, M., El-Rahman, T. A., Sansregret, R., Bathily, H., and Bouarab, K. **2009**. Cross talk between induced plant immune systems. In Molecular plant-microbe interactions (pp. 163-178). Wallingford UK: CABI.

Gossen, B.D., Adhikari, K.K.C., and McDonald M.R. **2012**. Effect of seeding date on development of clubroot in short-season *Brassica* crops. Can. J. Plant Pathol. **34**:516-523. https://doi.org/10.1080/07060661.2012.722129.

Gravot, A., Deleu, C., Wagner, G., Lariagon, C., Lugan, R., Todd, C., Wendehenne, D., Delourme, R., Bouchereau, A. and Manzanares-Dauleux, M.J. **2012**. Arginase induction represses gall development during clubroot infection in Arabidopsis. Plant Cell Physiol. **53**(5):901-911. https://doi.org/10.1093/pcp/pcs037.

Gupta, S. K., and Pratap, A. **2007**. History, origin, and evolution. Adv. Bot. Res. **45**:1-20. https://doi.org/10.1016/S0065-2296(07)45001-7.

Gururani, M.A., Venkatesh, J., Upadhyaya, C.P., Nookaraju, A., Pandey, S.K., and Park, S.W. **2012**. Plant disease resistance genes: Current status and future directions. Physiol Mol Plant Pathol. **78**:51-65. https://doi.org/10.1016/j.pmpp.2012.01.002.

Han, G., Qiao, Z., Li, Y., Wang, C., and Wang, B. **2021**. The roles of CCCH zinc-finger proteins in plant abiotic stress tolerance. Int. J. Mol. Sci. **22**(15):8327. https://doi.org/10.3390/ijms22158327.

Han, G., Yuan, F., Guo, J., Zhang, Y., Sui, N., and Wang, B. **2019b**. AtSIZ1 improves salt tolerance by maintaining ionic homeostasis and osmotic balance in Arabidopsis. Plant sci. **285**:55-67. https://doi.org/10.1016/j.plantsci.2019.05.002

Han, G.L., Wei, X., Dong, X., Wang, C., Sui, N., Guo, J.R. et al. **2019a**. Arabidopsis ZINC FINGER PROTEIN1 Acts Downstream of GL2 To Repress Root Hair Initiation and Elongation by Directly Suppressing bHLH Genes. Plant Cell. **32**:206-225. https://doi.org/10.1105/tpc.19.00226.

Han, X., and Kahmann, R. **2019**. Manipulation of phytohormone pathways by effectors of filamentous plant pathogens. Frontiers in Plant Science. **10**:822. https://doi.org/10.3389/fpls.2019.00822.

Hanano, A., Blee, E., and Murphy, D.J. **2023**. Caleosin/peroxygenases: multifunctional proteins in plants. *Annals of Botany.* **131**(3):387409. https://doi.org/10.1093/aob/mcad001.

Harvey, B.L., and Downey, R.K. **1964**. The inheritance of erucic acid content in rapeseed (*Brassica napus*). Can. J. Plant Sci. **44**:104-111. https://doi.org/10.4141/cjps64-019.

Hasan, M. J., Shaikh, R., Basu, U., and Rahman, H. **2021c**. Mapping clubroot resistance of *Brassica rapa* introgressed into *Brassica napus* and development of molecular markers for the resistance. Crop Sci. **61**:4112-4127. https://doi.org/10.1002/csc2.20626.

Hasan, J., Shaikh, R., Megha, S., Herrmann, D.T., Kebede, B., and Rahman, H. **2021b**. Mapping of flowering time, seed quality and clubroot resistance in rutabaga × spring canola

populations and their association. Euphytica. **217**(8):160. https://doi.org/10.1007/s10681021-02889-7.

Hasan, J., Megha, S., and Rahman, H. **2021a**. Clubroot in Brassica: recent advances in genomics, breeding, and disease management. Genome. **64**:735-760. https://dx.doi.org/10.1139/gen-2020-0089.

Hasan, M.J., and Rahman, H. **2016**. Genetics and molecular mapping of resistance to *Plasmodiophora brassicae* pathotypes 2, 3, 5, 6, and 8 in rutabaga (*Brassica napus* var. *napobrassica*). Genome. **59**:805-815. https://doi.org/10.1139/gen-2016-0034.

Hasan, M.J., Strelkov, S., Howard, R., and Rahman, H. **2012**. Screening of Brassica germplasm for resistance to *Plasmodiophora brassicae* pathotypes prevalent in Canada for broadening diversity in clubroot resistance. Can. J. Plant Sci. **92**:501-515. https://doi.org/10.4141/cjps2010-006.

Hashiba, T., Morita, S., Narisawa, K., and Usuki, K. **2003**. Mechanisms of symbiosis and disease suppression of root endophytic fungus. Nippon Bogeikagaku Kaishi- Japan Society for Bioscience, Biotechnology and Agrochemistry. **77**:130-133.

Hatakeyama, K., Fujimura, M., Ishida, M., and 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 Sci. **54**:197-201. https://doi.org/10.1270/jsbbs.54.197.

Hatakeyama, K., Niwa, T., Kato, T., Ohara, T., Kakizaki, T., and Matsumoto, S. **2017**. The tandem repeated organization of NB-LRR genes in the clubroot-resistant CRb locus in *Brassica rapa* L. Mol. Genet. Genomics. **292**(2):397-405. https://doi.org/10.1007/s00438-016-12811.

Hatakeyama, K., Suwabe, K., Tomita, R.N., Kato, T., Nunome, T., Fukuoka, H., and 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. https://doi.org/10.1371/journal. pone.0054745.

Hejna, O., Havlickova, L., He, Z., Bancroft, I., and Curn, V. **2019**. Analysing the genetic architecture of clubroot resistance variation in *Brassica napus* by associative transcriptomics. Mol. Breed. **39**(8):112. https://doi.org/10.1007/s11032-019-1021-4.

Hichri, I., Muhovski, Y., Žižková, E., Dobrev, P. I., Franco-Zorrilla, J. M., Solano, R., ... and Lutts, S. **2014**. The *Solanum lycopersicum* Zinc Finger2 cysteine-2/histidine-2 repressor-like transcription factor regulates development and tolerance to salinity in tomato and Arabidopsis. Plant Physiol. **164**(4):1967-1990. https://doi.org/10.1104/pp.113.225920.

Hirai, M. **2006**. Genetic analysis of clubroot resistance in Brassica crops. Breed. Sci. **56**:223-229. https://doi.org/10.1270/jsbbs.56.223.

Hirai, M., Harada, T., Kubo, N., Tsukada, M., Suwabe, K., and Matsumoto, S. **2004**. A novel locus for clubroot resistance in *Brassica rapa* and its linkage markers. Theor. Appl. Genet. **108**(4): 639-643. https://doi.org/10.1007/s00122-003-1475-x.

Hirani, A.H., Gao, F., Liu, J., Fu, G., Wu, C., McVetty, P.B.E., Duncan, R.W., and Li, G. **2018**. Combinations of independent dominant loci conferring clubroot resistance in all four turnip accessions (*Brassica rapa*) from the European clubroot differential set. Front. Plant Sci. **9**: 1628. https://doi.org/10.3389/fpls.2018.01628.

Hirani, A.H., Gao, F., Liu, J., Fu, G., Wu, C., Yuan, Y., Li, W., Hou, J., Duncan, R., and Li, G. **2016**. Transferring clubroot resistance from Chinese cabbage (*Brassica rapa*) to canola (*B. napus*). Can. J. Plant Pathol. **38**:82-90. https://doi.org/10.1080/07060661.2016.1141799.

Hollman, K. B., Hwang, S. F., Manolii, V. P., and Strelkov, S. E. **2021**. Pathotypes of *Plasmodiophora brassicae* collected from clubroot resistant canola (*Brassica napus* L.) cultivars in western Canada in 2017-2018. Can. J. Plant Pathol. **43**:622-630. https://doi.org/10.1080/07060661.2020.1851893.

Horton, P., Park, K.J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C.J., and Nakai, K. **2007**. WoLF PSORT: protein localization predictor. Nucleic Acids Res. **35**(2):W585-W587. https://doi.org/10.1093/nar/gkm259.

Howard, R.J., Strelkov, S.E., and Harding, M.W. **2010**. Clubroot of cruciferous crops- New perspectives on an old disease. Can. J. Plant Pathol. **32**:43-57. https://doi.org/10.1080/07060661003621761.

Hu, B., Jin, J., Guo, A.Y., Zhang, H., Luo, J., and Gao, G. **2015**. GSDS 2.0: an updated gene feature visualization server. Bioinformatics. **31**(8):1296-1297. https://doi.org/10.1093/bioinformatics/btu817.

Hu, X., Zhu, L., Zhang, Y., Xu, L., Li, N., Zhang, X., and Pan, Y. **2019**. Genome-wide identification of C2H2 zinc-finger genes and their expression patterns under heat stress in tomato (Solanum lycopersicum L.). Peer J. **7**:e7929. https://doi.org/10.7717/peerj.7929.

Huang, H., Ullah, F., Zhou, D.X., Yi, M., and Zhao, Y. **2019**. Mechanisms of ROS regulation of plant development and stress responses. Front. Plant Sci. **10**:800. https://doi.org/10.3389/fpls.2019.00800.

Huang, J., Sun, S. J., Xu, D. Q., Yang, X., Bao, Y. M., Wang, Z. F., ... and Zhang, H. **2009**. Increased tolerance of rice to cold, drought and oxidative stresses mediated by the overexpression of a gene that encodes the zinc finger protein ZFP245. Biochem. Biophy. Res. Co. **389**(3):556-561. https://doi.org/10.1016/j.bbrc.2009.09.032.

Huang, Z., Peng, G., Gossen, B.D., and Yu, F. 2019. Fine mapping of a clubroot resistance gene from turnip using SNP markers identified from bulked segregant RNA-Seq. Mol. Breeding. 39:131. https://doi.org/10.1007/s11032-019-1038-8.

Huang, Z., Peng, G., Liu, X., Deora, A., Falk, K.C., Gossen, B.D., McDonald, M.R., and Yu, F. 2017. Fine mapping of a clubroot resistance gene in Chinese cabbage using SNP markers identified from bulked segregant RNA sequencing. Front. Plant Sci. 8:1-9. https://doi.org/10.3389/fpls.2017.01448.

Hung, S., Umemura, T., Yamashiro, S., Slinger, S.J., and Holub, B.J. **1977**. The effects of original and randomized rapeseed oils containing high or very low levels of erucic acid on cardiac lipids and myocardial lesions in rats. Lipids. **12**:215221. https://doi.org/10.1007/BF02533297.

Hwang, S.F., Ahmed, H.U., Strelkov, S.E., Zhou, Q., Gossen, B.D., McDonald, M.R., Peng, G., and Turnbull, G.D. **2017a**. Suppression of clubroot by dazomet fumigant. Can. J. Plant Sci. **98**(1):172-182. https://doi.org/10.1139/cjps-2017-0099.

Hwang, S.F., Ahmed, H.U., Zhou, Q., Strelkov, S.E., Gossen, B.D., Peng, G., and Turnbull, G.D. **2014a**. Efficacy of Vapam fumigant against clubroot (*Plasmodiophora brassicae*) of canola. Plant Pathol. **63**(6):1374-1383. https://doi.org/10.1111/ppa.12207.

Hwang, S.F., Howard, R.J., Strelkov, H.E., Gossen, B.D., and Peng, G. **2014**. Management of clubroot (*Plasmodiophora brassicae*) on canola (*Brassica napus*) in western Canada. Can. J. Plant Pathol. **36**:49-65. https://doi.org/10.1080/07060661.2013.863806.

Hwang, S.F., Strelkov, S.E., Feng, J., Gossen, B.D., and Howard, R.J. **2012**. *Plasmodiophora brassicae*: a review of an emerging pathogen of the Canadian canola (*Brassica napus*) crop. Mol. Plant Pathol. **13**:105-113. https://doi.org/10.1111/j.1364-3703.2011.00729.x.

Iftikhar, R., Wang, X., and Rahman, H. **2018**. Broadening the genetic base of *Brassica napus* canola by interspecific crosses with different variants of *B. oleracea*. Euphytica. **214**:133. https://doi.org/10.1007/s10681-018-2213-4.

Irani, S., Trost, B., Waldner, M., Nayidu, N., Tu, J., Kusalik, A.J., Todd, C.D., Wei, Y. and Bonham-Smith, P.C. **2018**. Transcriptome analysis of response to *Plasmodiophora brassicae* infection in the Arabidopsis shoot and root. BMC Genomics. **19**:1-19. https://doi.org/10.1186/s12864-017-4426-7.

Iversen, G.R., and Gergen, M. **1997**. *Statistics: The Conceptual Approach*; Springer: NewYork, NY, USA, ISBN 13-978-1-4612-2244-6.

Jahn, L., Mucha, S., Bergmann, S., Horn, C., Staswick, P., Steffens, B., Siemens, J., and Ludwig-Müller, J. **2013**. The clubroot pathogen (*Plasmodiophora brassicae*) influences auxin signaling to regulate auxin homeostasis in Arabidopsis. Plants. **2**(4):726-749. https://doi.org/10.3390/plants2040726.

Ji, R., Gao, S., Bi, Q., Wang, Y., Lv, M., Ge, W., and Feng, H. **2021**. The salicylic acid signaling pathway plays an important role in the resistant process of *Brassica rapa* L. ssp. pekinensis to *Plasmodiophora brassicae* Woronin. Journal of Plant Growth Regulation. **40**:405-422. https://doi.org/10.1007/s00344-020-10105-4.

Ji, R., Wang, Y., Wang, X., Liu, Y., Shen, X., and Feng, H. **2018**. Proteomic analysis of the interaction between *Plasmodiophora brassicae* and Chinese cabbage (*Brassica rapa* L. ssp. pekinensis) at the initial infection stage. Sci. Hortic. **233**:386-393. https://doi.org/ 10.1016/j.scienta.2018.02.006.

Jia, H., Wei, X., Yang, Y., Yuan, Y, Wei, F., and Zhao, Y., et al. **2017**. Root RNA-seq analysis reveals a distinct transcriptome landscape between clubroot-susceptible and clubroot-

resistant Chinese cabbage lines after *Plasmodiophora brassicae* infection. Plant Soil. **421**:93-105. https://doi.org/10.1007/s11104-017-3432-5.

Jiang, J., Fredua-Agyeman, R., Hwang, S.F., and Strelkov, S.E. **2021**. Differentially expressed genes in canola (*Brassica napus*) during infection by virulent and avirulent *Plasmodiophora brassicae* pathotypes. Plant Pathol. **70**:50-60. https://doi.org/10.1111/ppa.13267.

Johnston, T. **1970**. A new factor for resistance to club root in *Brassica napus* L. Plant Pathol. **19**(4):156-158. doi:10.1111/ j.1365-3059. 1970.tb01007. x.

Jones, J.D., and Dangl, J.L. **2006**. The plant immune system. Nature. **444**(7117):323. https://doi.org/10.1038/nature05286.

Jönsson, R. **1978**. Breeding for improved oil and meal quality in rape (*Brassica napus* L.) and turnip rape (*Brassica campestris* L.). Hereditas. **87**:205-218. https://doi.org/10.1111/j.1601-5223. 1978.tb01264.x.

Jorrin-Novo, J.V. **2014**. Plant proteomics methods and protocols (pp. 3-13). Humana Press.

Joseph, M.P., Papdi, C., Kozma-Bognár, L., Nagy, I., López-Carbonell, M., Rigó, G., Koncz, C., and Szabados, L. **2014**. The Arabidopsis ZINC FINGER PROTEIN3 Interferes with Abscisic Acid and Light Signaling in Seed Germination and Plant Development. *Plant Physiol.* **165**(3):1203-1220. https://doi.org/10.1104/pp.113.234294.

Joshi, R.K., Megha, S., Rahman, M.H., Basu, U., and Kav, N.N. **2016**. A global study of transcriptome dynamics in canola (*Brassica napus* L.) responsive to *Sclerotinia sclerotiorum* infection using RNA-Seq. Gene. **590**(1):57-67. https://doi.org/10.1016/j.gene.2016.06.003.

Jr, G. D., Sherman, B. T., Hosack, D. A., Yang, J., Gao, W., Lane, H. C., et al. **2003**. DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biology. https://doi.org/10.1186/gb-2003-4-9-r60.

Kageyama, K., and Asano, T. **2009**. Life cycle of *Plasmodiophora brassicae*. J. Plant Growth Regul. **28**(3):203-211. https://doi.org/10.1007/ s00344-009-9101-z.

Kam, J., Gresshoff, P.M., Shorter, R., and Xue, G.P. **2008**. The Q-type C2H2 zinc finger subfamily of transcription factors in *Triticum aestivum* is predominantly expressed in roots and enriched with members containing an EAR repressor motif and responsive to drought stress. Plant Mol. Biol. **67**(3):305-322. https://doi.org/10.1007/s11103-008-9319-3.

Kanehisa, M., and Goto, S. **2000**. KEGG: kyoto encyclopedia of genes and genomes[J]. Nucleic acids research. **28**(1):27-30. https://doi.org/10.1093/nar/28.1.27.

Karim, M.M., Dakouri, A., Zhang, Y., Chen, Q., Peng, G., Strelkov, S.E., Gossen, B.D., and Yu, F. **2020**. Two clubroot-resistance genes, Rcr3 and Rcr9wa, mapped in *Brassica rapa* using bulk segregant RNA sequencing. Int. J. Mol. Sci. **21**(14):1-15. https://doi.org/10.3390/ijms21145033.

Karling, J. **1969**. Plasmodiophorales. 2nd ed. Hafner Publishing Co Ltd., New York.

Kato, T., Hatakeyama, K., Fukino, N., and Matsumoto, S. 2012. Identification of a clubroot resistance locus conferring resistance to a *Plasmodiophora brassicae* classified into pathotype group 3 in Chinese cabbage (*Brassica rapa* L.). Breed. Sci. 62:282-287. https://doi.org/10.1270/jsbbs.62.282.

Kato, T., Hatakeyama, K., Fukino, N., and Matsumoto, S. **2013**. Fine mapping of the clubroot resistance gene *CRb* and development of a useful selectable marker in *Brassica rapa*. Breed. Sci. **63**(1):116-24. https://doi.org/10.1270/jsbbs.63.116.

Kaur, K., Liu, Y., and Rahman, H. **2022**. Introgression of resistance to multiple pathotypes of *Plasmodiophora brassicae* from turnip (*Brassica rapa* ssp. *rapifera*) into spring *B. napus* canola. Agronomy. **12**:1225. https://doi.org/10.3390/agronomy12051225.

Kawai, Y., Ono, E., and Mizutani, M. **2014**. Evolution and diversity of the 2–oxoglutarate-dependent dioxygenase superfamily in plants. The Plant Journal. **78**(2):328-343. https://doi.org/10.1111/tpj.12479.

Kiełbowicz-Matuk, A. **2012**. Involvement of plant C2H2-type zinc finger transcription factors in stress responses. Plant Sci. **185**:78-85. https://doi.org/10.1016/j.plantsci.2011.11.015.

Kim S.H., Hong J.K., Lee S.C., Sohn K.H., Jung H.W., and Hwang B.K. **2004**. *CAZFP1*, Cys2/His2-type zinc-finger transcription factor gene functions as a pathogen-induced early-defense gene in *Capsicum annuum*. Plant Mol. Biol. **55**:883-904. https://doi.org/10.1007/s11103-004-2151-5.

Kim, D., Langmead, B. and Salzberg, S. **2015**. HISAT: a fast spliced aligner with low memory requirements. Nat Methods**. 12**:357-360. https://doi.org/10.1038/nmeth.3317.

Kim, J.C., Lee, S.H., Cheong, Y.H., Yoo, C.M., Lee, S.I., Chun, H.J., ... and Cho, M.J. **2001**. A novel cold-inducible zinc finger protein from soybean, SCOF-1, enhances cold tolerance in transgenic plants. The Plant Journal. **25**(3):247-259. https://doi.org/10.1046/j.1365-313x.2001.00947.x.

Kim, K.H., Kang, Y.J., Kim, D.H., Yoon, M.Y., Moon, J.K., Kim, M.Y., Van, K., and Lee, S.H. **2011**. RNA-seq analysis of a soybean near-isogenic line carrying bacterial leaf pustule-resistant and -susceptible alleles. DNA Research. **18**(6):483-497. https://doi.org/10.1093/dnares/dsr033.

Kim, S.H., Hong, J.K., Lee, S.C., Sohn, K.H., Jung, H.W., and Hwang, B.K. **2004**. *CAZFP1*, Cys2/His2-type zinc-finger transcription factor gene functions as a pathogen-induced early-defense gene in *Capsicum annuum*. Plant Mol. Biol. **55**:883-904. https://doi.org/10.1007/s11103-004-2151-5.

Kissen, R., and Bones, A. M. **2009**. Nitrile-specifier proteins involved in glucosinolate hydrolysis in *Arabidopsis thaliana*. Journal of Biological Chemistry. **284**(18):12057-12070. https://doi.org/10.1074/jbc.M807500200.

Knepper, C., Savory, E.A. and Day, B. **2011**. Arabidopsis NDR1 is an integrin-like protein with a role in fluid loss and plasma membrane-cell wall adhesion. Plant physiol. **156**(1):286-300. https://doi.org/10.1104/pp.110.169656.

Kobe, B., and Kajava, A.V. **2001**. The leucine-rich repeat as a protein recognition motif. Curr Opin Struct Biol. **11**:725-732. https://doi.org/10.1016/S0959-440X(01)00266-4.

Koboldt, D.C., Zhang, Q., Larson, D.E., Shen, D., Mclellan, M.D., Lin, L., Miller, C.A., Mardis, E.R., Ding, L., and Wilson, R.K. **2012**. VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Research. **22**:568-576. http://genome.cshlp.org/cgi/doi/10.1101/gr.129684.111.

Koornneef, A. and Pieterse, C.M. **2008**. Cross talk in defense signaling. Plant physiology. **146**(3): 839-844. https://doi.org/10.1104/pp.107.112029.

Korbas, M., Jajor, E., and Budka, A. **2009**. Clubroot (*Plasmodiophora brassicae*) - a threat for oilseed rape. J. Plant Protect. Res. **49**:446-451. https://doi.org/10.2478/v10045- 009- 0071-78.

Kosambi, D.D. **1943**. The estimation of map distances from recombination values. Annals of Eugenics. **12**:172-175. https://doi.org/10.1007/978-81-322-3676-4_16.

Koscielny, C.B., Hazebroek, J., and Duncan, R.W. **2018**. Phenotypic and metabolic variation among spring *Brassica napus* genotypes during heat stress. Crop and Pasture Sci. **69**(3):284-295. https://doi.org/10.1071/CP17259.

Kotchoni, S.O., and Gachomo, E.W. **2009**. A rapid and hazardous reagent free protocol for genomic DNA extraction suitable for genetic studies in plants. Molecular Biology Reports. **36**:633-1636. https://doi.org/10.1007/s11033-008-9362-9.

Kubo, K.I., Sakamoto, A., Kobayashi, A., Rybka, Z., Kanno, Y., Nakagawa, H. et al. **1998**. Cys2/His2 zinc-finger protein family of petunia: evolution and general mechanism of target-sequence recognition. Nucleic Acids Res. **26**:608-615. https://doi.org/10.1093/nar/26.2.608.

Kuginuki, Y., Yoshikawa, H., and Hirai, M. **1999**. Variation in virulence of *Plasmodiophora brassicae* in Japan tested with clubroot-resistant cultivars of Chinese cabbage (*Brassica rapa* L. ssp. *pekinensis*). Eur. J. Plant Pathol. **105**:327-332. https://doi.org/10.1023/A:1008705413127.

Kumar, S., Stecher, G., and Tamura, K. **2016**. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. **33**(7):1870-1874. https://doi.org/10.1093/molbev/msw054.

Kumari, P., Kumari, R., Sharma, B., Prasad, S.K., and Singh, R.K. **2019**. Abiotic stress response of field crops: Recent approach. Int. J. Curr. Microbiol. App. Sci. **8**:1761-1769. https://doi.org/10.20546/ijcmas.2019.804.205.

Kundu, A., Das, S., Basu, S., Kobayashi, Y., Kobayashi, Y., Koyama, H., and Ganesan, M. **2019**. GhSTOP1, a C2H2 type zinc finger transcription factor is essential for Aluminum and proton stress tolerance and lateral root initiation in cotton. Plant Biol. **21**(1):35-44. https://doi.org/10.1111/plb.12895.

Lahlali, R., McGregor, L., Song, T., Gossen, B.D., Nariswa, K., and Peng. G. **2014**. *Heteroconium chaetospira* induces resistance to clubroot via upregulation of host genes involved in

jasmonic acid, ethylene and auxin biosynthesis. PLoS ONE. **9**: e94144. https://doi.org/10.1371/journal.pone.0094144.

Laila, R., Park, J., Robin, A.H.K., Natarajan, S., Vijayakumar, H., Shirasawa, K., Isobe, S., Kim, H., and Nou, I. **2019**. Mapping of a novel clubroot resistance QTL using ddRAD-seq in Chinese cabbage (*Brassica rapa* L.). BMC Plant Biol. **19**:13-21. https://doi.org/10.1186/s12870-018-1615-8.

Laity, J.H., Lee, B.M., and Wright, P.E. **2001**. Zinc finger proteins: new insights into structural and functional diversity. Curr. Opin. Struc Biol. **11**(1):39-46. https://doi.org/10.1016/S0959-440X(00)00167-6.

Laloi, C., Mestres-Ortega, D., Marco Y, Meyer Y, and Reichheld, J.P. **2004**. The Arabidopsis cytosolic thioredoxin h5 gene induction by oxidative stress and its W-box-mediated response to pathogen elicitor. Plant Physiol. **134**(3):1006-1016. https://doi.org/10.1104/pp.103.035782.

Lan, M., Li, G., Hu, J., Yang, H., Zhang, L., Xu, X., Liu, J., He, J., and Sun, R. **2019**. iTRAQ-based quantitative analysis reveals proteomic changes in Chinese cabbage (*Brassica rapa* L.) in response to *Plasmodiophora brassicae* infection. Sci. Rep. **9**(1):12058. https://doi.org/10.1038/s41598-019-48608-0.

Landry, BS., Hubert, N., Crete, R., Chang, M.S., Lincoln, S.E., and Etoh, T. **1992**. A genetic map for *Brassica oleracea* based on RFLP markers detected with expressed DNA sequences and mapping of resistance genes to race 2 of *Plasmodiophora brassicae* (Woronin). Genome. **35**:409-420. https://doi.org/10.1139/g92-061.

Langridge, P., and Fleury, D. **2011**. Making the most of 'omics' for crop breeding. Trends Biotech. **29**(1):33-40. https://doi.org/10.1016/j.tibtech.2010.09.006

Laperche, A., Aigu, Y., Jubault, M., Ollier, M., Guichard, S., Glory, P., et al. **2017**. Clubroot resistance QTL are modulated by nitrogen input in *Brassica napus*. Theor. Appl. Genet. **130**(4):669-684. doi:10.1007/s00122-016-2842-8.

Latgé, J.P. and Beauvais, A. **2014**. Functional duality of the cell wall. Current opinion in microbiology. **20**:111-117. https://doi.org/10.1016/j.mib.2014.05.009.

Laurens, F., and Thomas, G. **1993**. Inheritance of resistance to clubroot (*Plasmodiophora brassicae Wor.*) in kale (*Brassica oleracea* ssp. *acephala*). Hereditas. **119**:253-262. https://doi.org/10.1111/j.1601-5223.1993.00253.x.

Lawrence, S.D., and Novak, N.G. **2018**. Comparative analysis of the genetic variability within the Q-type C2H2 zinc-finger transcription factors in the economically important cabbage, canola and Chinese cabbage genomes. Hereditas. **155**:29. https://doi.org/10.1186/ s41065-018-0065-5.

LeBoldus, J.M., Manolii, V.P., Turkington, T.K., and Strelkov, S.E. **2012**. Adaptation to *Brassica* host genotypes by a single spore isolate and population of *Plasmodiophora brassicae* (clubroot). Plant Dis. **96**:833-838. https://doi.org/10.1094/PDIS-09-11-0807.

LeClere, S., Tellez, R., Rampey, R. A., Matsuda, S. P., and Bartel, B. **2002**. Characterization of a family of IAA- amino acid conjugate hydrolases from Arabidopsis. J. Biol. Chem. **277**:20446–20452. https://doi.org/10.1074/jbc.M111955200.

Lee, G.P., Baek, N.K., and Park, K.W. **2002**. Genetic mapping of resistant genes in *Brassica pekinensis* against *Plasmodiophora brassicae* race 6. Plant Pathol. J. **18**(5):266-270. https://doi.org/10.5423/ PPJ.2002.18.5.266.

Lee, J., Izzah, N.K., Choi, B.S.S., Joh, H.J., Lee, S.C.C., Perumal, S., et al. **2016**. Genotyping-by-sequencing map permits identification of clubroot resistance QTLs and revision of the

reference genome assembly in cabbage (*Brassica oleracea* L.). DNA Res. **23**(1):29-41. https://doi.org/10.1093/dnares/dsv034. PMID:26622061.

Leflon, M., Eber, F., Letanneur, J.C., Chelysheva, L., Coriton, O., Huteau, V., Ryder, C.D., Barker, G., Jenczewski, E., and Chèvre, A.M. **2006**. Pairing and recombination at meiosis of *Brassica rapa* (AA) x *Brassica napus* (AACC) hybrids. Theor. Appl. Genet. **113**:1467-1480. https://doi.org/10.1007/s00122-006-0393-0.

Lescot, M., Déhais, P., and Thijs, G. **2002**. PlantCARE, a database of plant *cis*-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Res. **30**:325-327. https://doi.org/10.1093/nar/30.1.325.

Letunic, I., Khedkar, S., and Bork, P. **2020**. SMART: recent updates, new developments and status in 2020. Nucleic Acids Res. **49**(D1):D458-D460. https://doi.org/10.1093/nar/gkaa937.

Leutert, M., Rodríguez-Mias, R.A., Fukuda, N.K., and Villén, J. **2019**. R2-P2 rapid-robotic phosphoproteomics enables multidimensional cell signaling studies. Mol Syst Biol. **15**(12):e9021. https://doi.org/10.15252/msb.20199021.

Levy, M., Wang, Q., Kaspi, R., Parrella, M.P., and Abel, S. **2005**. Arabidopsis IQD1, a novel calmodulin-binding nuclear protein, stimulates glucosinolate accumulation and plant defense. Plant J. **43**(1):79-96. https://doi.org/10.1111/j.1365-313X.2005.02435.x.

Li, Q., Shah, N., Zhou, X., Wang, H., Yu, W., Luo, J., Liu, Y., Li, G., Liu, C., Zhang, C., Chen, P. **2021**. Identification of micro-ribonucleic acids and their targets in response to *Plasmodiophora brassicae* infection in *Brassica napus*. Front. Plant Sci. **12**:734419. https://doi.org/10.3389/fpls.2021.734419.

Li, B.B., Liu, Y.G., Tao, W.U., Wang, J.P., Xie, G.R., Chu, Z.H., and Ding, X.H. **2019**. OsBGLU19 and OsBGLU23 regulate disease resistance to bacterial leaf streak in rice. J. Integr. Agr. **18**(6):1199-1210. https://doi.org/10.1016/S2095-3119(18)62117-3.

Li, H., and Durbin, R. **2009a**. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. **25**:1754-1760. https: //doi.org/10.1093/bioinformatics/btp324.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. **2009b**. The Sequence Alignment/Map format and SAMtools. Bioinformatics. **25**: 2078-2079. https://doi.org/10.1093/bioinformatics/btp352.

Li, H., Ye, G., and Wang, J. **2007**. A modified algorithm for the improvement of composite interval mapping. Genetics. **175**:361-374. https://doi.org/10.1534/genetics.106. 066811.

Li, T., Gonzalez, N., Inze, D., and Dubois, M. **2020**. Emerging connections between small RNAs and phytohormones. Trends Plant Sci. **25**:912-929. https://doi.org/10.1016/j.tplants.2020.04.004.

Liao, J., Luo, L., Zhang, L., Wang, L., Shi, X., Yang, H., and et al. **2021**. Comparison of the effects of three fungicides on clubroot disease of tumorous stem mustard and soil bacterial community. J. Soils Sediments. **22**:256-271. https://doi.org10.1007/s11368-021-03073-z.

Liao, Y., Smyth, G.K., and Shi, W. **2014**. FeatureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. **30**(7):923-30. https://doi.org/10.1093/bioinformatics/btt656.

Linthorst, H.J. and Van Loon, L.C. **1991**. Pathogenesis-related proteins of plants. Crit. Rev. Plant Sci. **10**(2):123-150. https://doi.org/10.1080/07352689109382309.

Liu, H., Lyu, H.M., Zhu, K., Van de Peer, Y., and Max, Cheng, Z.M. **2021**. The emergence and evolution of intron-poor and intronless genes in intron-rich plant gene families. Plant J. **105**(4):1072-1082. https://doi.org/ 10.1111/tpj.15088.

Liu, J., Fu, C., Li, G., Khan, M. N., and Wu, H. **2021**. ROS homeostasis and plant salt tolerance: plant nanobiotechnology updates. Sustainability. **13**:3552. https://doi.org/10.3390/su13063552.

Liu, J.J., Sturrock, R. and Ekramoddoullah, A.K. **2010**. The superfamily of thaumatin-like proteins: its origin, evolution, and expression towards biological function. Plant cell reports. **29**:419-436. https://doi.org/10.1007/s00299-010-0826-8.

Liu, L., Qin, L., Zhou, Z., Hendriks, Wilhelmina G.H.M., Liu, S., and Wei, Y. **2020**. Refining the life cycle of *Plasmodiophora brassicae*. Phytopathol. **110**:1704-1712. https://doi.org/10.1094/phyto- 02- 20- 0029- r.

Liu, Q., Wang, Z., Xu, X., Zhang, H., and Li, C. **2015**. Genome-wide analysis of C2H2 Zinc-finger family transcription fac- tors and their responses to abiotic stresses in poplar (*Populus trichocarpa*). PLoS One. **10**: e0134753. https://doi.org/10.1371/journal.pone.0134753.

Liu, S., Fan, C., Li, J., Cai, G., Yang, Q., Wu, J., Yi, X., Zhang, Z.Y. **2016**. A genome-wide association study reveals novel elite allelic variations in seed oil content of *Brassica napus*. Theor. Appl. Genet. **129**:1203-1215. https://doi.org/10.1007/s00122-016-2697-z.

Liu, Y., Khan, A.R., and Gan, Y. **2022**. C2H2 zinc finger proteins response to abiotic stress in plants. Int. J. Mol. Sci. **23**(5):2730. https://doi.org/10.3390/ijms23052730.

Liu, Y., Lu, S., Liu, K. et al. **2019**. Proteomics: a powerful tool to study plant responses to biotic stress. Plant Methods. **15**:135. https://doi.org/10.1186/s13007-019-0515-8.

Livak, K.J., and Schmittgen, T.D. **2001**. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods. **25**:402-408. https://doi.org/10.1006 /meth.2001.1262.

Love, M.I., Huber, W., and Anders, S. **2014**. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome biology. **15**(12):1-21. https://doi.org/10.1186/s13059-014-0550-8.

Lu, D., Lin, W., Gao, X., Wu, S., Cheng, C., and Avila, J., et al. **2011**. Direct ubiquitination of pattern recognition receptor FLS2 attenuates plant innate immunity. Science. **332**:1439-1442. https://doi.org/10.1126/science.1204903.

Ludwig-Müller, J., Prinsen, E., Rolfe, S.A., and Scholes, J.D. **2009**. Metabolism and plant hormone action during clubroot disease. J. Plant Growth Regul. **28**:229-244. https://doi.org/10.1007/s00344-009-9089-4.

Luo, X., Bai, X., Zhu, D., Li, Y., Ji, W., Cai, H. et al. **2012**. GsZFP1, a new Cys2/His2-type zinc-finger protein, is a positive regulator of plant tolerance to cold and drought stress. Planta. **235**:1141-55. https://doi.org/ 10.1007/s00425-011-1563-0.

Luo, Y.X., Luo, C.Y., Du, D.Z., Fu, Z., Yao, Y.M., Xu, C.C., and Zhang, H.S. **2014**. Quantitative trait analysis of flowering time in spring rapeseed (*B. napus* L.). Euphytica. **200**:321-335. https://doi.org/10.1007/s10681-014-1140-2.

Lysak, M.A., Cheung, K., Kitschke, M., and Bureš, P. **2007**. Ancestral chromosomal blocks are triplicated in *Brassiceae* species with varying chromosome number and genome size. Plant Physiol. **145**:402-410. https:// doi.org/10.1104/pp.107.104380.

Lyu, T., and Cao, J. **2018**. Cys2/His2 zinc-finger proteins in transcriptional regulation of flower development. Int. J. Mol. Sci. **19**(9):2589. https://doi.org/10.3390/ijms19092589.

Mandaˊkovaˊ, T., and Lysak, M.A. **2008**. Chromosomal phylogeny and karyotype evolution in x = 7 crucifer species (*Brassicaceae*). Plant Cell. **20**:2559-2570. https://doi.org/10.1105/tpc.108.062166.

Manichaikul, A., Moon, J.Y., Sen, S.ˊ, Yandell, B.S., Broman, K.W. **2009**. A model selection approach for the identification of quantitative trait loci in experimental crosses, allowing epistasis. Genetics. **181**:1077-1086. https://doi.org/10.1534/genetics.108.094565.

Manzanares-Dauleux, M., Divaret, I., Baron, F., and Thomas, G. **2001**. Assessment of biological and molecular variability between and within field isolates of *Plasmodiophora brassicae*. Plant Pathol. **50**(2):165-173. doi:10.1046/j.1365-3059.2001.00557.x.

Manzanares-Dauleux, M.J., Delourme, R., and Baron, F. **2000**. Mapping of one major gene and of QTLs involved in resistance to clubroot in *Brassica napus*. Theor. Appl. Genet. **101**:885-891. https://doi.org/10.1007/s001220051557.

Marioni, J.C., Mason, C.E., Mane, S.M., Stephens, M., and Gilad, Y. **2008**. RNA-Seq: An assessment of technical reproducibility and comparison with gene expression assays. Genome Res. **18**: 1509-1517. https://doi.org/10.1101/gr.079558.108

Matschi, S., Katharina, H., Marco, H., Bettina, H., and Tina, R. **2015**. "The calcium-dependent protein kinase CPK28 regulates development by inducing growth phase-specific, spatially restricted alterations in jasmonic acid levels independent of defense responses in Arabidopsis." The Plant Cell. **27**(3):591-606. https://doi.org/10.1105/tpc.15.00024.

Matsumoto, E., Ueno, H., Aruga, D., Sakamoto, K., and Hayashida, N. **2012**. Accumulation of three clubroot resistance genes through marker-assisted selection in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*). J. Jpn. Soc. Hortic. Sci. **81**(2):184-190. https://doi.org/10.2503/jjshs1.81.184.

Matsumoto, E., Yasui, C., Ohi, M., and Tsukada, M. 1998. Linkage analysis of RFLP markers for clubroot resistance and pigmentation in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*). Euphytica. 104:79. https://doi.org/10.1023/A:1018370418201.

Mawson, R., Heaney, R.K., Zdunczyk, Z., and Kozlowska, H. **1993**. Rapeseed meal-glucosinolates and their antinutritional effects. Part II. Flavour and palatability. Die Nahrung. **37**:336-344. https://doi.org/10.1002/food.19930370405.

McDonald, M.R., Kornatowska, B., and McKeown, A.W. **2004**. Management of clubroot of Asian Brassica crops grown on organic soils. Acta Hortic. **635**:25-30. http://hdl.handle.net/10214/17918.

Megha, S., Basu, U., Joshi, R.K., and Kav, N.N.V. **2018**. Physiological studies and genome-wide microRNA profiling of cold-stressed *Brassica napus.* Plant Physiol. Bioch. **132**:1-17. https://doi.org/10.1016/j.plaphy.2018.08.027.

Megha, S., Wang, Z., Kav, N.N., and Rahman, H. **2022**. Genome-wide identification of Biotin Carboxyl Carrier subunits of Acetyl-CoA Carboxylase in Brassica and their role in stress tolerance in oilseed *Brassica napus*. BMC Genomics. **23**:107. https://doi.org/10.1186/s12864-022-08920-y.

Mehta D., Scandola S., and Uhrig R.G. **2022**. BoxCar and Library-free data-independent acquisition substantially improve the depth, range, and completeness of label-free quantitative proteomics. Anal. Chem. **94**(2):793-802. https://doi.org/10.1021/acs.analchem.1c03338.

Mei J., Guo Z., Wang J., Feng Y., Ma G., Zhang C., Qian W., and Chen G. **2019**. Understanding the resistance mechanism in *Brassica napus* to clubroot caused by *Plasmodiophora brassicae*. Phytopathol. **109**(5):810-818. https://doi.org/10.1094/PHYTO-06-18-0213-R.

Meng, L., Li, H., Zhang, L., and Wang, J. **2015**. QTL IciMapping: Integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. Crop J. **3**:269-283. https://doi.org/10.1016/j.cj.2015.01.001.

Meng, X., Zhou, J., and Sui, N. **2018**. Mechanisms of salt tolerance in halophytes: current understanding and recent advances. Open Life Sci. **13**:149-154. https://doi.org/10.1515/biol-2018-0020.

Meyers, B.C., Kaushik, S., and Nandety, R.S. **2005**. Evolving disease resisatnce genes. Curr. Opin. Plant Biol. **8**:129-134. https://doi.org/10.1016/j.pbi.2005.01.002.

Michelmore, R.W., Paran, I., and 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. PNAS. **88**(21):9828-9832. https://doi.org/10.1073/pnas.88.21.9828.

Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G.A., Sonnhammer, E.L.L., Tosatto, S.C.E., Paladin, L., Raj, S., Richardson, L.J., Finn, R.D., and Bateman, A. **2020**. Pfam: The protein families database in 2021. https://doi.org/10.1093/nar/gkaa913.

Moon, J., Kim, S., Choi, G., Kwon, S.Y., Cho, H.S., Kim, H.S., Moon, J.S., and Park, J.M. **2020**. Comparative proteomic analysis of host responses to *Plasmodiophora brassicae* infection in susceptible and resistant *Brassica oleracea*. Plant Biotechnol. Rep. **14**:263-274. https://doi.org/10.1007/s11816-020-00596-8.

Moriguchi, K., Kimizuka-Takagi, C., Ishii, K., and Nomura, K. **1999**. A genetic map based on RAPD, RFLP, Isozyme, morphological markers and QTL analysis for clubroot resistance in *Brassica oleracea*. Breed. Sci. **49**:257-265. https://doi.org/10.1270/jsbbs.49.257.

Morita, M.T., Sakaguchi, K., Kiyose, Si., Taira, K., Kato, T., Nakamura, M. et al. **2006**. A C2H2-type zinc finger protein, SGR5, is involved in early events of gravitropism in *Arabidopsis* inflorescence stems. Plant J. **47**:619-28. https://doi.org/10.1111/j.1365-313X.2006.02807.x.

Moxham, S.E., Fraser, R., and Buczacki, S. **1983**. Spore wall proteins of *Plasmodiophora brassicae*. T. Brit. Mycol. Soc. **80**(3): 497-506. https://doi.org/10.1016/S0007-1536(83)80046-1.

Müller, M., and Munné-Bosch, S. **2015**. Ethylene Response Factors: A Key Regulatory Hub in Hormone and Stress Signaling. Plant Physiol. **169**(1):32-41. https://doi.org/10.1104/pp.15.00677.

Nagaoka, T., Doullah, M.A.U., Matsumoto, S., Kawasaki, S., Ishikawa, T., Hori, H., and 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*. Theor. Appl. Genet. **120**:1335-1346. https://doi.org/10.1007/s00122-010-1259-z.

Naiki, T., Kawaguchi, C., and Ikegami, H. **1984**. Root hair reinfection in Chinese cabbage seedlings by the secondary zoospores of *Plasmodiophora brassicae* Woronin. Ann. Phytopathol. Soc. Jpn. **50**:216-220. https://doi.org/10.3186/jjphytopath.50.216.

Narisawa, K., Shimura, M., Usuki, F., Fukuhara, S., and Hashiba, T. **2005**. Effects of pathogen density, soil moisture, and soil pH on biological control of clubroot in Chinese cabbage by *Heteroconium chaetospira*. Plant Dis. **89**:285-290. https://doi.org/10.1094/PD-89-0285.

Neuhauser, S., Bulman, S., and Kirchmair, M. **2010**. Plasmodiophorids: the challenge to understand soil-borne, obligate biotrophs with a multiphasic life cycle. Molecular identification of fungi. 51-78. https://doi.org/10.1007/978-3-642-05042-8_3.

Neuhauser, S., Kirchmair, M., and Gleason, F.H. **2011**. Ecological roles of the parasitic phytomyxids (plasmodiophorids) in marine ecosystems - A review. Mar. Freshwater Res. **62**:365-371. https://doi.org/10.1071/MF10282.

Ng, D.W.K., Abeysinghe, J.K., and Kamali, M. **2018**. Regulating the regulators: the control of transcription factors in plant defense signaling. Int. J. Mol. Sci. **19**:3737. https://doi.org/10.3390/ijms19123737.

Nguyen, M.L., Monakhos, G.F., Komakhin, R.A., and Monakhos, S.G. **2018**. The New Clubroot resistance locus is located on chromosome A05 in Chinese cabbage (*Brassica rapa* L.). Russ. J. Genet. **54**(3):296-304. https://doi.org/10.1134/ S1022795418030080.

Ning, Y., Wang, Y., Fang, Z., Zhuang, M., Zhang, Y., Lv, H., ... and Yang, L. **2019**. Comparative transcriptome analysis of cabbage (Brassica oleracea var. capitata) infected by *Plasmodiophora brassicae* reveals drastic defense response at secondary infection stage. Plant and Soil. **443**:167-183. https://doi.org/10.1007/s11104-019-04196-6.

Nitta, Y., Ding, P. and Zhang, Y. **2014**. Identification of additional MAP kinases activated upon PAMP treatment. Plant signaling & behavior. **9**(11): p.e976155. https://doi.org/10.4161/15592324.2014.976155.

Nordberg, H., Cantor, M., Dusheyko, S., Hua, S., Poliakov, A., Shabalov, I., Smirnova, T., Grigoriev, IV., and Dubchak, I. **2014**. The genome portal of the Department of Energy Joint Genome Institute: 2014 updates. Nucleic Acids Res. **42(1)**:D26-31. https://doi.org/10.1093/nar/gkt1069.

Olsen, O., and Sorensen, H. **1980**. Sinalbin and other glucosinolates in seeds of Double low Rape Species and *Brassica napus* cv. Bronowski. J. Agr. Food Chem. **28**:43-48.

Oshlack, A., Robinson, M.D., and Young, M.D. **2010**. From RNA-seq reads to differential expression results. Genome Biol. **11**:220. https://doi.org/10.1186/gb-2010-11-12-220.

Ozsolak, F., and Milos, P. **2011**. RNA sequencing: advances, challenges and opportunities. Nat. Rev. Genet. **12**:87-98. https://doi.org/10.1038/nrg2934.

Pageau, D., Lajeunesse, J., and Lafond, J. **2006**. Impact de l'hernie des crucifères [*Plasmodiophora brassicae*] sur la productivité et la qualité du canola [Impact of cruciferous hernia (*Plasmodiophora brassicae*) on canola productivity and quality]. Can. J. Plant Pathol. **28**: 137-143. https://doi.org/10.1080/07060660609507280.

Pandey, M.K., Agarwal, G., Kale, S.M., Clevenger, J., and Nayak, S.N., et al. **2017**. Development and evaluation of a high-density genotyping 'Axiom-Arachis' array with 58K SNPs for accelerating genetics and breeding in groundnut. Sci. Rep. **7**:40577. https://doi.org/10.1038/srep40577.

Pang, W., Fu, P., Li, X., Zhan, Z., Yu, S., and Piao, Z. **2018**. Identification and mapping of the clubroot resistance gene CRd in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*). Front. Plant Sci. **9**:653. https://doi.org/10.3389/fpls.2018.00653.

Parkin, I.A., Koh, C., Tang, H., Robinson, S.J., Kagale, S., Clarke, W.E., Town, C.D., Nixon, J., Krishnakumar, V., Bidwell, S.L. and Denoeud, F. **2014**. Transcriptome and methylome profiling reveals relics of genome dominance in the mesopolyploid *Brassica oleracea*. Genome Biol. **15**(6):1-18. https://doi.org/10.1186/gb-2014-15-6-r77.

Partridge, M., and Murphy, D.J. **2009**. Roles of a membrane-bound caleosin and putative peroxygenase in biotic and abiotic stress responses in *Arabidopsis*. Plant Physiology and Biochemistry. **47**:796-806. https://doi.org/10.1016/j.plaphy.2009.04.005.

Peng, G., Falk, K.C., Gugel, R.K., Franke, C., Yu, F., James, B., Strelkov, S.E., Hwang, S.F., and McGregor, L. **2014b**. Sources of resistance to *Plasmodiophora brassicae* (clubroot) pathotypes virulent on canola. Can. J. Plant Pathol. **36**:89-99. https://doi.org/10.1080/07060661.2013.863805.

Peng, G., Lahlali, R., Hwang, S.F., Pageau, D., Hynes, R.K., McDonald, M.R., Gossen, B.D., and Strelkov, S.E. **2014a**. Crop rotation, cultivar resistance, and fungicides/biofungicides for managing clubroot (*Plasmodiophora brassicae*) on canola. Can. J. Plant Pathol. **36**(1):99-112. https://doi.org/10.1080/07060661.2013.860398.

Peng, G., McGregor, L., Lahlali, R., Gossen, B.D., Hwang, S.F., Adhikari, K.K., Strelkov, S.E., and McDonald, M.R. **2011a**. Potential biological control of clubroot on canola and crucifer vegetable crops. Plant Pathol. **60**:566-574. https://doi.org/10.1111/j.1365-3059.2010.02400.x.

Peng, G., Pageau, D., Strelkov, S.E., Gossen, B.D., Hwang, S.-F., and Lahlali, R. **2015**. A >2-year crop rotation reduces resting spores of *Plasmodiophora brassicae* in soil and the impact of clubroot on canola. Eur. J. Agron. **70**:78-84. https://doi.org/10.1016/j. eja.2015.07.007.

Peng, L., Zhou, L., Li, Q., Wei, D., Ren, X., Song, H., et al. **2018b**. Identification of quantitative trait loci for clubroot resistance in *Brassica oleracea* with the use of Brassica SNP microarray. Front. Plant Sci. **9**:822. https://doi.org/10.3389/fpls.2018.00822.

Peng, Y., van Wersch, R. and Zhang, Y. **2018a**. Convergent and divergent signaling in PAMP-triggered immunity and effector-triggered immunity. Mol. Plant Microbe In. **31**(4):403-409. https://doi.org/10.1094/MPMI-06-17-0145-CR.

Pertea, M., Pertea, G., Antonescu, C. *et al.* **2015**. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol **33**:290-295. https://doi.org/10.1038/nbt.3122.

Piao, Z., Deng, Y., Choi, S., Park, Y., and Lim, Y. **2004**. SCAR and CAPS mapping of CRb, a gene conferring resistance to *Plasmodiophora brassicae* in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*). Theor. Appl. Genet. **108**(8):1458-1465. https://doi.org/10.1007/s00122-0031577-5.

Prerostova, S., Dobrev, P.I., Konradyoya, V., Knirsch, V., Gaudinova, A., Kramna, B., Kazda, J., Ludwig-Muller, J., and Vankova, R. **2018**. Hormonal responses to *Plasmodiophora brassicae* infection in *Brassica napus* cultivars differing in their pathogen resistance. Int. J. Mol. Sci. **19**(12):4024. https://doi.org/10.3390/ijms19124024.

Przybylski, R., and Mag, T. **2011**. Canola/rapeseed oil. In Vegetable Oils in Food Technology: Composition, Properties and Uses. Issue Blackwell Publishing Ltd., Hoboken. https://doi.org/10.1002/9783527809080.cataz14026.

Przybylski, R., Mag, T., Eskin, N. A. M., and McDonald, B. E. **2005**. Canola Oil. B_a_i_l_e_y_'s_ _Industrial Oil and Fat Products, Volumes 1-6 (6th Edition) - 2.2.4 Composition. **3**:61–121. https://doi.org/https://app.knovel.com/hotlink/pdf/id:kt003Y1332/baileys-industrial-oil/composition.

Qasim, M.U., Zhao, Q., Shahid, M., Samad, R.A., Ahmar, S., Wu, J., Fan, C., Zhou, Y. **2020**. Identification of QTLs containing resistance genes for Sclerotinia Stem Rot in *Brassica napus* using comparative transcriptomic studies. Front. Plant Sci. **11**:776. https://doi.org/10.3389/fpls.2020.00776.

R Core Team. 2020 R: A language and environment for statistical computing. R Foundation for statistical computing, Vienna, Austria. Available online: https://www.R-project.org/ (accessed on 10 November, 2021).

Rahman, H., Kebede, B., Zimmerli, C., and Yang, R.C. **2014**. Genetic study and QTL mapping of seed glucosinolate content in *Brassica rapa* L. Crop Sci. **54**:537-543. https://doi.org/10.2135/cropsci2013.06.0391.

Rahman, H., Peng, G., Yu, F., Falk, K.C., Kulkarni, M., and Selvaraj, G. **2014b**. Genetics and breeding for clubroot resistance in Canadian spring canola (*Brassica napus* L.). Can. J. Plant Pathol. **36**:122-134. https://doi.org/10.1080/07060661.2013.862571.

Rahman, H., Shakir, A., and Jakir Hasan, M. **2011b**. Breeding for clubroot resistant spring canola (*Brassica napus* L.) for the Canadian prairies: Can the European winter canola cv. Mendel be used as a source of resistance? Can. J. Plant Sci. **91**(3):447-458. https://doi.org/10.4141/cjps10073.

Rahman, M.H., Joersbo, M., and Poulsen, M.H. **2001**. Development of yellow-seeded *Brassica napus* of double low quality. Plant Breed. **120**:473-478. https://doi.org/10.1046/j.1439-0523.2001. 00639.x.

Rahman, M.H., Sun, Z., McVetty, P.B.E., and Li, G. **2008**. High throughput genome-specific and gene-specific molecular markers for erucic acid genes in *Brassica napus* (L.) for marker-assisted selection in plant breeding. Theor. Appl. Genet. **117**:895-904. https://doi.org/10.1007/s00122-008-0829-9.

Rai, A.C., Singh, M., and Shah, K. **2013**. Engineering drought tolerant tomato plants over-expressing BcZAT12 gene encoding a C2H2 zinc finger transcription factor. Phytochemistry. **85**:44-50. https://doi.org/10.1016/j.phytochem.2012.09.007.

Raymer, P.L. **2002**. Canola: An emerging oilseed crop. In: J. Janick and A. Whipkey, Trends in new crops and new uses. ASHS Press. Alexandria, VA, pp. 122–126.

Ren, H., Zhao, X., Li, W., Hussain, J., Qi, G. and Liu, S. **2021**. Calcium signaling in plant programmed cell death. Cells. **10**(5):1089. https://doi.org/10.3390/cells10051089.

Robin, A.H.K., Saha, G., Laila, R., Park, J.I., Kim, H.T., and Nou, I.S. **2020**. Expression and role of biosynthetic, transporter, receptor, and responsive genes for auxin signaling during clubroot disease development. Int. J. Mol. Sci. **21**(15):5554. https://doi.org/10.3390/ijms21155554.

Rolfe, S.A., Strelkov, S.E., Links, M.G., Clarke, W.E., Robinson, S.J., Djavaheri, M., Malinowsky, R., Haddadi, P., Kagale, S., Parkin, I.A.P., Taheri, A., and Borhan, M.H. **2016**. The compact genome of the plant pathogen *Plasmodiophora brassicae* is adapted to intacellular interactions with host Brassica spp. BMC Genomics. **17**:272. https://doi.org/10.1186/s12864-016-2597-2.

Rousseau-Gueutin, M., Belser, C., Da Silva, C., Richard, G., Istace, B., Cruaud, C., Falentin, C., Boideau, F., Boutte, J., and Delourme, R. **2020**. Long-read assembly of the *Brassica napus* reference genome Darmor-bzh. GigaScience. **9**(12):137. https://doi.org/10.1093/gigascience/giaa137.

Saito, M., Kubo, N., Matsumoto, S., Suwabe, K., Tsukada, M., and Hirai, M. **2006**. Fine mapping of the clubroot resistance gene, Crr3, in *Brassica rapa*. Theor. Appl. Genet. **114**(1):81-91. https://doi.org/10.1007/s00122-006-0412-1.

Saitou, N., and Nei, M. **1987**. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. **4(4)**:406-425. https://doi.org/10.1093/oxfordjournals.molbev.a040454.

Sakamoto, K., Saito, A., Hayashida, N., Taguchi, G., and Matsumoto, E. **2008**. Mapping of isolate-specific QTLs for clubroot resistance in Chinese cabbage (*Brassica rapa* L. ssp. *pekinensis*). Theor. Appl. Genet. **117**(5):759-767. https://doi.org/10.1007/s00122-008-0817-0.

Sarkar, A., Kisiala, A., Adhikary, D., Basu, U., Emery, R.N., Rahman, H., and Kav, N.N. **2023**. Silicon ameliorates clubroot responses in canola (*Brassica napus*): A "multi-omics"-based investigation into possible mechanisms. Physiologia Plantarum. **175**(2): e13900. https://doi.org/10.1111/ppl.13900.

Schranz, M.E., and Mitchell-Olds, T. **2006**. Independent ancient polyploidy events in the sister families Brassicaceae and Cleomaceae. The Plant Cell. **18**(5):1152-1165. https://doi.org/10.1105/tpc.106.041111.

Schuller, A., Kehr, J., and Ludwig-Müller, J. **2014**. Laser microdissection coupled to transcriptional profiling of Arabidopsis roots inoculated by *Plasmodiophora brassicae* indicates a role for brassinosteroids in clubroot formation. Plant Cell Physiol. **55**:392-411. https://doi.org/10.1093/pcp/pct174.

Schwelm, A., Fogelqvist, J., Knaust, A., Jülke, S., Lilja, T., Bonilla-Rosso, G., et al. **2015**. The *Plasmodiophora brassicae* genome reveals insights in its life cycle and ancestry of chitin synthases. Sci. Rep. **5**:11153. https://doi.org/10.1038/srep11153.

Sedaghatkish, A., Gossen, B.D., Yu, F., Torkamaneh, D., and McDonald, M.R. **2019**. Whole-genome DNA similarity and population structure of *Plasmodiophora brassicae* strains from Canada. BMC Genomics. **20**:744. https://doi.org/10.1186/s12864-019-6118-y.

Selote, D., Robin, G.P., and Kachroo, A. **2013**. GmRIN4 protein family members function nonreduntantly in soybean race-specific resistance against *Pseudomonas syringae*. New Phytol. **197**:1225-1235. https://doi.org/10.1111/nph.12093.

Seo, E., and Choi, D. **2015**. Functional studies of transcription factors involved in plant defenses in the genomics era. Brief Funct. Genom. **14**:260-267. https://doi.org/10.1093/bfgp/elv011.

Shah, N., Sun, J., Yu, S., Yang, Z., Wang, Z., Huang, F., Dun, B., Gong, J., Liu, Y., Li, Y., Li, Q., Yuan, L., Baloch, A., Li, G., Li, S., and Zhang, C. **2019**. Genetic variation analysis of field isolates of clubroot and their responses to *Brassica napus* lines containing resistant genes CRb and PbBa8.1 and their combination in homozygous and heterozygous state. Mol. Breed. **39**(10-11):153. https://doi.org/10.1007/s11032-019-1075-3.

Shahidi, F. **1990**. Canola and rapeseed: Production, chemistry, nutrition and processing technology, Springer Science and Bussiness Media. Van Nostrand Reinhold, New York, USA.

Sharma, K., Gossen, B. D., and McDonald, M. R. **2011b**. Effect of temperature on cortical infection by *Plasmodiophora brassicae* and clubroot severity. Phytopathol. **101**:1424-1432. https://doi.org/10.1094/PHYTO-04-11-0124.

Sharma, K., Gossen, B. D., McDonald, M. R. **2011a**. Effect of temperature on primary infection by *Plasmodiophora brassicae* and initiation of clubroot symptoms. Plant Pathol. **60**:830-838. https://doi.org/10.1111/j.1365-3059.2011.02458.x.

Sharma, R., Mahanty, B., Mishra, R., and Joshi, R.K. **2021**. Genome wide identification and expression analysis of pepper C2H2 zinc finger transcription factors in response to anthracnose pathogen *Colletotrichum truncatum*. Biotech. **11**(3):1-18. https://doi.org/10.1007/s13205-020-02601-x.

Shi, H., Wang, X., Ye, T., Chen, F., Deng, J., Yang, P. et al. **2014**. The Cysteine2/Histidine2-Type transcription factor ZINC FINGER OF *ARABIDOPSIS THALIANA* 6 modulates biotic and abiotic stress responses by activating salicylic acid-related genes and *C-REPEAT-BINDING*

FACTOR genes in *Arabidopsis*. Plant Physiol. **165**:1367-79. https://doi.org/10.1104/pp.114.242404.

Shokri-Gharelo, R., and Noparvar, P.M. **2018**. Molecular response of canola to salt stress: insights on tolerance mechanisms. PeerJ. **6**: e4822. https://doi.org/10.7717/peerj.4822.

Some, A., Manzanares, M. J., Laurens, F., Baron, F., Thomas, G., and Rouxel, F. **1996**. Variation for virulence on *Brassica napus* L. amongst *Plasmodiophora brassicae* collections from France and derived single-spore isolates. Plant Pathol. **45**:432-439. https://doi.org/10.1046/j.1365-3059.1996.d01-155.x.

Song, J.M., Guan, Z., Hu, J., Guo, C., Yang, Z., Wang, S., Liu, D., Wang, B., Lu, S., and Zhou, R. **2020**. Eight high-quality genomes reveal pan-genome architecture and ecotype differentiation of *Brassica napus*. Nature Plants. **6**(1):34-45. https://doi.org/10.1038/s41477-019-0577-7.

Song, T., Chu, M., Lahlali, R., Yu, F., and Peng, G. **2016**. Shotgun label-free proteomic analysis of clubroot (*Plasmodiophora brassicae*) resistance conferred by the gene *Rcr1* in *Brassica rapa*. Front. Plant Sci. **7**:1013. https://doi.org/10.3389/fpls.2016.01013.

Stael, S., Kmiecik, P., Willems, P., Van Der Kelen, K., Coll, N.S., Teige, M. and Van Breusegem, F. **2015**. Plant innate immunity–sunny side up? Trends in plant science. **20**(1):3-11. https://doi.org/10.1016/j.tplants.2014.10.002.

Stamm, S., Ben-Ari, S., Rafalska, I., Tang, Y., Zhang, Z., Toiber, D. et al. **2005**. Function of alternative splicing. Gene. **344**:1-20. https://doi.org/10.1016/j.gene.2004.10.022.

Statista. *Worldwide oilseed production by type 2021/22*. https://www.statista.com/statistics/267271/worldwide-oilseed-production-since-2008/ (accessed March 27, 2023).

Stegmann, M., Monaghan, J., Smakowska-Luzan, E., Rovenich, H., Lehner, A., Holton, N., and Zipfel, C. **2017**. The receptor kinase FER is a RALF-regulated scaffold controlling plant immune signaling. Science. **355**(6322):287-289. https://doi.org/10.1126/science.aal2541.

Strelkov, S.E. and Dixon, G.R. **2014**. Clubroot (*Plasmodiophora brassicae*) on canola and other Brassica species – disease development, epidemiology and management. Canadian Journal of Plant Pathology. **36**(1):1-4. https://doi.org/10.1080/07060661.2013.875338.

Strelkov, S.E., and Hwang, S. F. **2014**. Clubroot in the Canadian canola crop: 10 years into the outbreak. Can. J. Plant Pathol. **36**:27-36. https://doi.org/10.1080/07060661.2013.863807.

Strelkov, S.E., Hwang, S., Manolii, V.P., Cao, T., FreduaAgyeman, R, Harding, M.W., Peng, G., Gossen, B.D., Mcdonald, M.R., and Deindel, D. **2018**. Virulence and pathotype classification of *Plasmodiophora brassicae* populations collected from clubroot resistant canola (*Brassica napus*) in Canada. Can J Plant Pathol. **40**:284-298. https://doi.org/10.1080/07060661.2018.1459851.

Strelkov, S.E., Hwang, S.F., Manolii, V.P., Cao, T., and Feindel, D. **2016b**. Emergence of new virulence phenotypes of *Plasmodiophora brassicae* on canola (*Brassica napus*) in Alberta, Canada. Eur J Plant Pathol. **145**:517-529. https://doi.org/10.1007/s10658-016-0888-8.

Strelkov, S.E., Hwang, S.F., Manolii, V.P., Turnbull, G., Fredua-Agyeman, R., Hollman, K., and Kaus, S. **2021**. Characterization of clubroot (*Plasmodiophora brassicae*) from canola (*Brassica napus*) in the Peace Country of Alberta, Canada. Can. J. Plant Pathol. **43**:155-161. https://doi.org/10.1080/07060661.2020.1776931.

Strelkov, S.E., Manolii, V.P., Cao, T., Xue, S., and Hwang, S.F. **2007**. Pathotype classification of *Plasmodiophora brassicae* and its occurrence in *Brassica napus* in Alberta, Canada. Phytopathol. **155**:706-712. https://doi.org/10.1111/j.14390434.2007.01303.x.

Strelkov, S.E., Tewari, J.P., and Smith-Degenhardt, E. **2006**. Characterization of *Plasmodiophora brassicae* populations from Alberta, Canada. Can. J. Plant Pathol. **28**:467-474. https://doi.org/10.1080/07060660609507321.

Su, T., Yu, S., Wang, W., Li, P., Zhang, F., Yu, Y., Zhang, D., and Zhao, X. **2018**. iTRAQ analysis of protein profile during the secondary stage of infection of *Plasmodiophora brassicae* in Chinese cabbage (*Brassica rapa*). J. Plant Pathol. **100**:533-542. https://doi.org/10.1007/s42161-018-0121-z.

Sultan, M., Schulz, M. H., Richard, H., Magen, A., Klingenhoff, A., and Yaspo, M. **2008**. A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. Science. **321**:956-960. https://doi.org/10.1126/science.1160342.

Summanwar, A., Basu, U., Kav, N.N.V., and Rahman, H. **2020**. Identification of lncRNAs in response to infection by *Plasmodiophora brassicae* in *Brassica napus* and development of lncRNA-based SSR markers. Genome. **64**(5):547-566. https://doi.org/10.1139/gen-2020-0062.

Summanwar, A., Basu, U., Rahman, H., and Kav N. **2019**. Identification of lncRNAs responsive to infection by *Plasmodiophora brassicae* in clubroot-susceptible and -resistant *Brassica napus* lines carrying resistance introgressed from rutabaga. Mol. Plant Microbe Interact. **32**(10):1360-1377. https://doi.org/10.1094/MPMI-12-18-0341-R.

Summanwar, A., Farid, M., Basu, U., Kav, N., and Rahman H. **2021**. Comparative transcriptome analysis of canola carrying clubroot resistance from 'Mendel' or rutabaga and the development of molecular markers. Physiol Mol Plant P. **114**:101640. https://doi.org/10.1016/j.pmpp.2021.101640.

Sun, S.J., Guo, S.Q., Yang, X., Bao, Y.M., Tang, H.J., Sun, H., ... and Zhang, H.S. **2010**. Functional analysis of a novel Cys2/His2-type zinc finger protein involved in salt tolerance in rice. J. Exp. Bot. **61**(10):2807-2818. https://doi.org/10.1093/jxb/erq120.

Suwabe, K., Tsukazaki, H., Iketani, H., Hatakeyama, K., Fujimura, M., Nunome, T., Fukuoka, H., Matsumoto, S., and Hirai, M. **2003**. Identification of two loci for resistance to clubroot (*Plasmodiophora brassicae* Woronin) in *Brassica rapa* L. Theor. Appl. Genet. **107**(6):997-1002. https://doi.org/10.1007/s00122-003-1309-x.

Suwabe, K., Tsukazaki, H., Iketani, H., Hatakeyama, K., Kondo, M., Fujimura, M., Nunome, T., Fukuoka, H., Hirai, M., and Matsumoto, S. **2006**. Simple sequence repeat-based comparative genomics between *Brassica rapa* and *Arabidopsis thaliana*: the genetic origin of clubroot resistance. Genetics. **173**(1):309-19. https://doi.org/10.1534/genetics.104.038968.

Suzuki, K., Sugimoto, K., Hayashi, H., and Komyoji, T. **1995**. Biological mode of action of fluazinam, a new fungicide, for Chinese cabbage clubroot. Ann. Phytopathol. Soc. Jpn. **61**: 395-398 (in Japanese). https://doi.org/10.3186/jjphytopath.61.395.

Swiderski, M.R., Birker, D. and Jones, J.D., **2009**. The TIR domain of TIR-NB-LRR resistance proteins is a signaling domain involved in cell death induction. Molecular plant-microbe interactions. **22**(2):157-165. https://doi.org/10.1094/MPMI-22-2-0157.

Takatsuji H. **1999**. Zinc-finger proteins: the classical zinc finger emerges in contemporary plant science. Plant Mol. Biol. **39**:1073-1078. https://doi.org/10.1023/A:1006184519697.

Takatsuji, H. **1998**. Zinc-finger transcription factors in plants. Cell Mol. Life Sci. **54**:582–96. https://doi.org/10.1007/s000180050186.

Takatsuji, H., Mori, M., Benfey, P.N., Ren, L., and Chua, N.H. **1992**. Characterization of a zinc finger DNA-binding protein expressed specifically in *Petunia* petals and seedlings. EMBO J. **11**:241-249. https://doi.org/10.1002/j.1460-2075.1992.tb05047.x.

Tao, Z., Huang, Y., Zhang, L., Wang, X., Liu, G., and Wang, H. **2017**. BnLATE, a Cys2/His2-type zinc-finger protein, enhances silique shattering resistance by negatively regulating lignin accumulation in the silique walls of *Brassica napus*. PloS One. **12**: e0168046. https://doi.org/10.1371/journal.pone.0168046.

Tewari, J.P., Strelkov, S.E., Orchard, D., Hartman, M., Lange, R.M., and Turkington, T.K. **2005**. Identification of clubroot of crucifers on canola (*Brassica napus*) in Alberta. Can. J. Plant Pathol. **27**:143-144. https://doi.org/10.1080/07060660509507206.

Tian, Z.D., Zhang, Y., Liu, J., and Xie, C.H. **2010**. Novel potato C2H2-type zinc finger protein gene, *StZFP1*, which responds to biotic and abiotic stress, plays a role in salt tolerance. Plant Biol. **12**(5):689-697. https://doi.org/10.1111/j.1438-8677.2009.00276.x.

Tilman, D., Cassman, K.G., Matson, P.A., Naylor. R., and Polasky, S. **2002**. Agricultural sustainability and intensive production practices. Nature. **418**:671-677. https://doi.org/10.1038/nature01014.

Timothy, L. Bailey, Johnson, J., Grant, C.E., and Noble, W.S. **2015**. The MEME Suite. Nucleic Acid Res. **43**(W1): W39-W49. https://meme-suite.org/meme/tools/meme.

Ting, J.P.Y., Willingham, S.B., and Bergstralh, D.T. **2008**.NLRs at the intersection of cell death and immunity. Nat.Rev. Immunol. **8**(5):372–379. https://doi.org/10.1038/nri2296.

Tomita, H., Shimizu, M., and Asad-ud Doullah, M. **2013**. Accumulation of quantitative trait loci conferring broad-spectrum clubroot resistance in *Brassica oleracea*. Mol. Breeding**. 32**:889-900. https://doi.org/10.1007/s11032-013-9918-9.

Torres, M. A., Jones, J. D., and Dangl, J. L. **2006**. Reactive oxygen species signaling in response to pathogens. Plant physiol. **141**:373-378. https://doi.org/10.1104/pp.106.079467.

Trapnell, C., Roberts, A., and Goff, L. **2012**. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc. **7**:562-578. https://doi.org/10.1038/nprot.2012.016.

Tremblay, N., Belec, C., Coulombe, J., and Godin, C. **2005**. Evaluation of calcium cyanamide and liming for control of clubroot disease in cauliflower. Crop Prot. **24**:798-803. https://doi.org/10.1016/j.cropro.2004.12.013.

Trujillo, M., and Shirasu K. **2010**. Ubiquitination in plant immunity. Curr. Opin Plant Biol. **13**: 402-408. https://doi.org/10.1016/j.pbi.2010.04.002.

Tundo, S., Mandala, G., Sella, L., Favaron, F., Bedre, R., and Kalunke, R.M. **2022**. Xylanase inhibitors: Defense players in plant immunity with implications in agro-industrial processing. Int. J. Mol. Sci. **23**(23):14994. https://doi.org/10.3390/ijms232314994.

U., Nagaharu. **1935**. Genome analysis in *Brassica* with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. Jpn. J. Bot. **7**:389-452.

Uehara, Y., Takahashi, Y., Berberich, T. et al. **2005**. Tobacco ZFT1, a Transcriptional Repressor with a Cys_2/His_2 Type Zinc Finger Motif that Functions in Spermine-Signaling Pathway. Plant Mol. Biol. **59**:435-448. https://doi.org/10.1007/s11103-005-0272-0.

Ueno, H., Matsumoto, E., Aruga, D., Kitagawa, S., Matsumura, H., and Hayashida, N. **2012**. Molecular characterization of the CRa gene conferring clubroot resistance in *Brassica rapa*. Plant Mol. Biol. **80**(6):621-629. doi:10.1007/s11103-012-9971-5.

Uhrig, R.G., Schläpfer, P., Roschitzki, B., Hirsch-Hoffmann, M., and Gruissem, W. **2019**. Diurnal changes in concerted plant protein phosphorylation and acetylation in Arabidopsis organs and seedlings. Plant J. **99**(1):176-194. https://doi.org/10.1111/tpj.14315.

Ülker, B., Shahid Mukhtar, M. and Somssich, I.E. **2007**. The WRKY70 transcription factor of *Arabidopsis* influences both the plant senescence and defense signaling pathways. Planta. **226**:125-137. https://doi.org/10.1007/s00425-006-0474-y.

USDA Foreign Agricultural Service **2022**. *Oilseeds: World Markets and Trade*. https://www.fas.usda.gov/data/oilseeds-world-markets-and-trade (accessed 2023-04-03).

Varshney, R. K., Nayak, S. N., May, G. D., and Jackson, S. A. **2009**. Next-generation sequencing technologies and their implications for crop genetics and breeding. Trends Biotech. **27**(9):522-530. https://doi.org/10.1016/j.tibtech.2009.05.006.

Verma, V., Ravindran, P., and Kumar, P.P. **2016**. Plant hormone-mediated regulation of stress responses. BMC Plant Biol. **16**:86. https://doi.org/10.1186/s12870-016-0771-y.

Vidova, V., and Spacil, Z. **2017**. A review on mass spectrometry-based quantitative proteomics: Targeted and data independent acquisition. Analytica chimica acta. **964**: 7-23. https://doi.org/10.1016/j.aca.2017.01.059.

Voorrips, R. **2002**. MapChart: software for the graphical presentation of linkage maps and QTLs. J. hered. **93**(1):77-78. https://doi.org/10.1093/jhered/93.1.77.

Voorrips, R. E., and Kanne, H. J. **1997**. Genetic analysis of resistance to clubroot (*Plasmodiophora brassicae*) in *Brassica oleracea*. I. Analysis of symptom grades. Euphytica. **93**:31-39. https://doi.org/10.1023/A:1002970103683.

Voorrips, R.E., and Visser, D.L. **1993**. Examination of resistance to clubroot in accessions of *Brassica oleracea* using a glasshouse seedling test. Neth. J. Plant Pathol. **99**:269-276. https://doi.org/10.1007/BF01974308.

Vorwerk, S., Schiff, C., Santamaria, M. et al. **2007**. *EDR2* negatively regulates salicylic acid-based defenses and cell death during powdery mildew infections of *Arabidopsis thaliana*. BMC Plant Biol. **7**:35. https://doi.org/10.1186/1471-2229-7-35.

Wagner, G., Charton, S., Lariagon, C., Laperche, A., Lugan, R., Hopkins, J., Frendo, P., Bouchereau, A., Delourme, R., Gravot, A., and Manzanares-Dauleux, M.J. **2012**. Metabotyping: a new approach to investigate rapeseed (*Brassica napus* L.) genetic diversity in the metabolic response to clubroot infection. MPMI. **25**:1478-1491. https://doi.org/10.1094/MPMI-02-12-0032-R.

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 Pathol. **45**:710-719. https://doi.org/10.1046/j.1365-3059. 1996.d01-173.x.

Wallenhammar, A.C., Almquist, C., Schwelm, A., Roos, J., Marzec-Schmidt, K., Jonsson, A., and *Dixelius*, C. **2014**. Clubroot, a persistent threat to Swedish oilseed rape production. Can. J. Plant Pathol. **36**:135-141. https://doi.org/10.1080/07060661.2013.870606.

Wang, F., Tong, W., Zhu, H. et al. **2016**. A novel Cys_2/His_2 zinc finger protein gene from sweetpotato, *IbZFP1*, is involved in salt and drought tolerance in transgenic *Arabidopsis*. Planta, **243**:783-797. https://doi.org/10.1007/s00425-015-2443-9.

Wang, S., Yu, F., Zhang, W., Tang, J., Li, J., Yu, L., Wang, H., and Jiang, J. **2019**. Comparative transcriptomic analysis reveals gene expression changes during early stages

of *Plasmodiophora brassicae* infection in cabbage (*Brassica oleracea* var. *capitata* L.). Can J Plant Pathol. **2**:188–199. https://doi.org/10.1080/07060661.2019.1567592.

Wang, X., Wang, H., Wang, J., Sun, R., Wu, J., Liu, S., Bai, Y., Mun, J.H., Bancroft, I., Cheng, F., Huang, S., Li, X., Hua, W., Wang, J., Wang, X., Freeling, M., Pires, J.C., Paterson, A.H., Chalhoub, B., … Zhang, Z. **2011**. The genome of the mesopolyploid crop species *Brassica rapa*. Nat Genet. **43**:1035-1040. https://doi.org/10.1038/ng.919.

Wang, Z., Fang, H., Chen, Y., Chen, K., Li, G., Gu, S., and Tan, X. **2014**. Overexpression of BnWRKY33 in oilseed rape enhances resistance to *Sclerotinia sclerotiorum*. Mol. Plant Pathol. **15**(7):677-689. https://doi.org/10.1111/mpp.12123.

Wang, Z., Gerstein, M., and Snyder, M. **2009**. RNA-Seq: a revolutionary tool for transcriptomics. Nature reviews genetics. **10**(1):57-63. https://doi.org/10.1038/nrg2484.

Wang, Z., Megha, S., Kebede, B., Kav, N.N., and Rahman, H. **2022**. Genetic and molecular analysis reveals that two major loci and their interaction confer clubroot resistance in canola introgressed from rutabaga. The Plant Genome. e20241. https://doi.org/10.1002/tpg2.20241.

Warwick, S., Francis, A., and Al-Shehbaz, I. **2006**. Brassicaceae: Species checklist and 1692 database on CD-Rom. Plant Syst. Evol. **259**(2-4):249-258. https://doi.org/10.1007/s00606-006-0422-0.

Watson, A.G., and Baker, K.F. **1969**. Possible gene centres for resistance in the genus *Brassica* to *Plasmodiophora brassicae*. Econ. Bot. **23**:245-252. https://doi.org/10.1007/BF02860456.

Webster, M.A., and Dixon, G.R. **1991**. Boron, pH and inoculum concentration infuencing colonization by *Plasmodiophora brassicae*. Myc. Res. **95**:64-73. https://doi.org/10.1016/S0953-7562(09)81363-4.

Wei, X., Li, J., Zhang, X., Zhao, Y., Nath, U.K., Mao, L., Xie, Z., Yang, S., Shi, G., Wang, Z. and Tian, B. **2022**. Fine mapping and functional analysis of major QTL, *CRq* for clubroot resistance in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*). Agronomy. **12**(5): p.1172. https://doi.org/10.3390/agronomy12051172.

Wei, X., Zhang, Y., Zhao, Y., Xie, Z., Hossain, M.R., Yang, S., Shi, G., Lv, Y., Wang, Z., Tian, B., Su, H., Wei, F., Zhang, X., and Yuan, Y. **2021**. Root transcriptome and metabolome profiling reveal key phytohormone-related genes and pathways involved clubroot resistance in *Brassica rapa* L. Front. Plant Sci. **12**:759623. https://doi.org/10.3389/fpls.2021.759623.

Werner, S., Diederichsen, E., Frauen, M., Schondelmaier, J., and Jung, C. **2008**. Genetic mapping of clubroot resistance genes in oilseed rape. Theor. Appl. Genet. **116**(3):363-372. https://doi.org/10.1007/s00122-007-0674-2.

Westermann, A.J., Gorski, S.A., and Vogel, J. **2012**. Dual RNA-seq of pathogen and host. Nat. Rev. Microbiol. **10**(9):618-630. https://doi.org/10.1038/nrmicro2852.

Wilkins, M.R., Sanchez, J.C., Gooley, A.A., Appel, R.D., Humphery-Smith, I., Hochstrasser, D. F., and Williams, K.L. **1996**. Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. Biotechnol. Genet. Eng. **13**(1):19-50. https://doi.org/10.1080/02648725.1996.10647923.

Williams, P.H. **1966**. A system for the determination of races of *Plasmodiophora brassicae* that infect cabbage and rutabaga. Phytopathol. **56**:624-626.

Wilton, M., Subramaniam, R., Elmore, J., Felsensteiner, C., Coaker, G., and Desveaux, D. **2010**. The type III effector HopF2(Pto) targets *Arabidopsis* RIN4 protein to promote *Pseudomonas syringae* virulence. Proc. Natl. Acad. Sci. USA. **107**:2349-2354. https://doi.org/10.1073/pnas.0904739107.

Wu, W., and Ma, B.L. **2021**. The stage sensitivity of short–term heat stress to lodging–resistant traits and yield determination in canola (*Brassica napus L.*). J. Agron. Crop Sci. **207**:74-87. https://doi.org/10.1111/jac.12464.

Xu, D.Q., Huang, J., Guo, S.Q., Yang, X., Bao, Y.M., Tang, H.J., and Zhang, H.S. **2008**. Overexpression of a TFIIIA-type zinc finger protein gene ZFP252 enhances drought and salt tolerance in rice (*Oryza sativa* L.). FEBS letters. **582**(7):1037-1043. https://doi.org/10.1016/j.febslet.2008.02.052.

Xu, L., Yang, H., Ren, L., Chen, W., Liu, L., Liu, F., et al. **2018**. Jasmonic acid-mediated aliphatic glucosinolate metabolism is involved in clubroot disease development in *Brassica napus* L. Front. Plant Sci. **9**: 750. https://doi.org/10.3389/fpls.2018.00750.

Xu, X., Chen, C., Fan, B. and Chen, Z. **2006**. Physical and functional interactions between pathogen-induced Arabidopsis WRKY18, WRKY40, and WRKY60 transcription factors. The Plant Cell. **18**(5):1310-1326. https://doi.org/10.1105/tpc.105.037523.

Xu, Z., Raza, Q., Xu, L., He, X., Huang, Y., Yi, J., ... and Ali, Z. **2018**. GmWRKY49, a salt-responsive nuclear protein, improved root length and governed better salinity tolerance in transgenic *Arabidopsis*. Front. Plant Sci. **9**:809. https://doi.org/10.3389/fpls.2018.00809.

Xue, S., Cao, T., Howard, R.J., Hwang, S.F., and Strelkov, S.E. **2008**. Isolation and variation in virulence of single-spore isolates of *Plasmodiophora brassicae* from Canada. Plant Dis. **92**: 456-462. https://doi.org/10.1094/PDIS-92-3-0456.

Yamauchi, Y., Hasegawa, A., Taninaka, A., Mizutani, M., and Sugimoto, Y. **2011**. NADPH-dependent reductases involved in the detoxification of reactive carbonyls in plants. J Biol Chem. **286**(9):6999-7009. https://doi.org/10.1074/jbc.M110.202226.

Yang, B., Jiang, Y., Rahman, M.H., Deyholos, M.K., and Kav, N.N. **2009**. Identification and expression analysis of WRKY transcription factor genes in canola (*Brassica napus* L.) in response to fungal pathogens and hormone treatments. BMC Plant Biol. **9**(1):1-19. https://doi.org/10.1186/1471-2229-9-68.

Yang, F., Zhang, Y., Huang, Q. et al. **2015**. Analysis of key genes of jasmonic acid mediated signal pathway for defense against insect damages by comparative transcriptome sequencing. Sci Rep **5**:16500. https://doi.org/10.1038/srep16500.

Yang, J., Liu, D., Wang, X., Ji, C., Cheng, F., Liu, B., Hu, Z., Chen, S., Pental, D., Ju, Y., Yao, P., Li, X., Xie, K., Zhang, J., Wang, J., Liu, F., Ma, W., Shopan, J., Zheng, H., … Zhang, M. **2016a**. The genome sequence of allopolyploid *Brassica juncea* and analysis of differential homoeolog gene expression influencing selection. Nat. Genet. **48**:1225-1232. https://doi.org/10.1038/ng.3657.

Yang, J., Liu, G., Zhao, N., Chen, S., Liu, D., Ma, W., ... and Zhang, M. **2016**. Comparative mitochondrial genome analysis reveals the evolutionary rearrangement mechanism in Brassica. Plant biology. **18**(3): 527-536. https://doi.org/10.1111/plb.12414.

Yatusevich, R., Mugford, S.G., Matthewman, C., Gigolashvili, T., Frerigmann, H., Delaney, S., Koprivova, A., Flügge, U.I. and Kopriva, S. **2010**. Genes of primary sulfate assimilation are part of the glucosinolate biosynthetic network in *Arabidopsis thaliana*. The Plant Journal. **62**(1):1-11. https://doi.org/10.1111/j.1365-313X.2009.04118.x.

Yin M., Wang Y., Zhang L., Li J., Quan W., Yang L., Wang Q., and Chan Z. **2017**. The Arabidopsis Cys2/His2 zinc finger transcription factor ZAT18 is a positive regulator of plant tolerance to drought stress. *J. Exp. Bot.* **68**:2991-3005. https://doi.org/10.1093/jxb/erx157.

Yin, J., Wang, L., Zhao, J., Li, Y., Huang, R., Jiang, X., ... and Zhu, Y. **2020**. Genome-wide characterization of the C2H2 zinc-finger genes in Cucumis sativus and functional analyses of four *CsZFPs* in response to stresses. BMC Plant Biol. **20**(1):1-22. https://doi.org/10.1186/s12870-020-02575-1.

Yoshikawa, H. **1993**. Studies on breeding of clubroot resistance in cole [Cruciferae] crops. Bull. Natl. Res. Inst. Veg. Ornam. Plants Tea. Ser. A **7**:1–165.

You, F.M., Huo, N., Gu, Y., Luo, M.C., Ma, Y., Hane, D., Lazo, G.R., Dvorak, J., Anderson, O.D. **2008**. BatchPrimer3: A high throughput web application for PCR and sequencing primer design. BMC Bioinformatics. **9**:253. https://doi.org/10.1186/ 1471- 2105- 9- 253.

Yu Z., Yan H., Liang L., Zhang Y., Yang H., Li W., Choi J., Huang J., and Deng S. **2021**. A C_2H_2-Type Zinc-Finger Protein from *Millettia pinnata*, MpZFP1, enhances Salt Tolerance in Transgenic *Arabidopsis*. Int. J. Mol. Sci. **22**:832. https://doi.org/10.3390/ijms221910832.

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., and Peng, G. **2016**. Identification of Genome-wide variants and discovery of variants associated with *Brassica rapa* clubroot resistance gene Rcr1 through bulked segregant RNA sequencing. PLoS One. **11**(4): e0153218. https://doi.org/10.1371/journal.pone.0153218.

Yu, F., Zhang, X., Peng, G., Falk, K.C., Strelkov, S.E., and Gossen, B.D. **2017**. Genotyping-by-sequencing reveals three QTL for clubroot resistance to six pathotypes of *Plasmodiophora brassicae* in *Brassica rapa*. Sci. Rep. **7**(1):1–11. Springer US. https://doi.org/10.1038/s41598-017-04903-2.

Yu, F., Zhang, Y., Wang, J., Chen, Q., Karim, M.M., Gossen, B.D., and Peng, G. **2022**. Identification of two major QTLs in Brassica napus lines with introgressed clubroot

resistance from turnip cultivar ECD01. Front. Plant Sci. **12**:785989. https://doi.org/10.3389/fpls.2021.785989.

Yu, G., Wang, L., Han, Y., and He, Q. **2012**. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS: a journal of integrative biology. **16**(5):284-287. https://doi.org/10.1089/omi.2011.0118.

Yu, Y.H., Li, X.Z., Wu, Z.J., Chen, D.X., Li, G.R., Li, X.Q., and Zhang, G.H. **2016**. VvZFP11, a Cys2His2-type zinc finger transcription factor, is involved in defense responses in Vitis vinifera. Biol. Plantarum. **60**:292-298. https://doi.org/10.1007/s10535-016-0598-2.

Yu, Z., Yan, H., Liang, L., Zhang, Y., Yang, H., Li, W., ... and Deng, S. **2021**. A C2H2-type zinc-finger protein from *Millettia pinnata*, MpZFP1, enhances salt tolerance in transgenic Arabidopsis. Int. J. Mol. Sci. **22**(19):10832. https://doi.org/10.3390/ijms221910832.

Yuan S., Li X., Li R. et al. **2018**. Genome-wide identification and classification of soybean C_2H_2 zinc finger proteins and their expression analysis in legume-*Rhizobium Symbiosis*. Front. Microbiol. **9**:126. https://doi.org/10.3389/fmicb.2018.00126.

Yuan, M., Jiang, Z., Bi, G., Nomura, K., and Xin, X.F. **2021a**. Pattern-recognition receptors are required for NLR-mediated plant immunity. Nature. **592**:105-109. https://doi.org/10.1038/s41586-021-03316-6.

Yuan, P., Jauregui, E., Du, L., Tanaka, K., Poovaiah, B.W. **2017**. Calcium signatures and signaling events orchestrate plant-microbe interactions. Curr. Opin. Plant Biol. **38**:173-183. https://doi.org/10.1016/j.pbi.2017.06.003.

Yuan, Y., Qin, L., Su, H., Yang, S., Wei, X., Wang, Z., Zhao, Y., Li, L., Liu, H., Tian, B., Zhang, X. **2021b**. Transcriptome and coexpression network analyses reveal hub genes in Chinese

cabbage (*Brassica rapa* L. ssp. *pekinensis*) during different stages of *Plasmodiophora brassicae* infection. Front. Plant Sci. **12**: 650252. https://doi.org/10.3389/fpls.2021.650252.

Yuan, Y., Wei, X., Zhang, Q., Zhao, Y., Jiang, W., and Yao, Q. **2015**. BSA-Seq technologies identify a major QTL for clubroot resistance in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*). In Proceedings of the KSM Spring Meeting & KSM-ICWG-GSP Joint Clubroot Symposium, Daejeon Convention Center, Daejeon, Korea, 13–15 May 2015; p. 41.

Zeilmaker, T., Ludwig, N.R., Elberse, J., Seidl, M.F., Berke, L., Van Doorn, A., Schuurink, R.C., Snel, B., and Van den Ackerveken, G. **2015**. DOWNY MILDEW RESISTANT 6 and DMR6-LIKE OXYGENASE 1 are partially redundant but distinct suppressors of immunity in Arabidopsis. Plant J. **81**(2):210-22. https://doi.org/10.1111/tpj.12719.

Zhan, Z., Jiang, Y., Shah, N., Hou, Z., Zhou, Y., Dun, B., Li, S., Zhu, L., Li, Z., Piao, Z., et al. **2020**. Association of clubroot resistance locus *PbBa8.1* with a linkage drag of high erucic acid content in the seed of the European turnip. Front. Plant Sci. **11**:810. https://doi.org/10.3389/fpls.2020.00810.

Zhang D., Tong J., Xu Z., Wei P., Xu L., Wan Q. et al. **2016**. Soybean C2H2-type zinc finger protein GmZFP3 with conserved QALGGH motif negatively regulates drought responses in transgenic *Arabidopsis*. Front. Plant Sci. **7**:325. https://doi.org/10.3389/fpls.2016.00325.

Zhang X., and Fernando W.G.D. **2018**. Insights into fighting against blackleg disease of *Brassica napus* in Canada. Crop and Pasture Sci. **69**:40-47. https://doi.org/10.1071/CP16401.

Zhang, H., Feng, J., Hwang, S.F., Strelkov, S.E., Falak, I., Huang, X., and Sun, R. **2016a**. Mapping of clubroot (*Plasmodiophora brassicae*) resistance in canola (*Brassica napus*). Plant Pathol. **65**(3):435-440. doi:10.1111/ppa.12422.

Zhang, H., Ma, X., Liu, X., Zhang, S., Li, F., Li, G., Sun, R. and Zhang, S. **2022**. Identification and fine mapping of clubroot (*Plasmodiophora brassicae*) resistant QTL in *Brassica rapa*. Horticulturae. **8**(1):66. https://doi.org/10.3390/horticulturae8010066.

Zhang, L., Du, L., and Poovaiah, B. W. **2014**. Calcium signaling and biotic defense responses in plants. Plant Signal. Behav. **9**: e973818. https://doi.org/10.4161/15592324.2014.973818.

Zhang, X., Liu, Y., Fang, Z., Li, Z., Yang, L., Zhuang, M., Zhang, Y. and Lv, H. **2016**. Comparative transcriptome analysis between broccoli (*Brassica oleracea* var. *italica*) and wild cabbage (*Brassica macrocarpa* Guss.) in response to *Plasmodiophora brassicae* during different infection stages. Front. Plant Sci. **7**:1929. https://doi.org/10.3389/fpls.2016.01929.

Zhang, Y., Tessaro, M.J., Lassner, M., and Li, X. **2003**. Knockout analysis of Arabidopsis transcription factors TGA2, TGA5, and TGA6 reveals their redundant and essential roles in systemic acquired resistance. Plant Cell. **15**(11):2647-53. https://doi.org/10.1105/tpc.014894.

Zhao B.Y., Hu Y.F., Li J.J. et al. **2016**. BnaABF2, a bZIP transcription factor from rapeseed (*Brassica napus* L.), enhances drought and salt tolerance in transgenic *Arabidopsis*. Bot. Stud. **57**:12. https://doi.org/10.1186/s40529-016-0127-9.

Zhao, C., Zayed, O., Yu, Z., and Zhu, J.K. **2018**. Leucine-rich repeat extension proteins regulate plant salt tolerance in Arabidopsis. Biological Sci. **115**(51):13123-13128. https://doi.org/10.1073/pnas.1816991115.

Zhao, J., Wang, J., An, L. et al. **2007**. Analysis of gene expression profiles in response to *Sclerotinia sclerotiorum* in *Brassica napus*. Planta. **227**:13-24. https://doi.org/10.1007/s00425-007-0586-z.

Zhao, T., Wu, T., Zhang, J., Wang, Z., Pei, T., Yang, H., ... and Xu, X. **2020**. Genome-wide analyses of the genetic screening of C2H2-type zinc finger transcription factors and abiotic and biotic stress responses in tomato (*Solanum lycopersicum*) based on RNA-Seq data. Front. Genet. **11**:540. https://doi.org/10.3389/fgene.2020.00540.

Zheng, F., Cui, L., Li, C., Xie, Q., Ai, G., Wang, J., et al. **2021**. Hair interacts with SlZFP8-like to regulate the initiation and elongation of trichomes by modulating SlZFP6 expression in tomato. J. Exp. Bot. **73**:228-244. https://doi.org/10.1093/jxb/erab417.

Zheng, X., Koopmann, B., Ulber, B., and von Tiedemann, A. **2020**. A global survey on diseases and pests in oilseed rape—current challenges and innovative strategies of control. Front. Agron. **2**:908. https://doi.org/10.3389/fagro.2020.59 0908.

Zhong H., Guo Q.Q., Chen L., Ren F., Wang Q.Q., Zheng Y., and Li X.B. **2012**. Two Brassica napus genes encoding NAC transcription factors are involved in response to high-salinity stress. Plant Cell Rep. **31**(11):1991-2003. https://doi.org/10.1007/s00299-012-1311-3.

Zhou, J. M., Trifa, Y., Silva, H., Pontier, D., Lam, E., Shah, J., et al. **2000**. NPR1 differentially interacts with members of the TGA/OBF family of transcription factors that bind an element of the PR-1 gene required for induction by salicylic acid. Mol. Plant Microbe Interact. **13**:191-202. https://doi.org/10.1094/MPMI.2000.13.2.191.

Zhou, J.M. and Zhang, Y. **2020**. Plant immunity: danger perception and signaling. Cell. **181**(5): 978-989.

Zhou, L., Zhang, L., He, Y., Liu, F., Li, M., Wang, Z., and et al. **2014**. Isolation and characterization of bacterial isolates for biological control of clubroot on Chinese cabbage. Eur. J. Plant Pathol. **140**:159-168. https://doi.org/10.1007/s10658-014-0451-4.

Zhou, Q., Galindo-González, L., Hwang, S.F., and Strelkov, S.E. **2021**. Application of genomics and transcriptomics to accelerate development of clubroot resistant canola. Can. J. Plant Pathol. **43**:189-208. https://doi.org/10.1080/07060661.2020.1794541.

Zhou, Q., Galindo-Gonzalez, L., Manolii, V., Hwang, S.F., and Strelkov, S.E. **2020**.Comparative transcriptome analysis of rutabaga (*Brassica napus*) cultivars indicates activation of salicylic acid and ethylene-mediated defenses in response to *Plasmodiophora brassicae*. Int. J. Mol. Sci. **21**(21): 8381. https://doi.org/10.3390/ijms21218381.

Zhu J.K. **2016**. Abiotic stress signaling and responses in plants. Cell, **167**(2):313-324. https://doi.org/10.1016/j.cell.2016.08.029.

Zhu, H., Zhai, W., Li, X., and Zhu, Y. **2019b**. Two QTLs controlling Clubroot resistance identified from bulked segregant sequencing in Pakchoi (*Brassica campestris* ssp. *chinensis* Makino). Sci. Rep. **9**(1): 9228. https://doi.org/10.1038/s41598- 019-44724-z.

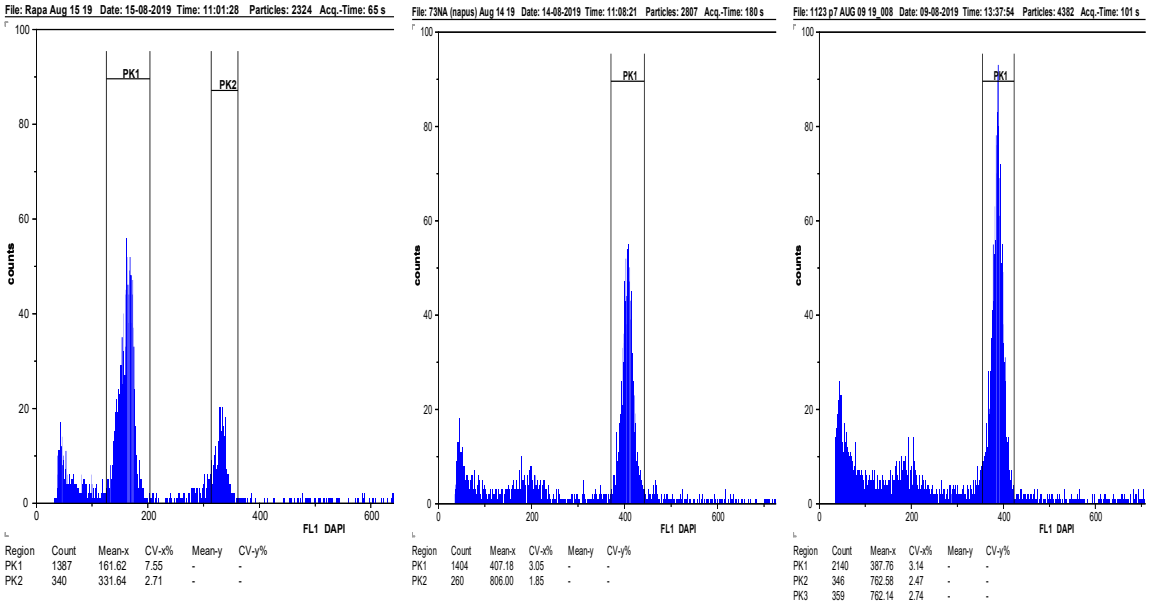
Zhu, J., Chen, J., Gao, F., Xu, C., Wu, H., Chen, K., Si, Z., Yan, H., and Zhang, T. **2017**. Rapid mapping and cloning of the virescent-1 gene in cotton by bulked segregant analysis–next generation sequencing and virus-induced gene silencing strategies. *J. Exp. Bot.* **68**(15): 4125-4135. https://doi.org/10.1093/jxb/erx240.

Zhu, M., He, Y., Li, Y., Ren, T., Liu, H., Huang, J., and et al. **2020**. Two new biocontrol agents against clubroot caused by *Plasmodiophora brassicae*. Front. Microbiol. **10**:3099. https://doi.org/10.3389/fmicb.2019.03099

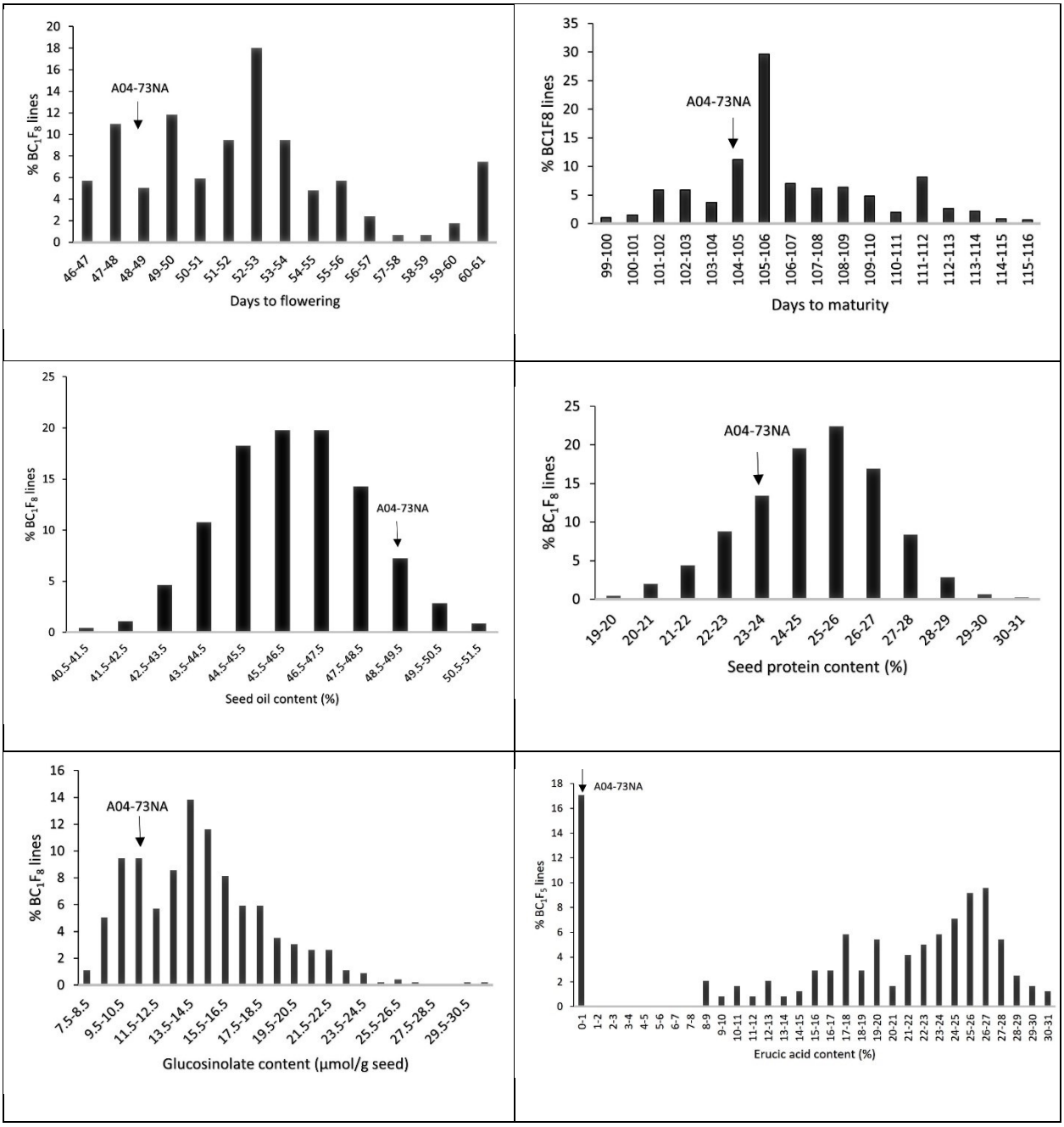
Zipfel, C. **2014**. Plant pattern-recognition receptors. Trends in immunology. **35**(7):345-351. https://doi.org/10.1016/j.it.2014.05.004.

Zipfel, C. and Robatzek, S. **2010**. Pathogen-associated molecular pattern-triggered immunity: veni, vidi…? Plant Physiol. **154**(2): 551-554. https://doi.org/10.1104/pp.110.161547.

Appendices

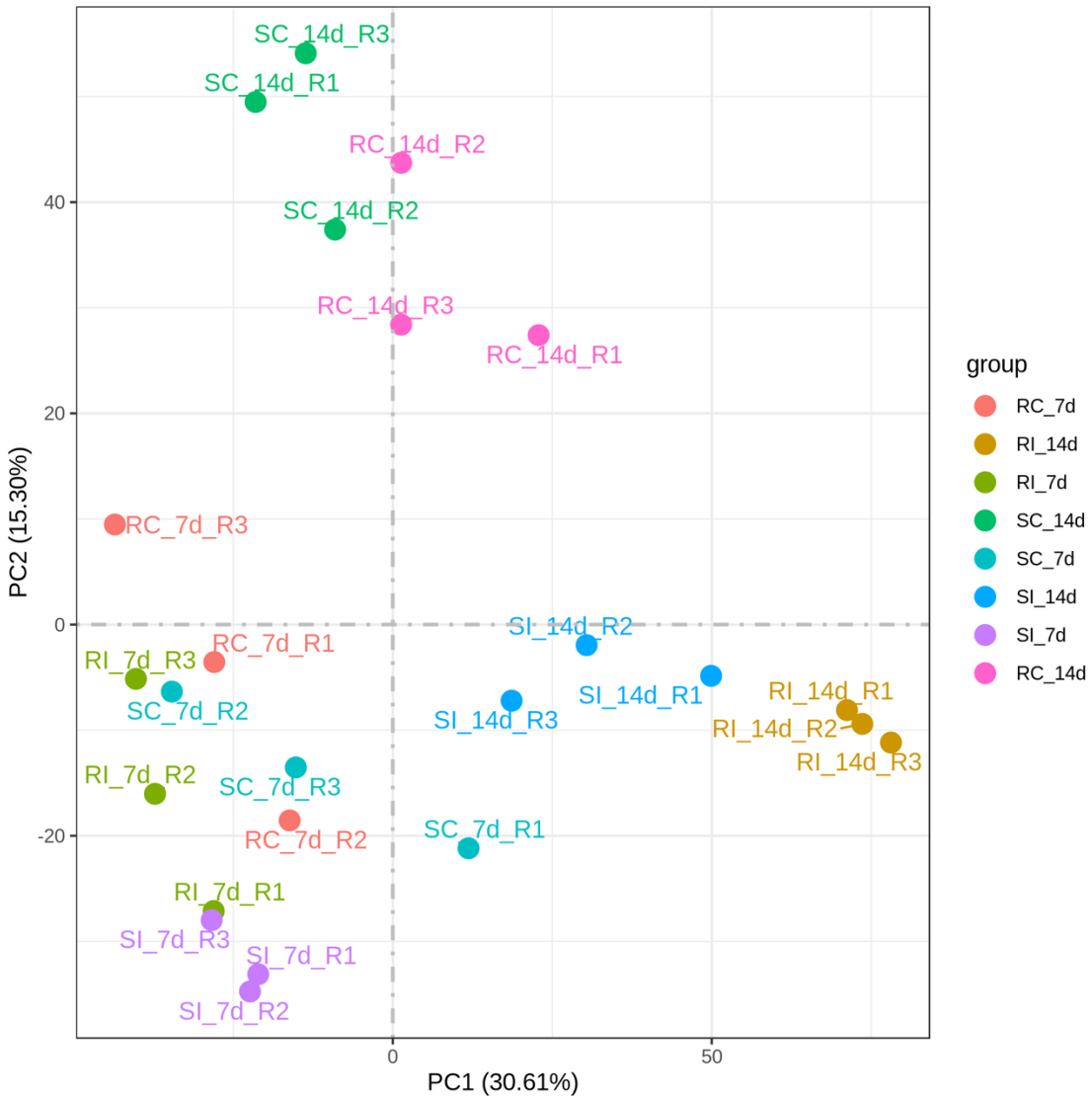


Supplementary Figure S2.1 DNA histogram showing relative nuclear DNA content in the *Brassica rapa* and *B. napus* parents and in a representative BC_1F_6 plant. The X-axis shows relative quantity of DNA in a cell and the Y-axis shows the number of cells counted.

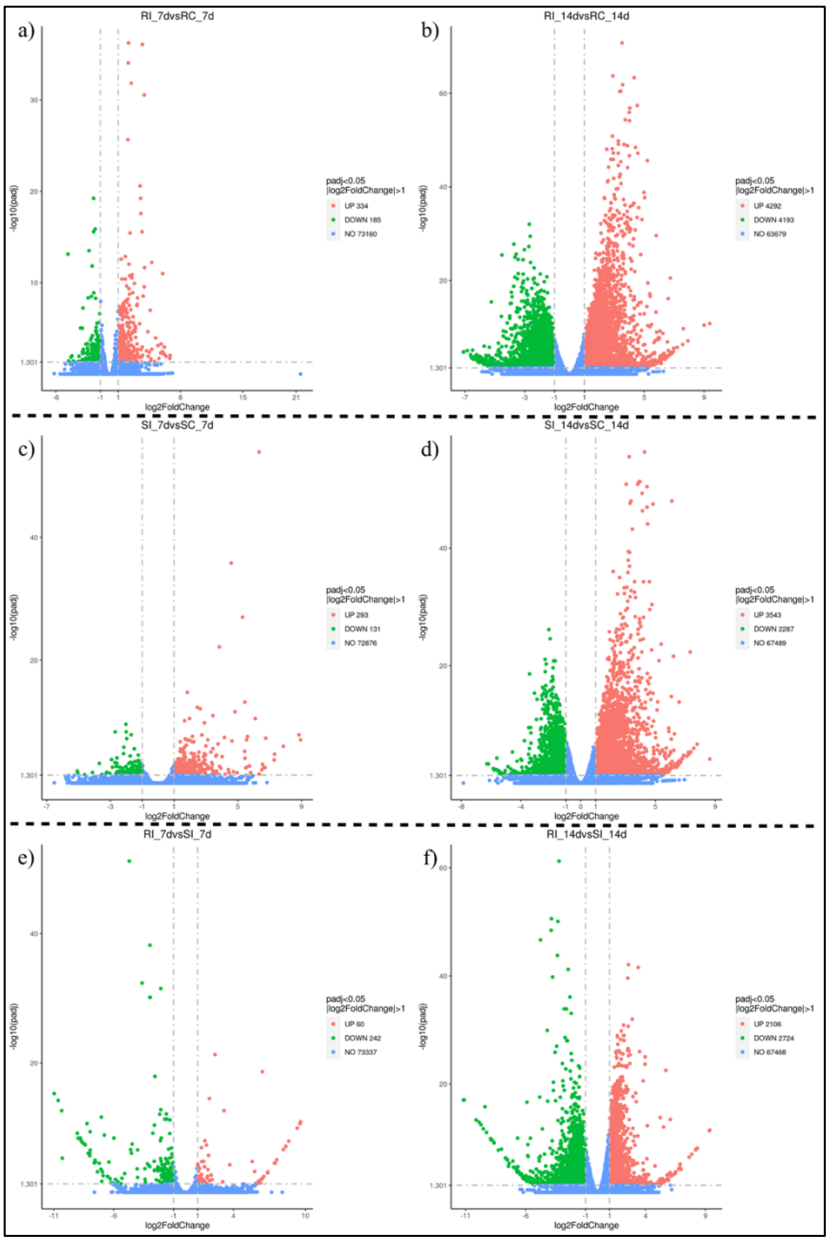


Supplementary Figure S2.2 Distribution of the BC_1F_8 lines of (*Brassica napus* × *B. rapa*) × *B. napus* interspecific cross, grown in nursery plots in 2020, for days to flowering (n = 452 lines), days to maturity (n = 452), and seed oil (n = 450), protein (n = 450), and glucosinolate (n = 450) contents. Erucic acid content in seed oil of the BC_1F_5 plants (n = 240) grown in greenhouse also presented. The value for the *B. napus* parent A04-73NA are indicated by vertical arrow.

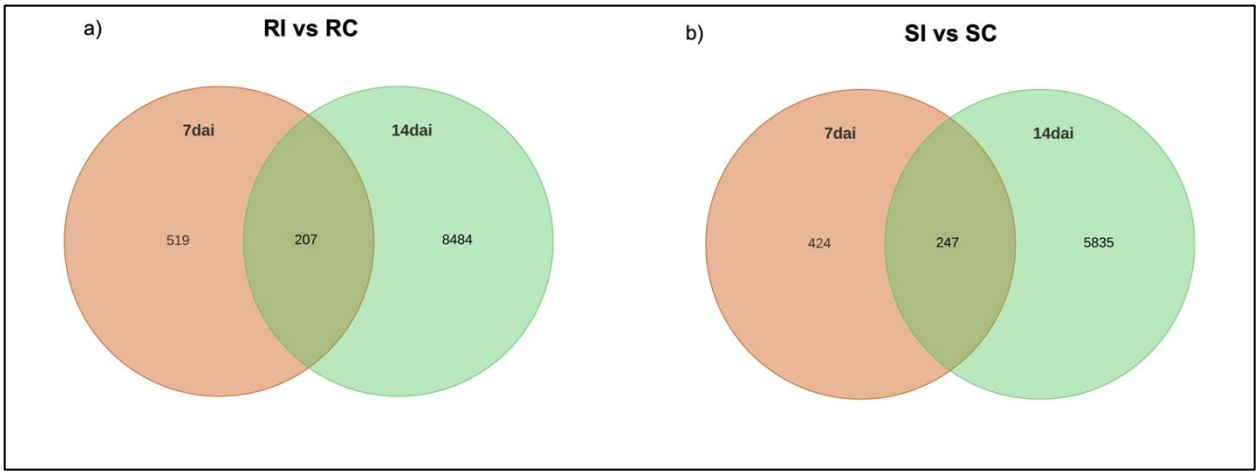
Supplementary Figure S4.1 The root phenotypes of uninoculated and inoculated clubroot resistant and susceptible *Brassica napus* near isogenic lines (NILs) at 7- and 14-days after infection (dai); inoculation done with *Plasmodiophora brassicae* pathotype 3H. RC: resistant control; RI: resistant inoculated; SC: susceptible control; SI: susceptible inoculated. The red circle indicates the development of galls on the roots.



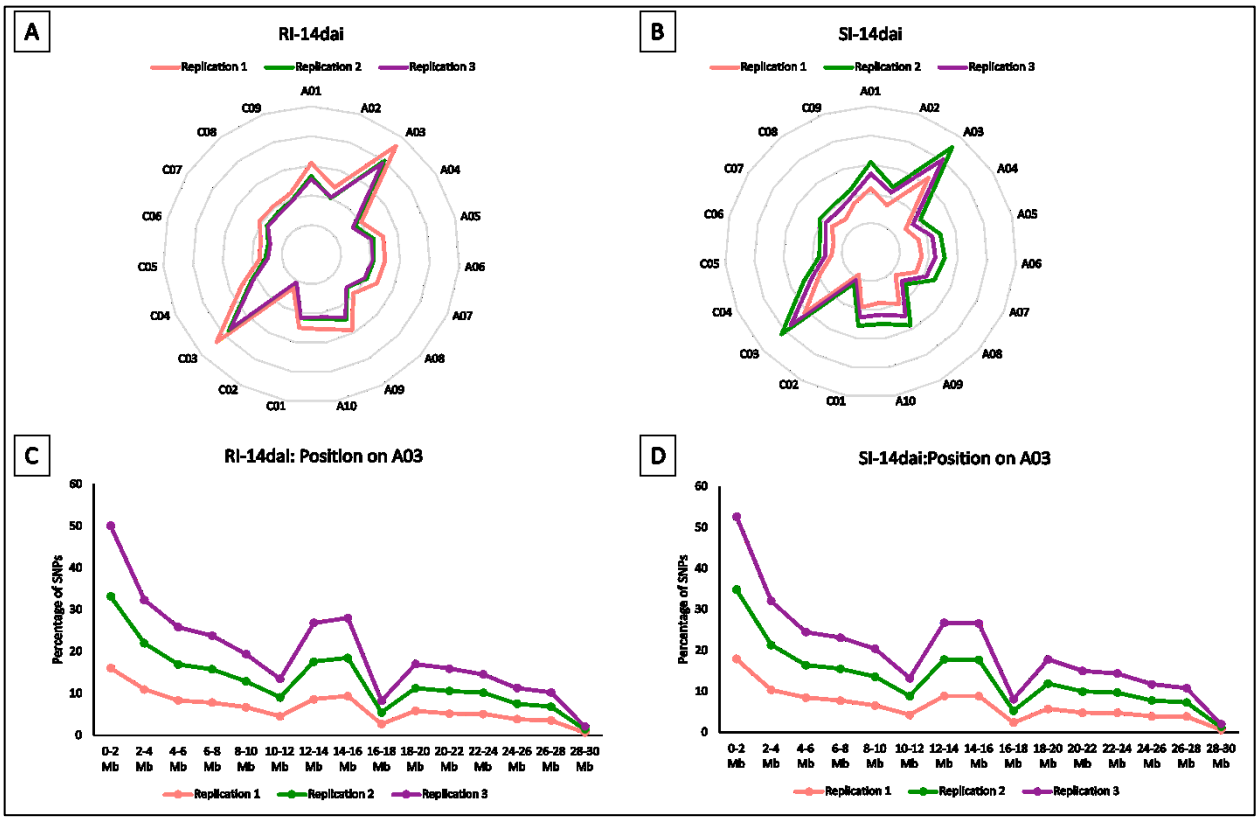
Supplementary Figure S4.2 Principal Component Analysis of the RNA-Seq data based on FPKM values of clubroot resistant and susceptible *Brassica napus* near isogenic lines (NILs) in response to *Plasmodiophora brassicae* pathotype 3H infection; where RC denotes the resistant control, RI is resistant inoculated, SC is susceptible control, SI is susceptible inoculated sample; 7d and 14d represents the two time points at 7 days after inoculation and 14 days after inoculation respectively; R1, R2 and R3 are the three biological replicates.



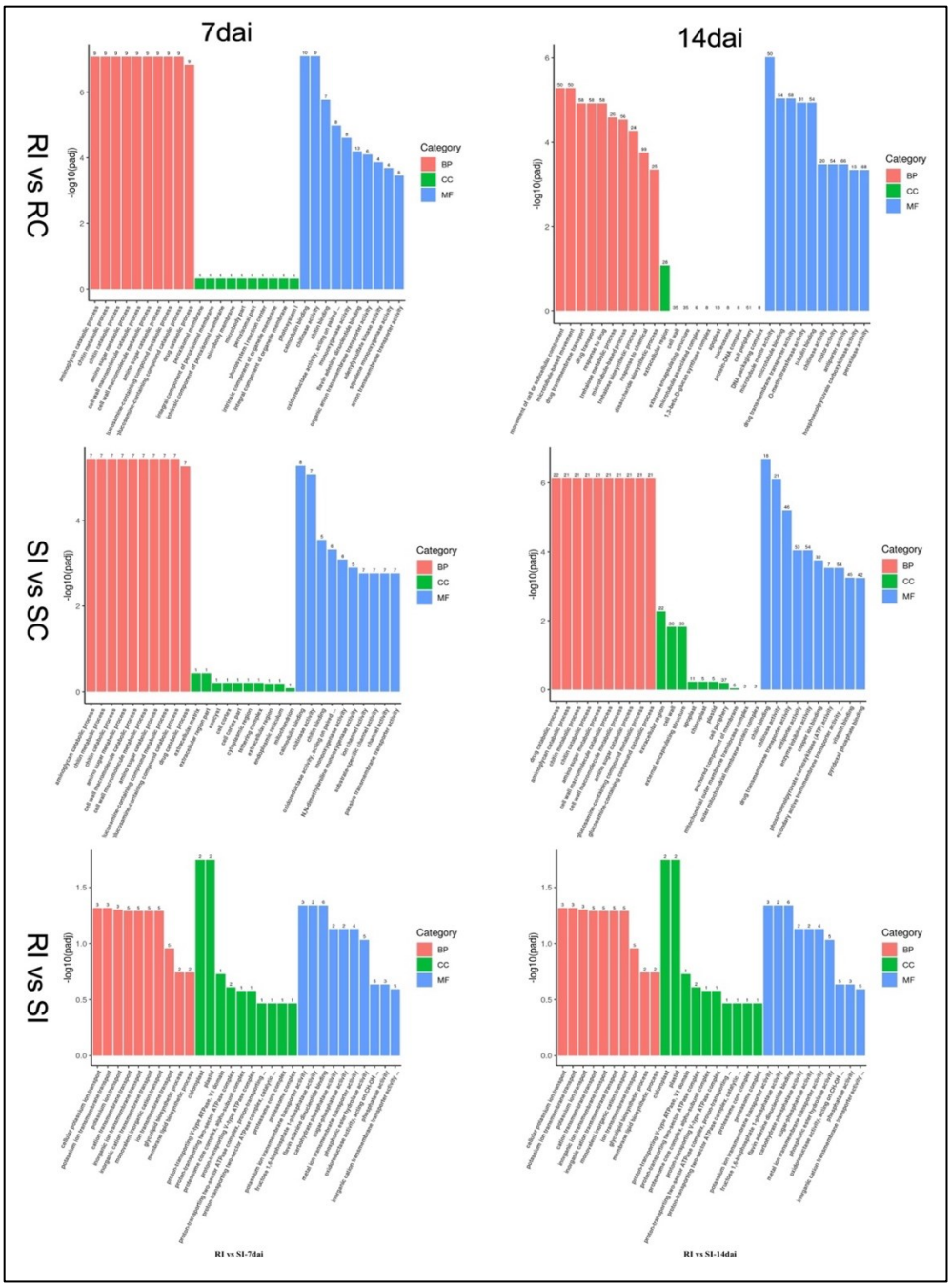
Supplementary Figure S4.3 Volcano plots of the differentially expressed genes in root samples of *Brassica napus* near isogenic lines (NILs) inoculated with *Plasmodiophora brassicae* pathotype 3H in three combinations at two time points. a) RI vs. RC 7 dai; b) RI vs. RC 14 dai; c) SI vs. SC 7 dai; d) SI vs. SC 14 dai; e) RI vs. SI 7 dai; f) RI vs. SI 14 dai; where RC is the resistant control, RI is resistant inoculated, SC is susceptible control, SI is susceptible inoculated sample; dai refers to the days after inoculation. The x-axis and y-axis, respectively, represent the log fold change and statistically significant degree of change in gene expression. The scattered dots represent the genes, where the blue dots represent the non-significant genes, red dots represent the upregulated genes, and green dots represents the downregulating genes.



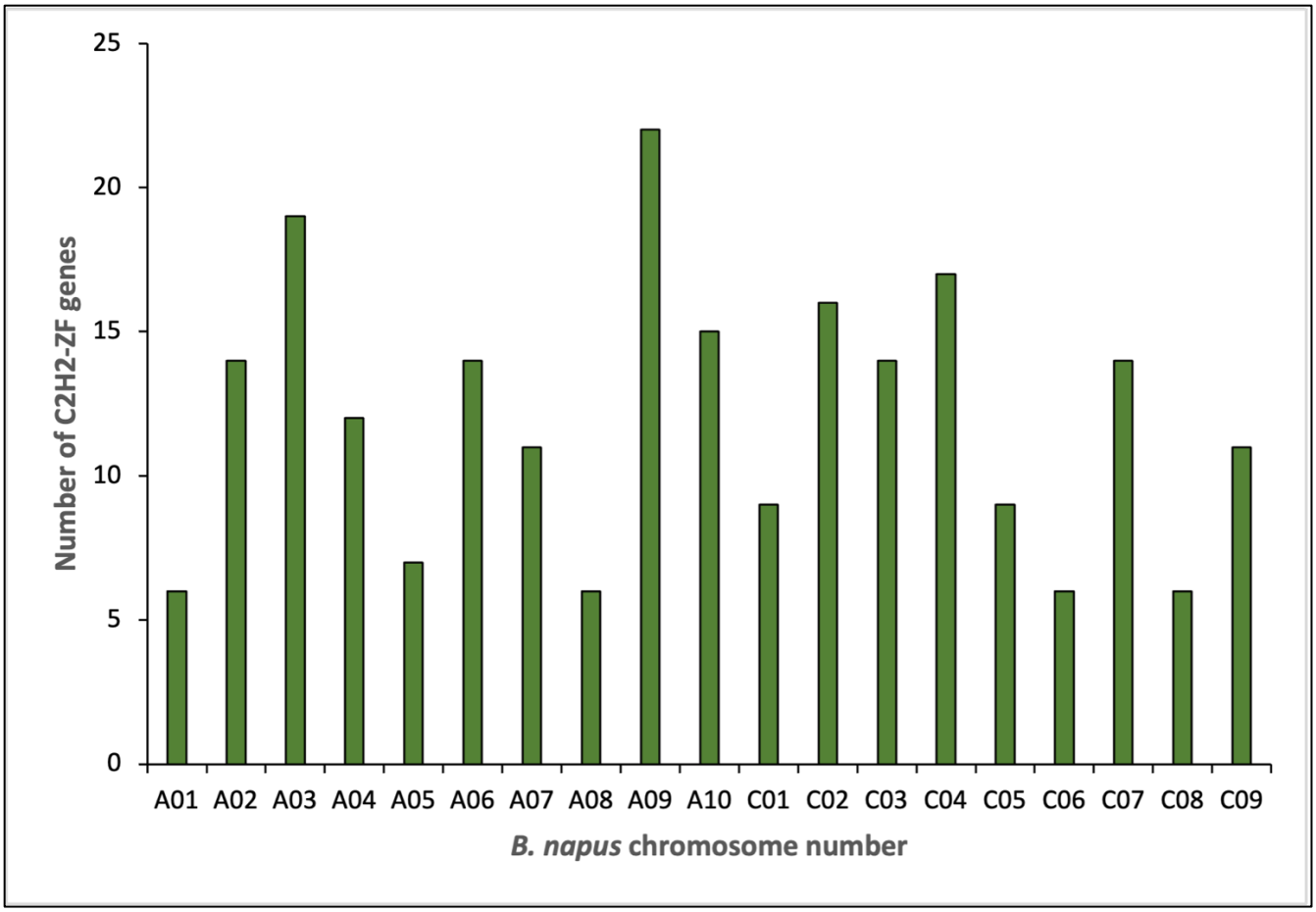
Supplementary Figure S4.4 Venn diagrams showing the significant differentially expressed genes (DEGs) (p-value < 0.05) in infected vs. control root samples of *Brassica napus* near isogenic lines (NILs) inoculated with *Plasmodiophora brassicae* pathotype 3H at 7 days after inoculation (dai) and 14 dai. (a) Inoculated resistant NIL vs. resistant control, and (b) Inoculated susceptible NIL vs. susceptible control. The overlapping regions shows the number of common DEGs from the comparisons.



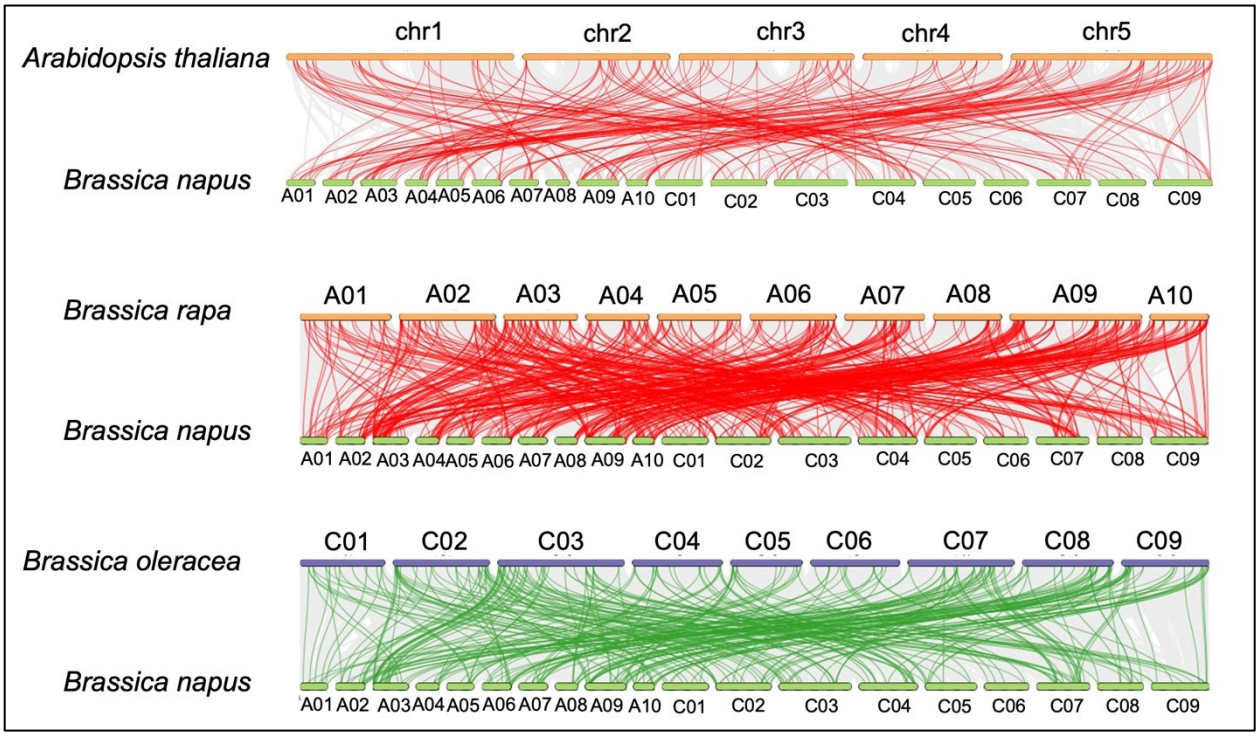
Supplementary Figure S4.5 Frequency distribution of SNPs from RNA-seq data of the resistant infected (RI) and susceptible infected (SI) root samples at 14 days after inoculation (dai) of *Brassica napus* near isogenic lines (NILs) inoculated with *Plasmodiophora brassicae* pathotype 3H. (A, B) The distribution of the SNPs on different *B. napus* chromosomes, and (C, D) the distribution of the SNPs of the chromosome A03.



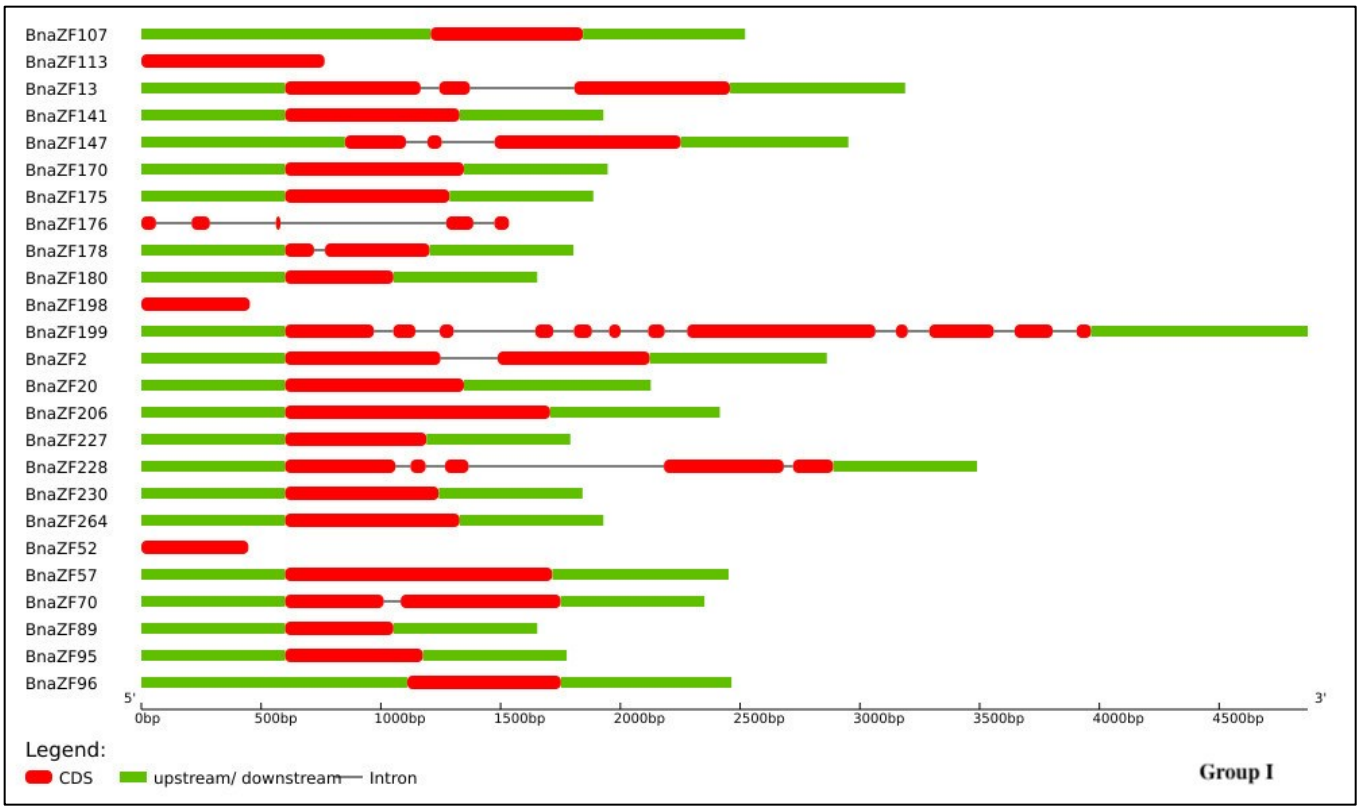
Supplementary Figure S4.6 The top 10 enriched gene ontology (GO) terms of each of the biological process (BP), cellular components (CC) and molecular function (MF) category for the differentially expressed genes (DEGs) identified from RI vs. RC, SI vs. SC and RI vs. SI comparisons of *Brassica napus* near isogenic lines (NILs) at 7- and 14-days after inoculation (dai) with *Plasmodiophora brassicae* pathotype 3H. RI = inoculated resistant NILs, RC = non-inoculated (control) resistant NILs; SI = inoculated susceptible NILs, and SC = non-inoculated (control) susceptible NILs. Different colors represent different categories of GO terms. The values at the top of bars denote the gene count.

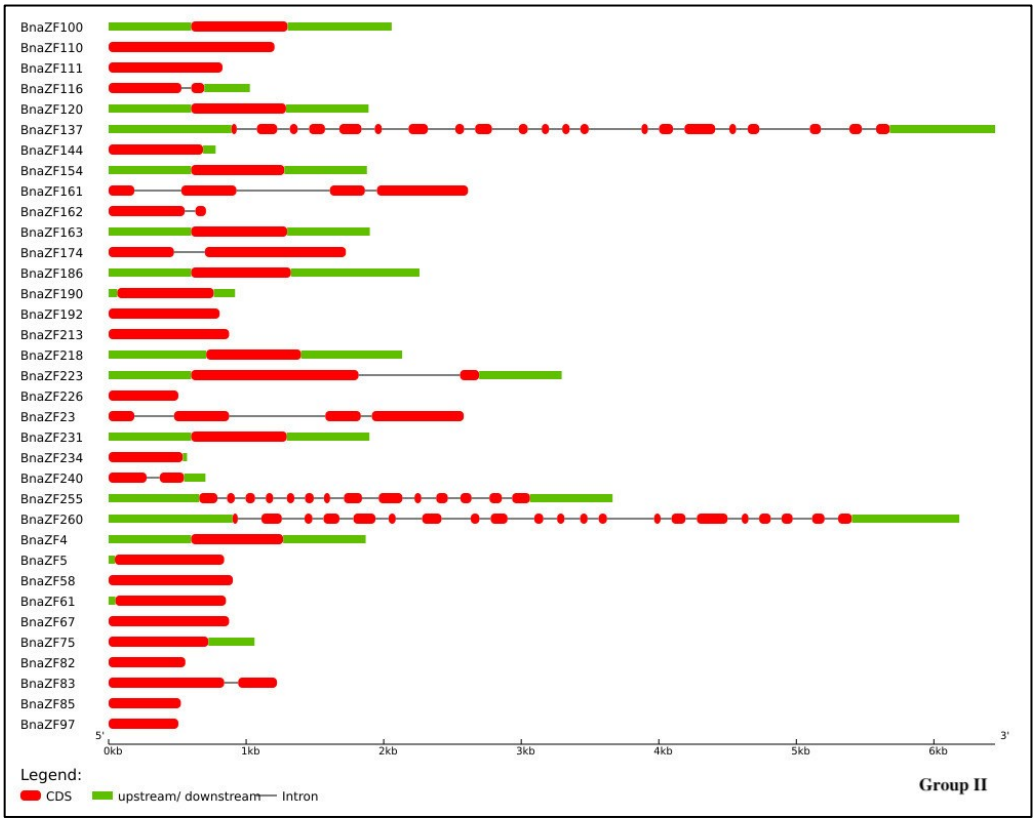


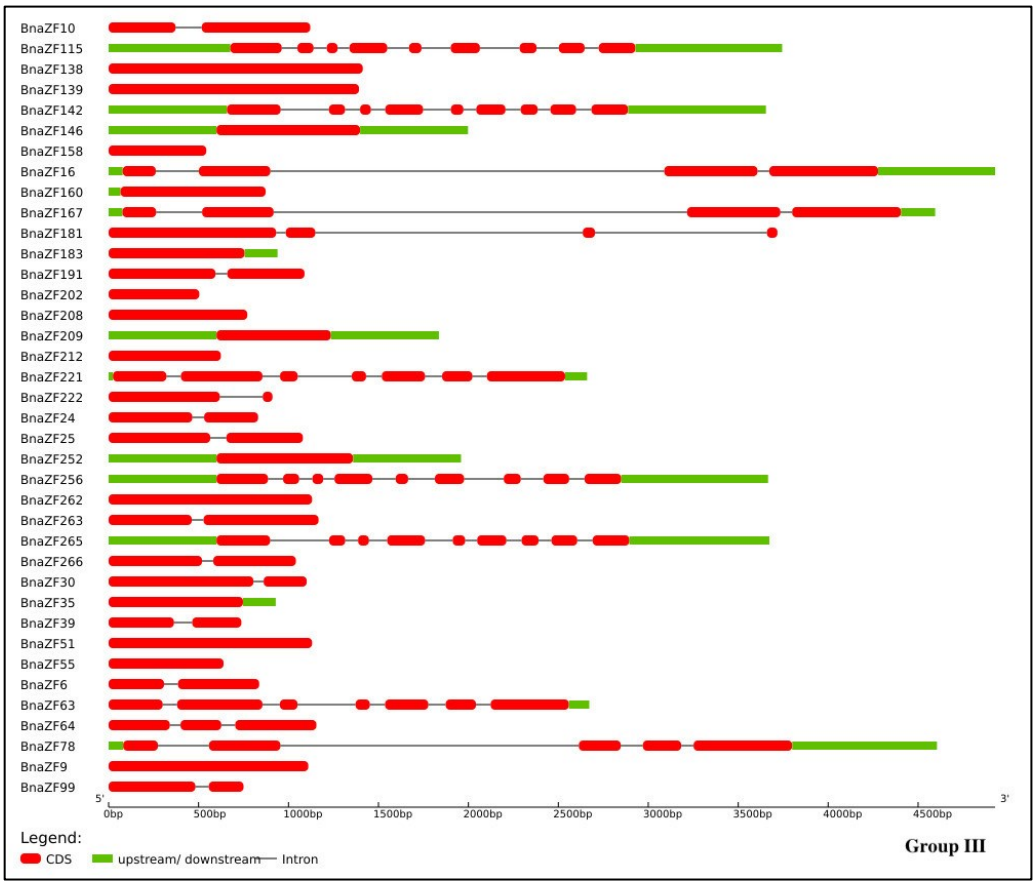
Supplementary Figure S6.1 Distribution of C2H2-ZF genes on different chromosomes of the A and C genomes of *Brassica napus.*



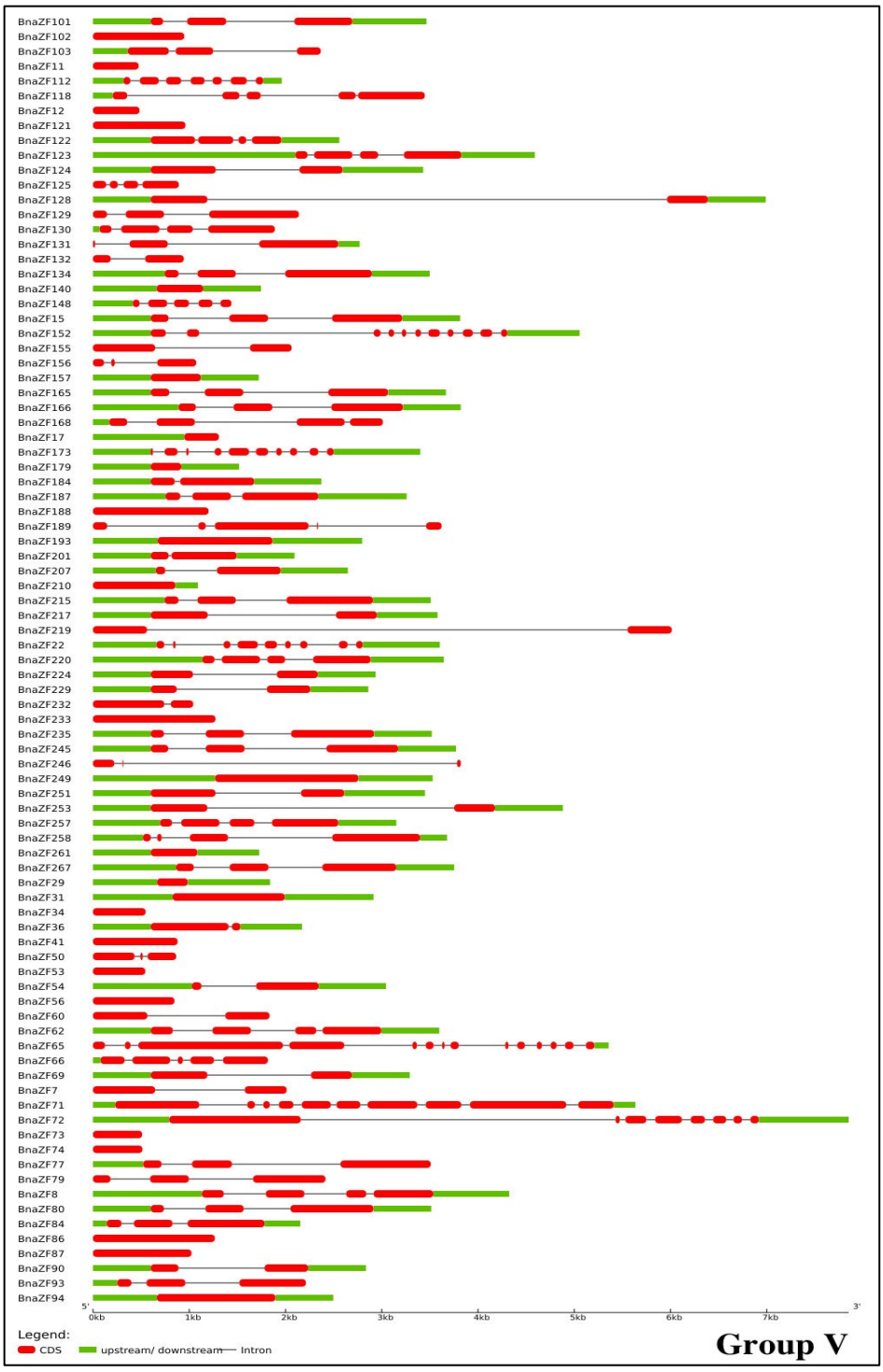
Supplementary Figure S6.2 Synteny analyses of C2H2-ZF genes between *Brassica napus* and three representative species. Grey lines represent the collinear regions between *Brassica napus* and other genomes, red and green lines indicate the syntenic C2H2-ZF gene pairs.


BnaZF107
BnaZF113
BnaZF13
BnaZF141
BnaZF147
BnaZF170
BnaZF175
BnaZF176
BnaZF178
BnaZF180
BnaZF198
BnaZF199
BnaZF2
BnaZF20
BnaZF206
BnaZF227
BnaZF228
BnaZF230
BnaZF264
BnaZF52
BnaZF57
BnaZF70
BnaZF89
BnaZF95
BnaZF96
5'
3'
0bp
500bp
1000bp
1500bp
2000bp
2500bp
3000bp
3500bp
4000bp
4500bp
Legend:
CDS
upstream/ downstream
Intron
Group I

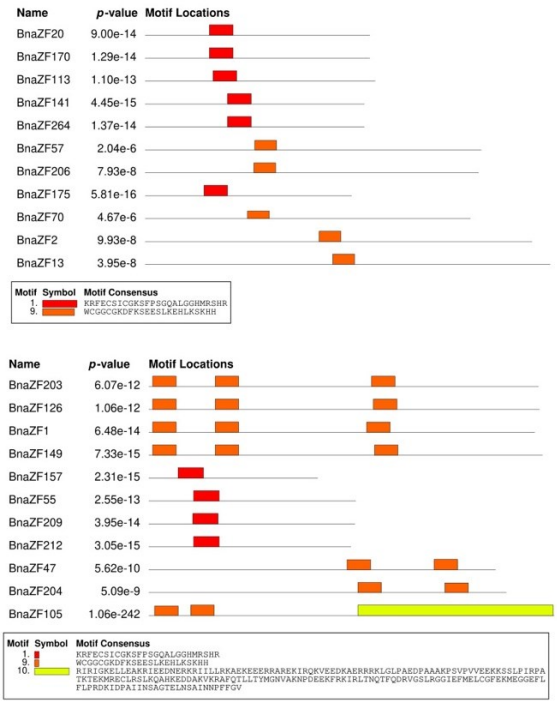

BnaZF100
BnaZF110
BnaZF111
BnaZF116
BnaZF120
BnaZF137
BnaZF144
BnaZF154
BnaZF161
BnaZF162
BnaZF163
BnaZF174
BnaZF186
BnaZF190
BnaZF192
BnaZF213
BnaZF218
BnaZF223
BnaZF226
BnaZF23
BnaZF231
BnaZF234
BnaZF240
BnaZF255
BnaZF260
BnaZF4
BnaZF5
BnaZF58
BnaZF61
BnaZF67
BnaZF75
BnaZF82
BnaZF83
BnaZF85
BnaZF97
5'
3'
0kb
1kb
2kb
3kb
4kb
5kb
6kb
Legend:
CDS
upstream/ downstream
Intron
Group II


BnaZF10
BnaZF115
BnaZF138
BnaZF139
BnaZF142
BnaZF146
BnaZF158
BnaZF16
BnaZF160
BnaZF167
BnaZF181
BnaZF183
BnaZF191
BnaZF202
BnaZF208
BnaZF209
BnaZF212
BnaZF221
BnaZF222
BnaZF24
BnaZF25
BnaZF252
BnaZF256
BnaZF262
BnaZF263
BnaZF265
BnaZF266
BnaZF30
BnaZF35
BnaZF39
BnaZF51
BnaZF55
BnaZF6
BnaZF63
BnaZF64
BnaZF78
BnaZF9
BnaZF99
5'
3'
0bp
500bp
1000bp
1500bp
2000bp
2500bp
3000bp
3500bp
4000bp
4500bp
Legend:
CDS
upstream/ downstream
Intron
Group III


BnaZF1
BnaZF104
BnaZF105
BnaZF106
BnaZF108
BnaZF109
BnaZF114
BnaZF117
BnaZF119
BnaZF126
BnaZF127
BnaZF133
BnaZF135
BnaZF136
BnaZF14
BnaZF143
BnaZF145
BnaZF149
BnaZF150
BnaZF151
BnaZF159
BnaZF164
BnaZF169
BnaZF171
BnaZF172
BnaZF177
BnaZF18
BnaZF182
BnaZF185
BnaZF19
BnaZF194
BnaZF195
BnaZF196
BnaZF197
BnaZF200
BnaZF203
BnaZF204
BnaZF205
BnaZF21
BnaZF211
BnaZF214
BnaZF216
BnaZF225
BnaZF236
BnaZF237
BnaZF238
BnaZF239
BnaZF241
BnaZF242
BnaZF243
BnaZF244
BnaZF247
BnaZF248
BnaZF250
BnaZF254
BnaZF259
BnaZF26
BnaZF27
BnaZF28
BnaZF3
BnaZF32
BnaZF33
BnaZF37
BnaZF38
BnaZF40
BnaZF42
BnaZF43
BnaZF44
BnaZF45
BnaZF47
BnaZF48
BnaZF49
BnaZF59
BnaZF68
BnaZF76
BnaZF81
BnaZF88
BnaZF91
BnaZF92
BnaZF98
5'
3'
0kb
1kb
2kb
3kb
4kb
5kb
6kb
7kb
8kb
9kb
Legend:
CDS
upstream/ downstream
Intron
Group IV

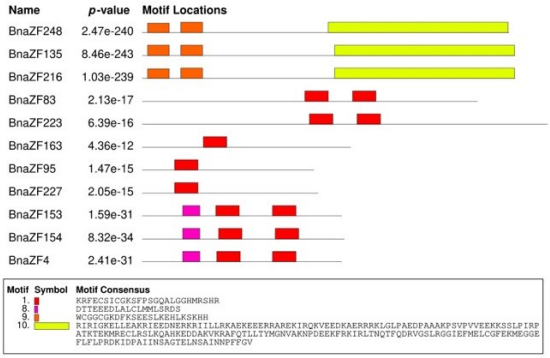


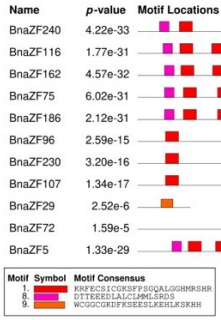
Supplementary Figure S6.3 Gene structure of the *Brassica napus* C2H2-ZF genes displaying organization of the exons and introns. Analysis was performed using Gene Structure Display Server (GSDS2.0). Red boxes represent exons, black lines represent introns and green boxes represent the untranslated regions (UTRs). The scale at the bottom can be used to estimate the length of the protein.

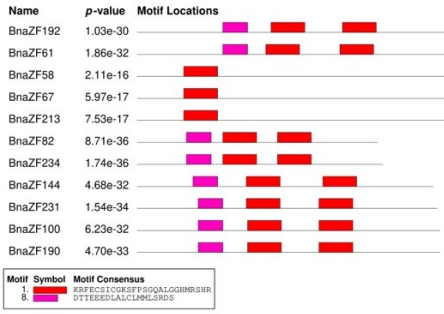

Name p-value Motif Locations
BnaZF20 9.00e-14
BnaZF170 1.29e-14
BnaZF113 1.10e-13
BnaZF141 4.45e-15
BnaZF264 1.37e-14
BnaZF57 2.04e-6
BnaZF206 7.93e-8
BnaZF175 5.81e-16
BnaZF70 4.67e-6
BnaZF2 9.93e-8
BnaZF13 3.95e-8
Motif Symbol Motif Consensus
1. KRFECSICGKSFPSGQALGGHMRSHR
9. WCGGCGKDFKSEESLKEHLKSKHH

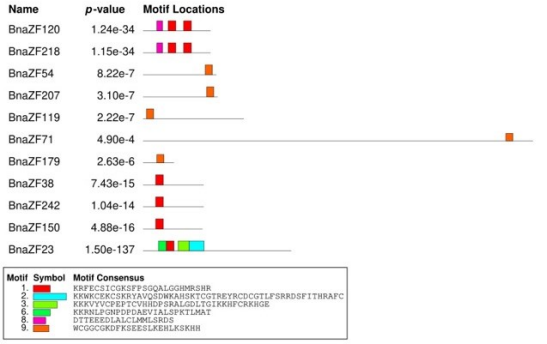

Name p-value Motif Locations
BnaZF42 3.40e-13
BnaZF59 2.28e-8
BnaZF194 6.70e-9
BnaZF91 4.69e-5
BnaZF3 2.20e-28
BnaZF151 4.88e-28
BnaZF44 2.59e-27
BnaZF241 3.00e-28
BnaZF106 6.01e-26
BnaZF254 3.04e-25
BnaZF48 2.95e-12
Motif Symbol Motif Consensus
1. KRFECSICGKSFPSGQALGGHMRSHR
8. DTTEEEDLALCLMMLSRDS
9. WCGGCGKDFKSEESLKEHLKSKHH

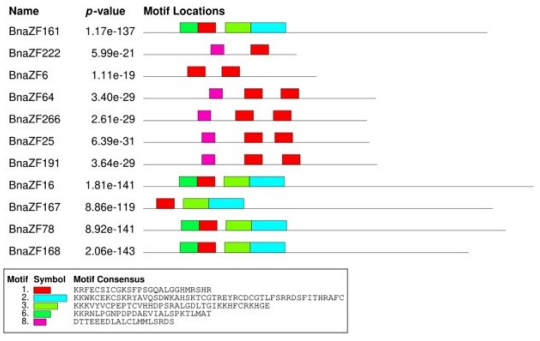
Name p-value Motif Locations
BnaZF203 6.07e-12
BnaZF126 1.06e-12
BnaZF1 6.48e-14
BnaZF149 7.33e-15
BnaZF157 2.31e-15
BnaZF55 2.55e-13
BnaZF209 3.95e-14
BnaZF212 3.05e-15
BnaZF47 5.62e-10
BnaZF204 5.09e-9
BnaZF105 1.06e-242
Motif Symbol Motif Consensus
1. KRFECSICGKSFPSGQALGGHMRSHR
9. WCGGCGKDFKSEESLKEHLKSKHH
10. RIRIGKELLEAKRIEEDNEKKRIILLRKAEKEEERRAREKIRQKVEEDKAERRRKLGLPAEDPAAAKPSVPVVEEKKSSLPIRPA TKTEKMRECLRSLKQAHKEDDAKVKRAFQTLLTYMGNVAKNPDEEKFRKIRLTNQTFQDRVGSLRGGIEFMELCGFEKMEGGEFL FLPRDKIDPAIINSAGTELNSAINNPFFGV

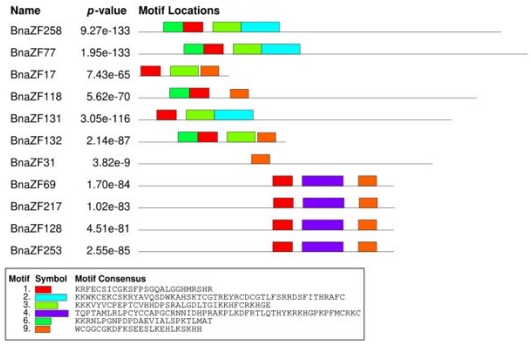

Name p-value Motif Locations
BnaZF248 2.47e-240
BnaZF135 8.46e-243
BnaZF216 1.03e-239
BnaZF83 2.13e-17
BnaZF223 6.39e-16
BnaZF163 4.36e-12
BnaZF95 1.47e-15
BnaZF227 2.05e-15
BnaZF153 1.59e-31
BnaZF154 8.32e-34
BnaZF4 2.41e-31
Motif Symbol Motif Consensus
1. KRFECSICGKSFPSGQALGGHMRSHR
8. DTTEEEDLALCLMMLSRDS
9. WCGGCGKDFKSEESLKEHLKSKHH
10. RIRIGKELLEAKRIEEDNERKRIILLRKAEKEEERRAREKIRQKVEEDKAERRRKLGLPAEDPAAAKPSVPVVEEKKSSLPIRP ATKTEKMRECLRSLKQAHKEDDAKVKRAFQTLLTYMGNVAKNPDEEKFRKIRLTNQTFQDRVGSLRGGIEFMELCGFEKMEGGE FLFLPRDKIDPAIINSAGTELNSAINNPFFGV

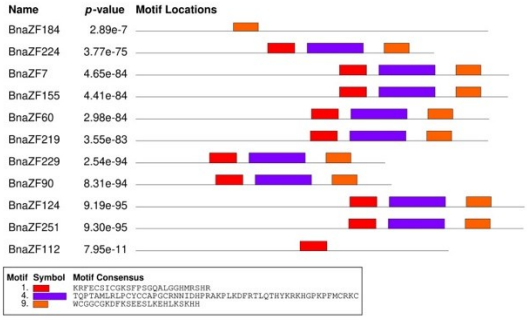

Name p-value Motif Locations
BnaZF240 4.22e-33
BnaZF116 1.77e-31
BnaZF162 4.57e-32
BnaZF75 6.02e-31
BnaZF186 2.12e-31
BnaZF96 2.59e-15
BnaZF230 3.20e-16
BnaZF107 1.34e-17
BnaZF29 2.52e-6
BnaZF72 1.59e-5
BnaZF5 1.33e-29
Motif Symbol Motif Consensus
1. KRFECSICGKSFPSGQALGGHMRSHR
8. DTTEEEDLALCLMMLSRDS
9. WCGGCGKDFKSEESLKEHLKSKHH

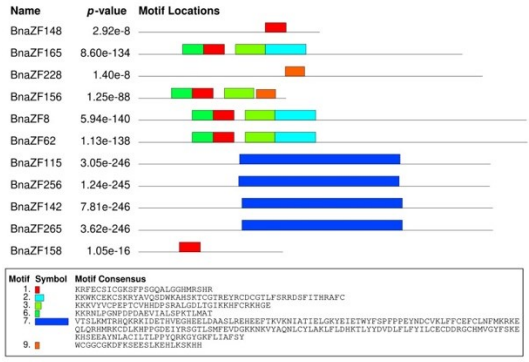

Name p-value Motif Locations
BnaZF192 1.03e-30
BnaZF61 1.86e-32
BnaZF58 2.11e-16
BnaZF67 5.97e-17
BnaZF213 7.53e-17
BnaZF82 8.71e-36
BnaZF234 1.74e-36
BnaZF144 4.68e-32
BnaZF231 1.54e-34
BnaZF100 6.23e-32
BnaZF190 4.70e-33
Motif Symbol Motif Consensus
1. KRFECSICGKSFPSGQALGGHMRSHR
8. DTTEEEDLALCLMMLSRDS

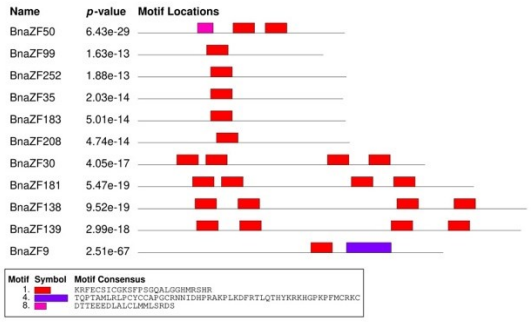

Name p-value Motif Locations
BnaZF120 1.24e-34
BnaZF218 1.15e-34
BnaZF54 8.22e-7
BnaZF207 3.10e-7
BnaZF119 2.22e-7
BnaZF71 4.90e-4
BnaZF179 2.63e-6
BnaZF38 7.43e-15
BnaZF242 1.04e-14
BnaZF150 4.88e-16
BnaZF23 1.50e-137
Motif Symbol Motif Consensus
1. KRFECSICGKSFPSGQALGGHMRSHR
2. KKWKCEKCSKRYAVQSDWKAHSKTCGTREYRCDCGTLFSRRDSFITHRAFC
3. KKKVYVCPEPTCVHHDPSRALGDLTGIKKHFCRKHGE
6. KKRNLPGNPDPDAEVIALSPKTLMAT
8. DTTEEEDLALCLMMLSRDS
9. WCGGCGKDFKSEESLKEHLKSKHH

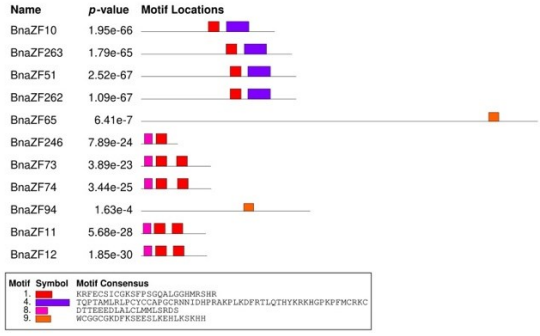

Name p-value Motif Locations
BnaZF161 1.17e-137
BnaZF222 5.99e-21
BnaZF6 1.11e-19
BnaZF64 3.40e-29
BnaZF266 2.61e-29
BnaZF25 6.39e-31
BnaZF191 3.64e-29
BnaZF16 1.81e-141
BnaZF167 8.86e-119
BnaZF78 8.92e-141
BnaZF168 2.06e-143
Motif Symbol Motif Consensus
1. KRFECSICGKSFPSGQALGGHMRSHR
2. KKWKCEKCSKRYAVQSDWKAHSKTCGTREYRCDCGTLFSRRDSFITHRAFC
3. KKKVYVCPEPTCVHHDPSRALGDLTGIKKHFCRKHGE
6. KKRNLPGNPDPDAEVIALSPKTLMAT
8. DTTEEEDLALCLMMLSRDS

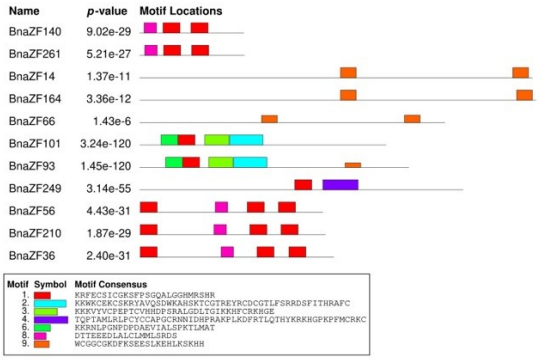

Name p-value Motif Locations
BnaZF258 9.27e-133
BnaZF77 1.95e-133
BnaZF17 7.43e-65
BnaZF118 5.62e-70
BnaZF131 3.05e-116
BnaZF132 2.14e-87
BnaZF31 3.82e-9
BnaZF69 1.70e-84
BnaZF217 1.02e-83
BnaZF128 4.51e-81
BnaZF253 2.55e-85
Motif Symbol Motif Consensus
1. KRFECSICGKSFPSGQALGGHMRSHR
2. KKWKCEKCSKRYAVQSDWKAHSKTCGTREYRCDCGTLFSRRDSFITHRAFC
3. KKKVYVCPEPTCVHHDPSRALGDLTGIKKHFCRKHGE
4. TQPTAMLRLPCYCCAPGCRNNIDHPRAKPLKDFRTLQTHYKRKHGPKPFMCRKC
6. KKRNLPGNPDPDAEVIALSPKTLMAT
9. WCGGCGKDFKSEESLKEHLKSKHH

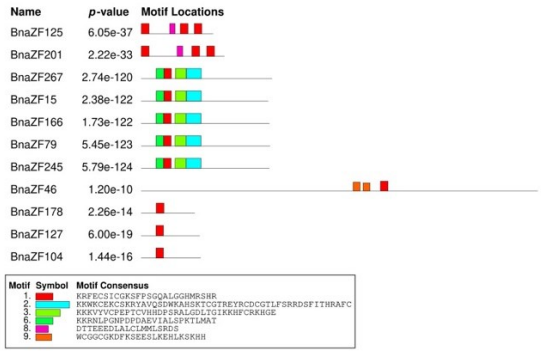

Name p-value Motif Locations
BnaZF184 2.89e-7
BnaZF224 3.77e-75
BnaZF7 4.65e-84
BnaZF155 4.41e-84
BnaZF60 2.98e-84
BnaZF219 3.55e-83
BnaZF229 2.54e-94
BnaZF90 8.31e-94
BnaZF124 9.19e-95
BnaZF251 9.30e-95
BnaZF112 7.95e-11
Motif Symbol Motif Consensus
1. KRFECSICGKSFPSGQALGGHMRSHR
4. TQPTAMLRLPCYCCAPGCRNNIDHPRAKPLKDFRTLQTHYKRKHGPKPFMCRKC
9. WCGGCGKDFKSEESLKEHLKSKHH

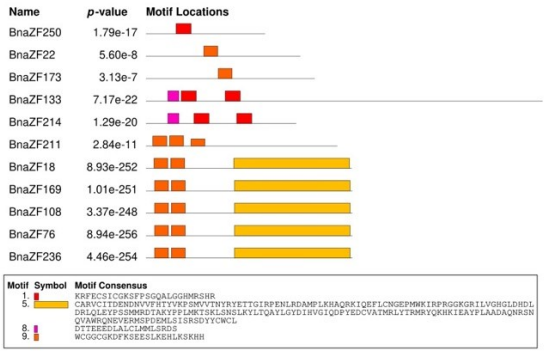

Name p-value Motif Locations
BnaZF148 2.92e-8
BnaZF165 8.60e-134
BnaZF228 1.40e-8
BnaZF156 1.25e-88
BnaZF8 5.94e-140
BnaZF62 1.13e-138
BnaZF115 3.05e-246
BnaZF256 1.24e-245
BnaZF142 7.81e-246
BnaZF265 3.62e-246
BnaZF158 1.05e-16
Motif Symbol Motif Consensus
1. KRFECSICGKSFPSGQALGGHMRSHR
2. KKWKCEKCSKRYAVQSDWKAHSKTCGTREYRCDCGTLFSRRDSFITHRAFC
3. KKKVYVCPEPTCVHHDPSRALGDLTGIKKHFCRKHGE
5. KKRNLPGNPDPDAEVIALSPKTLMAT
7. VTSLKMTRHQKRKIDETHVEGHEELDAASLREHEEFTKVKNIATIELGKYEIETWYFSPPPEYNDCVKLFFCEFCLNFMKRKE QLQRHMRKCDLKHPPGDEIYRSGTLSMFEVDGKKNKVYAQNLCYLAKLFLDHKTLYYDVDLFLFYILCECDDRGCHMVGYFSKE KHSEEAYNLACILTLPPYQRKGYGKFLIAFSY
9. WCGGCGKDFKSEESLKEHLKSKHH

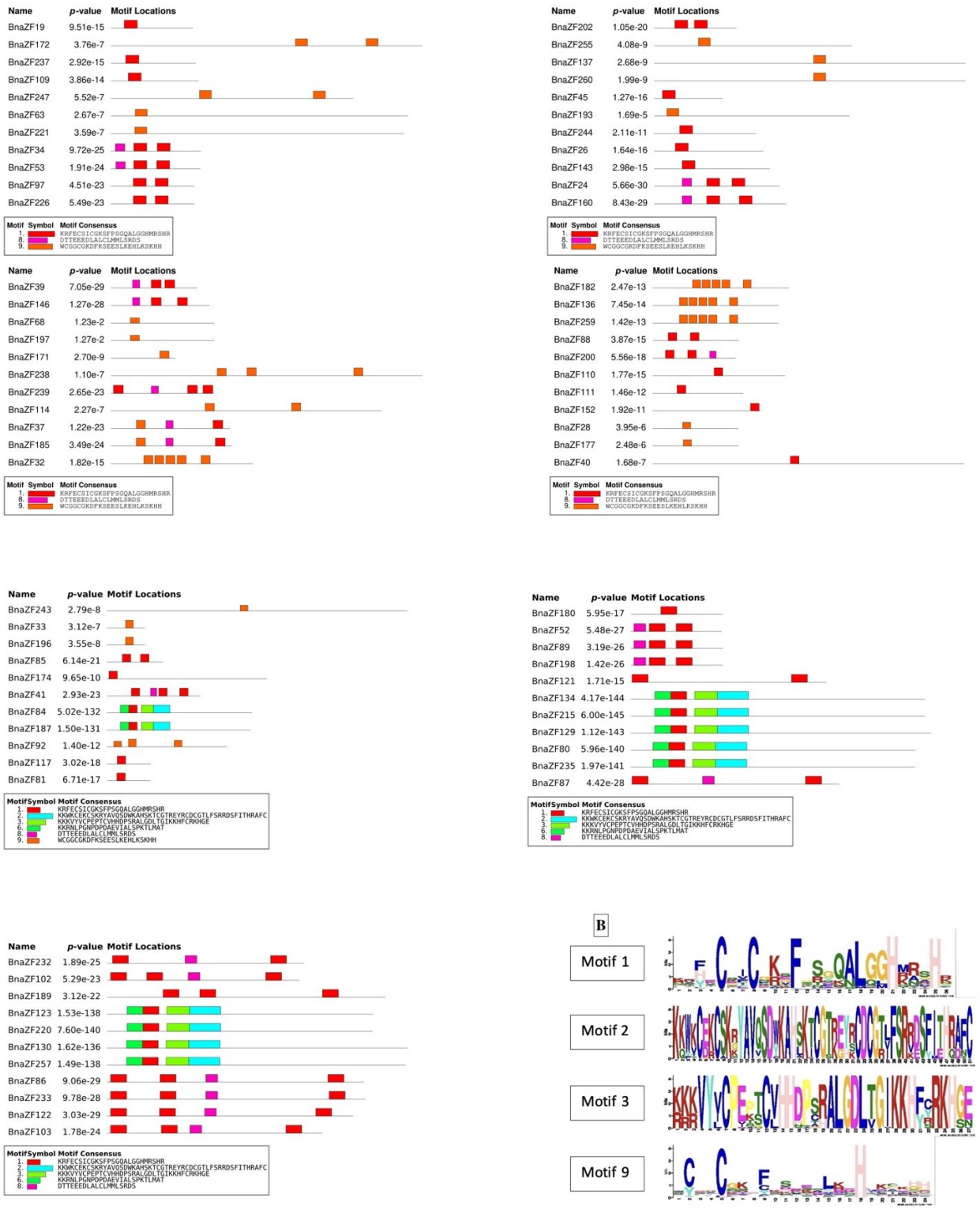

Name p-value Motif Locations
BnaZF50 6.43e-29
BnaZF99 1.63e-13
BnaZF252 1.88e-13
BnaZF35 2.03e-14
BnaZF183 5.01e-14
BnaZF208 4.74e-14
BnaZF30 4.05e-17
BnaZF181 5.47e-19
BnaZF138 9.52e-19
BnaZF139 2.99e-18
BnaZF9 2.51e-67
Motif Symbol Motif Consensus
1. KRFECSICGKSFPSGQALGGHMRSHR
4. TQPTAMLRLPCYCCAPGCRNNIDHPRAKPLKDFRTLQTHYKRKHGPKPFMCRKC
8. DTTEEEDLALCLMMLSRDS


Name p-value Motif Locations
BnaZF10 1.95e-66
BnaZF263 1.79e-65
BnaZF51 2.52e-67
BnaZF262 1.09e-67
BnaZF65 6.41e-7
BnaZF246 7.89e-24
BnaZF73 3.89e-23
BnaZF74 3.44e-25
BnaZF94 1.63e-4
BnaZF11 5.68e-28
BnaZF12 1.85e-30
Motif Symbol Motif Consensus
1. KRFECSICGKSFPSGQALGGHMRSHR
4. TQPTAMLRLPCYCCAPGCRNNIDHPRAKPLKDFRTLQTHYKRKHGPKPFMCRKC
8. DTTEEEDLALCLMMLSRDS
9. WCGGCGKDFKSEESLKEHLKSKHH

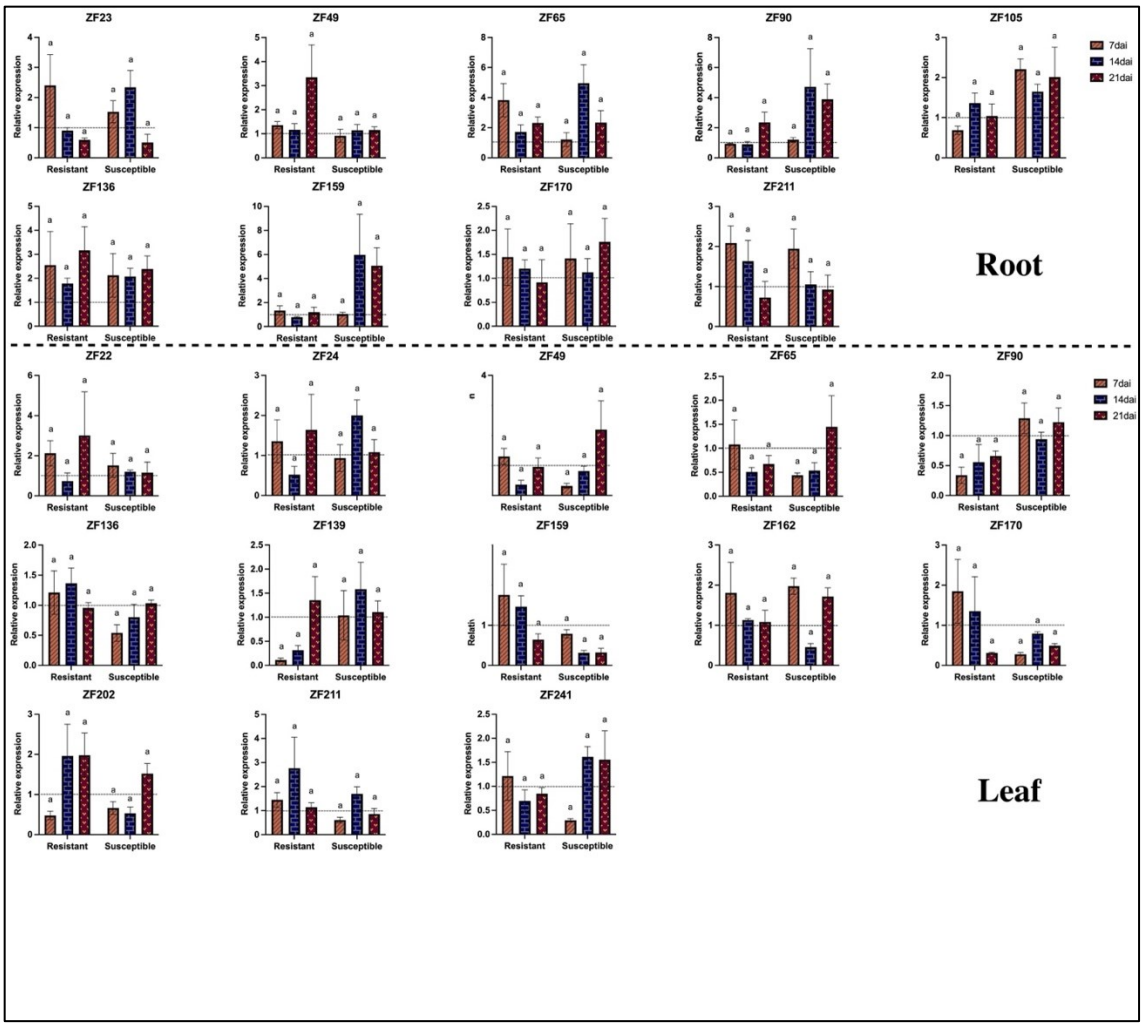

Name p-value Motif Locations
BnaZF140 9.02e-29
BnaZF261 5.21e-27
BnaZF14 1.37e-11
BnaZF164 3.36e-12
BnaZF66 1.43e-6
BnaZF101 3.24e-120
BnaZF93 1.45e-120
BnaZF249 3.14e-55
BnaZF56 4.43e-31
BnaZF210 1.87e-29
BnaZF36 2.40e-31
Motif Symbol Motif Consensus
1. KRFECSICGKSFPSGQALGGHMRSHR
2. KKWKCEKCSKRYAVQSDWKAHSKTCGTREYRCDCGTLFSRRDSFITHRAFC
3. KKKVYVCPEPTCVHHDPSRALGDLTGIKKHFCRKHGE
4. TQPTAMLRLPCYCCAPGCRNNIDHPRAKPLKDFRTLQTHYKRKHGPKPFMCRKC
6. KKRNLPGNPDPDAEVIALSPKTLMAT
8. DTTEEEDLALCLMMLSRDS
9. WCGGCGKDFKSEESLKEHLKSKHH


Name p-value Motif Locations
BnaZF125 6.05e-37
BnaZF201 2.22e-33
BnaZF267 2.74e-120
BnaZF15 2.38e-122
BnaZF166 1.73e-122
BnaZF79 5.45e-123
BnaZF245 5.79e-124
BnaZF46 1.20e-10
BnaZF178 2.26e-14
BnaZF127 6.00e-19
BnaZF104 1.44e-16
Motif Symbol Motif Consensus
1. KRFECSICGKSFPSGQALGGHMRSHR
2. KKWKCEKCSKRYAVQSDWKAHSKTCGTREYRCDCGTLFSRRDSFITHRAFC
3. KKKVYVCPEPTCVHHDPSRALGDLTGIKKHFCRKHGE
6. KKRNLPGNPDPDAEVIALSPKTLMAT
8. DTTEEEDLALCLMMLSRDS
9. WCGGCGKDFKSEESLKEHLKSKHH

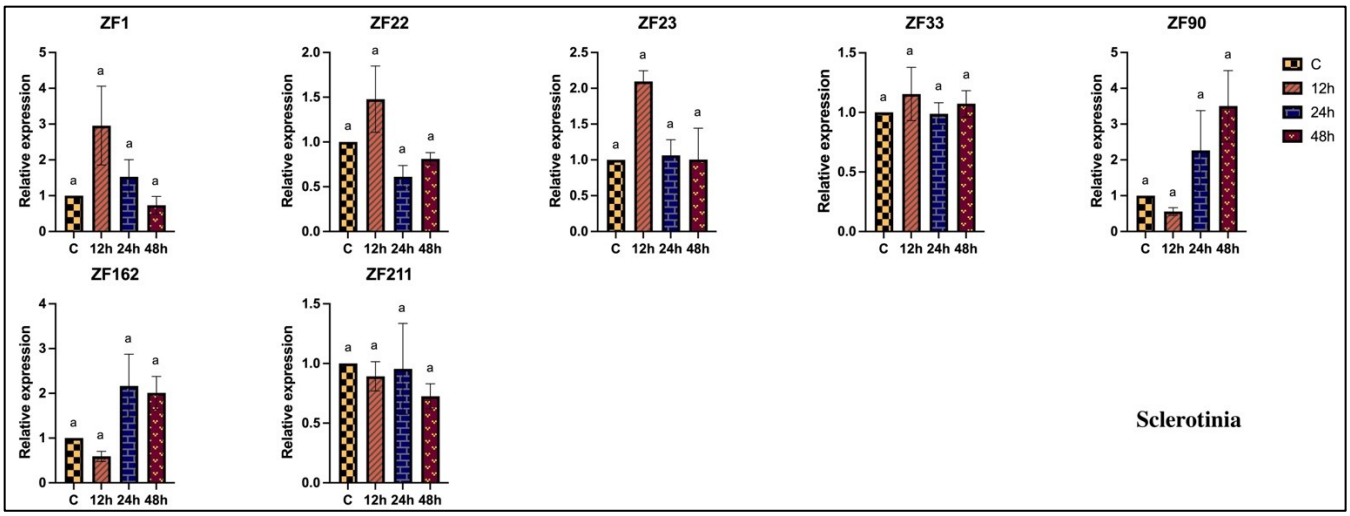

Name p-value Motif Locations
BnaZF250 1.79e-17
BnaZF22 5.60e-8
BnaZF173 3.13e-7
BnaZF133 7.17e-22
BnaZF214 1.29e-20
BnaZF211 2.84e-11
BnaZF18 8.93e-252
BnaZF169 1.01e-251
BnaZF108 3.37e-248
BnaZF76 8.94e-256
BnaZF236 4.46e-254
Motif Symbol Motif Consensus
1. KRFECSICGKSFPSGQALGGHMRSHR
5. CARVCITDENDNVVFHTVVKPSMVVTNYRYETTGIRPENLRDAMPLKHAQRKIQEFLCNGEPMWKIRPRGGKGRILVGHGLDHDL DRLQLEYPSSMMMDTAKYPPLMKTSKLSNSLKYLTQAYLGYDIHVGIQDPYEDCVATMRLYTRMRYQKHKIEAYPLAADAQNRSN QVAWHQNEVERMSPDEMLSISRSDYYCWCL
8. DTTEEEDLALCLMMLSRDS
9. WCGGCGKDFKSEESLKEHLKSKHH



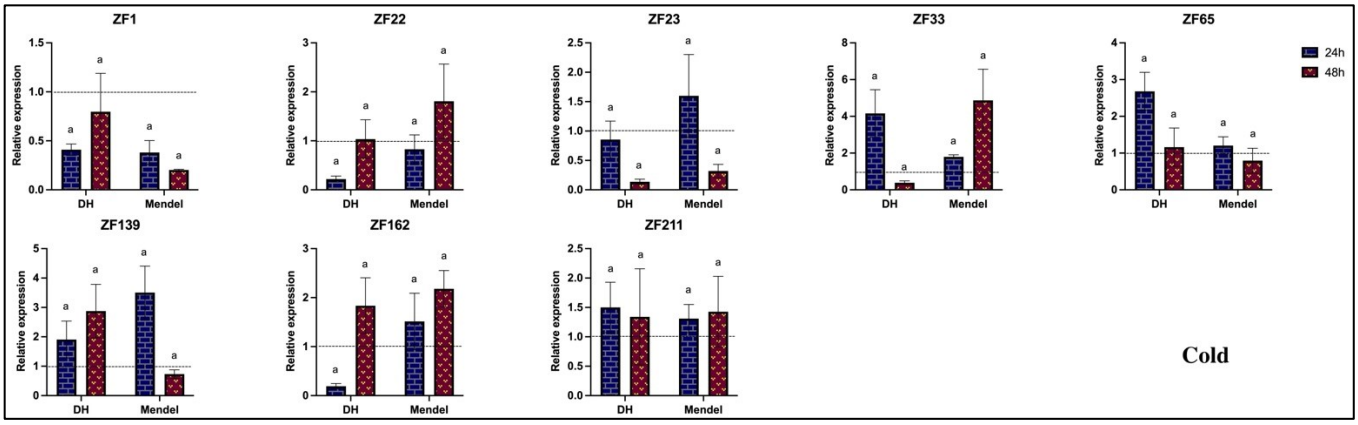
Supplementary Figure S6.4 Conserved motif analysis of the *Brassica napus* C2H2-ZF genes. A total of 10 motifs were identified. Motif analysis was performed using MEME5 suite. Part B of the figure represents the four motifs having Q, D, Z and M-type of C2H2-ZF domains.



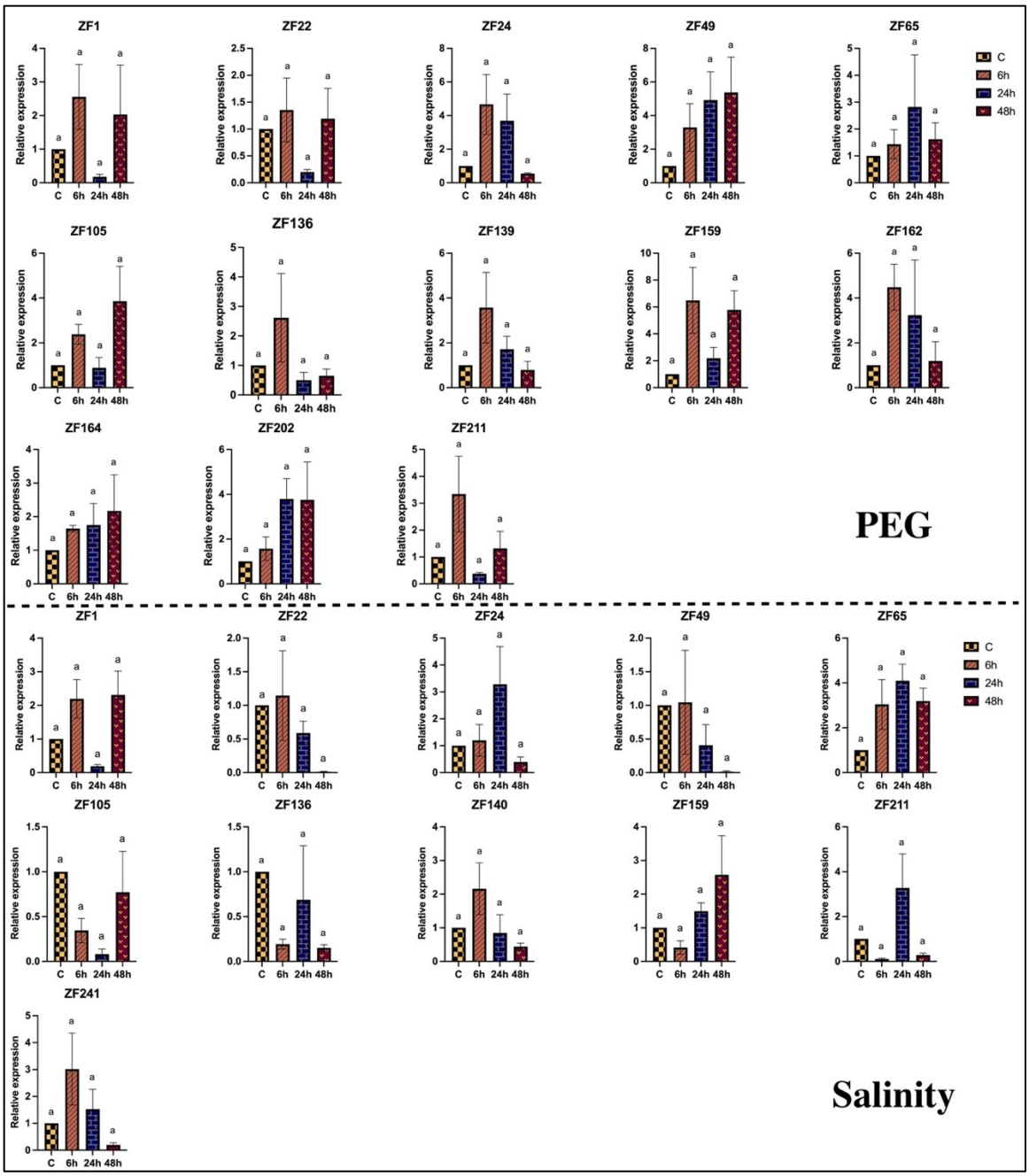
Supplementary Figure S6.5 Expression profiles of the *Brassica napus* C2H2-ZF genes which did not show significantly different expression upon *Plasmodiophora brassicae* infection in roots and leaves of clubroot resistant and susceptible lines (bulks). Error bars indicate means of three biological replicates ± standard errors (SEs). The same letter 'a' indicates the non-significant differences. The expression level was normalized to control of each time point and indicated by dotted line at expression level = 1.



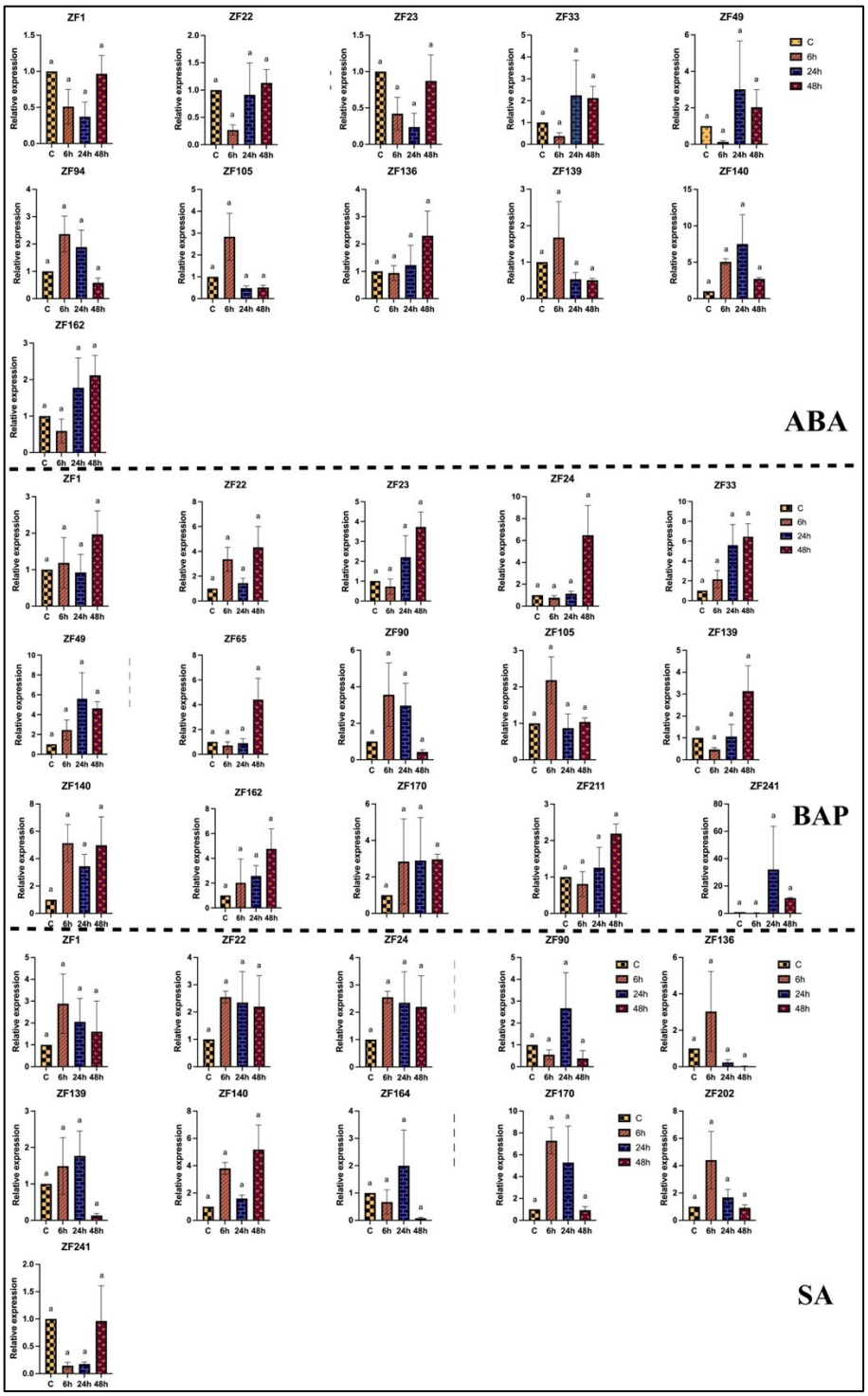
Supplementary Figure S6.6 Expression profiles of the *Brassica napus* C2H2-ZF genes which did not show significantly different expression upon *Sclerotinia sclerotium* infection. Error bars indicate means of three biological replicates ± standard errors (SEs). The same letter 'a' indicates the non-significant differences.



Supplementary Figure S6.7 Expression profiles of the *Brassica napus* C2H2-ZF genes which did not show significantly different expression upon cold stress. Error bars indicate means of three biological replicates ± standard errors (SEs). The same letter 'a' indicates the non-significant differences. The expression level was normalized to control of each time point and indicated by dotted line at expression level = 1.



Supplementary Figure S6.8 Expression profiles of the *Brassica napus* C2H2-ZF genes which did not show significantly different expression upon drought and salinity stress. Error bars indicate means of three biological replicates ± standard errors (SEs). The same letter 'a' indicates the non-significant differences.



Supplementary Figure S6.9 Expression profiles of the *Brassica napus* C2H2-ZF genes which did not show significantly different expression upon treatment with ABA (Abscisic acid), BAP (6-Benzylaminopurine) and SA (Salicylic acid). Error bars indicate means of three biological replicates ± standard errors (SEs). The same letter 'a' indicates the non-significant differences.

Please connect with the author for supplementary tables.

Supplementary Table S2.1 Performance of the BC_1F_8 lines (n = 450-452; see Figure S2) derived from (*B. napus* × *B. rapa*) × *B. napus* interspecific cross and their spring *B. napus* parent A04-73NA tested in nursery plots in 2020, and the BC_1F_9 lines (n = 112) and an open-pollinated cultivar UA CountyGold tested in replicated field trials in 2021 for different agronomic and seed quality traits.

Supplementary Table S2.2 Analysis of variance for agronomic and seed quality traits of the BC_1F_9 lines derived from (*B. napus* × *B. rapa*) × *B. napus* interspecific cross. Data from three trials conducted in 2021 were used in analysis.

Supplementary Table S3.1a, b List of SSR markers and allele-specific markers used for mapping of candidate regions conferring clubroot resistance.

Supplementary Table S3.2 Summary of whole-genome resequencing (WGRS) of the resistant (RP) and susceptible (SP) parental lines and the bulks of *B. napus* recombinant inbred lines derived from (*B. napus* ´ *B. rapa* var. *rapifera*) ´ *B. napus* interspecific cross for resistance to *Plasmodiophora brassicae* pathotype 3H and 3A.

Supplementary Table S3.3 Summary of parental polymorphism using SSR and allele-specific primers on 4 chromosomes for resistance to *Plasmodiophora brassicae* pathotype 3A.

Supplementary Table S3.4 List of putative *B. napus* genes to be affecting clubroot resistance.

Supplementary Table S4.1 Mapping statistics of the sequencing reads from *Brassica napus* near-isogenic line (NIL) libraries post inoculation with *Plasmodiophora brassicae* pathotype 3H.

Supplementary Table S4.2a The significantly enriched gene ontology (GO) terms of the differentially expressed genes (DEGs) obtained by comparing the inoculated clubroot resistant (CR) near-isogenic lines (NILs) with the control (RI vs. RC).

Supplementary Table S4.2b The significantly enriched gene ontology (GO) terms of the differentially expressed genes (DEGs) obtained by comparing the inoculated clubroot susceptible (CS) near-isogenic lines (NILs) with the control (SI vs. SC).

Supplementary Table S4.2c The significantly enriched gene ontology (GO) terms of the differentially expressed genes (DEGs) obtained by comparing the inoculated clubroot resistant (CR) near isogenic lines (NILs) with the inoculated clubroot susceptible (CS) NILs (RI vs. SI).

Supplementary Table S4.2d Common significantly enriched gene ontology (GO) terms of the differentially expressed genes (DEGs) obtained by comparing the inoculated clubroot resistant (CR) near-isogenic lines (NILs) and inoculated clubroot susceptible (CS) NILs (RI vs. SI) with their control. The GO terms marked with blue font are the most significant ones.

Supplementary Table S4.2e The significantly enriched gene ontology (GO) terms of the differentially expressed genes (DEGs) specific to the clubroot resistant (CR) near-isogenic lines (NILs).

Supplementary Table S4.2f The significantly enriched gene ontology (GO) terms of the differentially expressed genes (DEGs) specific to the clubroot susceptible (CS) near-isogenic lines (NILs).

Supplementary Table S4.3 List of common differentially expressed genes (DEGs) between the infected and control resistant and susceptible near-isogenic lines (NILs) (RI vs. RC and SI vs. SC) at 7- and 14-days after inoculation (dai).

Supplementary Table S4.4 List of common differentially expressed genes (DEGs) between infected resistant and susceptible near-isogenic lines (NILs) at 7- and 14-days after inoculation (dai).

Supplementary Table S4.5a List of differentially expressed genes (DEGs) involved in disease symptom (Sd) development and susceptibility (S) in the inoculated clubroot susceptible (CS) near-isogenic lines (NILs) of *Brassica napus.*

Supplementary Table S4.5b List of differentially expressed genes (DEGs) involved in pathogen-associated molecular pattern (PAMP)-triggered immunity (PTI) in the inoculated clubroot resistant near-isogenic lines (NILs) of *Brassica napus.*

Supplementary Table S4.5c List of differentially expressed genes (DEGs) involved in effector-triggered immunity (ETI) of the inoculated clubroot resistant near-isogenic lines (NILs) of *Brassica napus.*

Supplementary Table S4.6 List and description of the differentially expressed genes (DEGs) used for quantitative real-time (qRT) PCR analysis and their primer sequences.

Supplementary Table S4.7 List of differentially expressed genes (DEGs) in the inoculated clubroot resistant near-isogenic lines (NILs) of *Brassica napus.*

Supplementary Table S5.1a List of differentially accumulated proteins (DAPs) in roots of the Plasmodiophora brassicae-infected and control clubroot resistant near-isogenic Brassica napus lines (RI vs. RC) at 7-, 14- and 21 days post inoculation (dpi). The same proteins detectec at all three time points are listed here.

Supplementary Table S5.1b List of differentially accumulated proteins (DAPs) in roots of the Plasmodiophora brassicae-infected and control clubroot susceptible near-isogenic Brassica napus lines (SI vs. SC) at 7-, 14- and 21 days post inoculation (dpi). The same proteins detectec at all three time points are listed here.

Supplementary Table S5.1c List of differentially accumulated proteins (DAPs) in roots from *Plasmodiophora brassicae*-infected clubroot resistant CR) and susceptible (CS) (RI vs. SI) near-isogenic *Brassica napus* lines at 7-, 14- and 21 days post inoculation (dpi). The same proteins detected at all three time points are listed here.

Supplementary Table S5.2a Differentially accumulated proteins (DAPs) in roots from *Plasmodiophora brassicae*-inoculated and control clubroot resistant (RI vs RC) and clubroot susceptible (SI vs SC) near-isogenic *Brassica napus* lines at 7 dpi. The proteins showing a similar

pattern (increase or decrease) in both CR- and CS-NILs, and the proteins showing contrasting pattern (increase in one but decrease in the other) in the CR- and CS-NILs are presented.

Supplementary Table S5.2b Differentially accumulated proteins (DAPs) in roots from *Plasmodiophora brassicae*-inoculated and control clubroot resistant (RI vs RC) and clubroot susceptible (SI vs SC) near-isogenic *Brassica napus* lines at 14 dpi. The proteins showing a similar pattern (increase or decrease) in both CR- and CS-NILs, and the proteins showing contrasting pattern (increase in one but decrease in the other) in the CR- and CS-NILs are presented.

Supplementary Table S5.2c Differentially accumulated proteins (DAPs) in roots from *Plasmodiophora brassicae*-inoculated and control clubroot resistant (RI vs RC) and clubroot susceptible (SI vs SC) near-isogenic *Brassica napus* lines at 21 dpi. The proteins showing a similar pattern (increase or decrease) in both CR- and CS-NILs, and the proteins showing contrasting pattern (increase in one but decrease in the other) in the CR- and CS-NILs are presented.

Supplementary Table S5.3 Differentially accumulated proteins (DAPs) in roots of *Plasmodiophora brassicae*-inoculated clubroot resistant (CR) and susceptible (CS) near-isogenic *Brassica napus* lines (RI vs SI) at 7-, 14- and 21-days post inoculation (dpi). The proteins showing a similar pattern (increase or decrease) in both CR- and CS-NILs, and the proteins showing contrasting pattern (increase in one but decrease in the other) in the CR- and CS-NILs are presented.

Supplementary Table S5.4a The significantly enriched gene ontology (GO) terms of the differentially accumulated proteins (DAPs) in roots obtained by comparing the *Plasmodiophora brassicae*-inoculated clubroot resistant (CR) near-isogenic *Brassica napus* lines with the control (RI vs. RC). The GO terms marked with red font are the most significant ones.

Supplementary Table S5.4b The significantly enriched gene ontology (GO) terms of the differentially accumulated proteins (DAPs) in roots obtained by comparing the *Plasmodiophora brassicae*-inoculated clubroot susceptible (CS) near-isogenic *Brassica napus* lines with the control (SI vs. SC). The GO terms marked with red font are the most significant ones.

Supplementary Table S5.4c Significantly enriched gene ontology (GO) terms of the differentially accumulated proteins (DAPs) in roots obtained by comparing the inoculated clubroot resistant (CR) and inoculated susceptible (CS) of the near-isogenic *Brassica napus* lines (RI vs. SI). The GO terms marked with red font are the most significant ones.

Supplementary Table S5.5a The unique significantly enriched gene ontology (GO) terms of the differentially accumulated proteins (DAPs) in roots belonging to biological process (BP) and molecular function (MF) categories obtained by comparing the *Plasmodiophora brassicae*-inoculated clubroot resistant (CR) near-isogenic *Brassica napus* lines with the control (RI vs. RC). The GO terms marked with red font are discussed in the manuscript.

Supplementary Table S5.5b The unique significantly enriched gene ontology (GO) terms of the differentially accumulated proteins (DAPs) in roots belonging to biological process (BP) and molecular function (MF) categories obtained by comparing the *Plasmodiophora brassicae*-inoculated clubroot susceptible (CS) near-isogenic *Brassica napus* lines with the control (SI vs. SC). The GO terms marked with red font are discussed in the manuscript.

Supplementary Table S5.5c The unique significantly enriched gene ontology (GO) terms of the differentially accumulated proteins (DAPs) in roots belonging to biological process (BP) and molecular function (MF) categories obtained by comparing the *Plasmodiophora brassicae*-inoculated clubroot resistant (CR) near-isogenic *Brassica napus* lines with the *P. brassicae*-inoculated clubroot susceptible (CS) near-isogenic *B. napus* lines (RI vs. SI). The GO terms marked with red font are discussed in the manuscript.

Supplementary Table S5.6 The significantly enriched Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways of the differentially accumulated proteins (DAPs) in roots obtained by comparing the *Plasmodiophora brassicae*-inoculated clubroot resistant (CR) near-isogenic *Brassica napus* lines with the control (RI vs. RC), the *P. brassicae*-inoculated clubroot susceptible (CS) near-isogenic *B. napus* lines with the control (SI vs. SC) and the *P. brassicae*-inoculated clubroot resistant (CR) near-isogenic *B. napus* lines with the *P. brassicae*-inoculated clubroot susceptible (CS) near-isogenic *B. napus* lines (RI vs. SI). The KEGG pathways marked with red font are unique to the particular comparison.

Supplementary Table S5.7a The important and significantly ($q<0.05$) differentially accumulated proteins (DAPs) in roots to be involved in clubroot resistance in *Plasmodiophora brassicae*-inoculated clubroot resistant (CR) and *P. brassicae*-inoculated clubroot susceptible (CS) near-isogenic *Brassica napus* lines. The DAPs were identified the CR and CS lines while comparing with their respective controls (RI vs RC and SI vs SC).

Supplementary Table S5.7b The important and significantly ($q<0.05$) differentially accumulated proteins (DAPs) in roots to be involved in clubroot resistance in *Plasmodiophora brassicae*-inoculated clubroot resistant (CR) near-isogenic *Brassica napus* lines with respect to the *P. brassicae*-inoculated clubroot susceptible (CS) near-isogenic *B. napus* lines (RI vs SI).

Supplementary Table S6.1 The sequences of primers used in expression analysis of 20 *B. napus* C2H2-ZF genes.

Supplementary Table S6.2 List of 267 C2H2-ZF genes identified in *Brassica napus* and their relevant information.

Supplementary Table S6.3 Classification of the 267 *Brassica napus* C2H2-ZFPs based on organization of their C2H2 fingers.

Supplementary Table S6.4 Syntenic relationships between *Arabidopsis thaliana* and *Brassica napus* C2H2-ZF genes.

Supplementary Table S6.5 Ka, Ks and Ka/Ks ratios between *Arabidopsis thaliana* and *Brassica napus* orthologous C2H2-ZF genes.

Supplementary Table S6.6 Conserved motif sequences of the *Brassica napus* C2H2-ZFPs.

Supplementary Table S6.7 List of cis-elements identified in 2Kb upstream of *Brassica napus* C2H2-ZF genes.