Genetic and molecular analysis of clubroot resistance in canola introgressed from rutabaga cvs. Polycross and Brookfield

by

Zhengping Wang

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

Master of Science

in

Plant Science

Department of Agricultural, Food and Nutritional Science
University of Alberta

Abstract

Clubroot disease caused by *Plasmodiophora brassicae* is a serious threat to canola (*Brassica napus*) production. The evolution of new pathotypes has rendered available resistances ineffective and has necessitated the introgression of new resistance into canola and furthering our understanding of the genetic and molecular basis of the resistance. Clubroot resistance from rutabaga (*B. napus* var. *napobrassica*) cvs. Polycross and Brookfield has been introgressed into canola in a previous study. In the first study outlined in this thesis, I report the genetic and molecular basis of clubroot resistance in canola, introgressed from a rutabaga cv. Polycross, by using a doubled haploid (DH) mapping population. Whole genome resequencing-based bulked segregant analysis followed by genetic mapping was carried out to identify the genomic regions contributing to this resistance, and expression analysis of the genes from the quantitative trait loci (QTL) regions was carried out to understand the molecular basis of this resistance. Following this approach, two major QTL located at 14.41-15.44 Mb of A03 and at 9.96-11.09 Mb of A08 chromosomes and their interaction was observed to confer resistance to pathotypes 3H, 3A and 3D. Analysis of the genes from the two QTL regions suggested that decreased expression of sugar transporter genes may play an important role in resistance conferred by the A03 QTL, while increased expression of the TNL genes could be the major determinant of the resistance conferred by the A08 QTL. SNP allele-specific PCR-based markers, which could be detected by agarose gel electrophoresis, were also developed from the two QTL regions for use in breeding including pyramiding of multiple clubroot resistance genes. In the second study described in this thesis, I investigated four putative clubroot resistance genes and long noncoding RNAs (lncRNAs) from primary and secondary metabolic pathways through overexpression in *Arabidopsis thaliana*; these genes and lncRNAs were identified through transcriptome analysis of *B. napus* lines carrying clubroot resistance of the rutabaga cv. Brookfield. However, none of the homozygous transgenic *A. thaliana* lines carrying the above-mentioned genes or lncRNAs showed resistance to clubroot disease suggesting that these genes and lncRNAs may not be directly involved in clubroot resistance. Thus, the knowledge and materials generated from this thesis research can be used in breeding canola for durable resistance clubroot disease.

Preface

This thesis is the original submitted by Zhengping Wang for the degree of Master of Science. Zhengping Wang carried out all the experiments, collected and analyzed data for this thesis; with guidance, comments and suggestions from his supervisor Dr. Habibur Rahman, co-supervisor Dr. Nat Kav and mentor Dr. Swati Megha. Final version of the thesis was prepared by incorporating additional suggestions from the examination committee members Drs. Nat Kav and Glen Uhrig.

A version of chapter 2 in this dissertation has been submitted for publication in The Plant Genome journal. For this chapter, Zhengping Wang conducted experiments in growth chamber for phenotyping, DNA and RNA extraction, designing primers, genotyping, expression analysis, data collection, bioinformatic analysis and figure visualization. The population used in this experiment were developed by Dr. Berisso Kebede in the Canola program of the University of Alberta. Dr. Swati Megha helped in expression analysis and experiment design. This project is supervised and administered by Dr. Habibur Rahman.

For the chapter 3 in this dissertation, Zhengping Wang conducted all experiments, data collection and bioinformatic analysis under guidance of Dr. Habibur Rahman, Dr. Nat Kav and Dr. Swati Megha. The reference of putative lncRNAs and genes are provided by Dr. Aarohi Summanwar.

In addition, Dr. Habibur Rahman also provided valuable feedback on the chapter 1 and 4 in this dissertation.

Dedication

To

My younger sister Jieyu Wang

Acknowledgements

I'm extremely grateful to my supervisor Dr. Habibur Rahman for his guidance, patience, concern and financial support. I would also like to extend my deepest gratitude to my supervisory committee member Dr. Nat Kav and Dr. Glen Uhrig for their valuable suggestions and critical evaluation of my research. I am also grateful to Dr. Swati Megha for her guidance. It is treasure of my life to work with them.

I would like to thank to Dr. Muhammad Hasan, Dr. Berisso Kebede and Dr. Aarohi Summanwar for their previous study. I'd like to acknowledge the help from Dr. Enrico Scarpella, Dr. Urmila Basu, An Vo and Kelley Dunfield. I also want to thank other members in our Canola program.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AS	Allele-specific
BF	Rutabaga-Brookfield
bp	Base pair
BRAD	Brassicaceae database
BSA	Bulked-segregant analysis
BSR-Seq.	Bulked segregant RNA-seq
CCD	Canadian clubroot differential Set
CI	Confidence interval
cM	Centimorgan
DAI	Days after inoculation
DGE	Digital gene expression
DH	Doubled haploid
DSI	Disease severity index
ECD	European clubroot differential
ETI	Effector triggered immunity
gDNA	Genomic DNA
GSL	Glucosinolate
GWAS	Genome-wide association study
ICIM	Inclusive composite interval mapping
lncRNA	Long noncoding RNA
Mb	Mega base pair
MCS	Multiple cloning sites
mRNA	Messenger RNA
NGS	Next-generation sequencing
ncRNA	Noncoding RNA
PAMPs	Pathogen-associated-molecular-patterns
PC	Rutabaga-Polycross
PCA	Principal component analysis

PCR	Polymerase chain reaction
PLncDB	Plant long noncoding RNA database
PNRD	Plant ncRNA database
PRA	Resistant parent assembly
PRRs	Pattern recognition receptors
PTI	Pattern triggered immunity
PVE	Phenotypic variation explained
qRT-PCR	Quantitative reverse transcription PCR
QTL	Quantitative trait locus
RB	Resistant bulk
RILs	Recombinant inbred lines
RP	Resistant parental bulk
RPA	Resistant parent assembly
R protein	Resistance protein
SA	Salicylic acid
SB	Susceptible bulk
SBL	Sequencing by ligation
SBS	Sequencing by synthesis
SB-3D	Susceptible bulk-pathotype3D
SB-3H	Susceptible bulk-pathotype3H
SCD	Sinitic clubroot differential set
sncRNA	Small noncoding RNA
SNP	Single nucleotide polymorphism
SP	Susceptible parental bulk
SSRs	Simple-sequence repeats
TAIR	The Arabidopsis information resource
tPCK	Translocated Proto-Calepine Karyotype
vlincRNA	Very long intergenic RNA
WGRS	Whole-genome resequencing
$\chi 2$	Chi-square

Chapter 1. Introduction

1.1 *Brassica napus*

1.1.1 Introduction of *Brassica* genus

Brassicaceae (Cruciferae) is one of the most important families of flowering plants including at least 372 genera such as *Brassica* and *Arabidopsis*. Among them, the genus *Brassica* comprises a large and diverse group of vegetables, such as broccoli, brussels sprouts, cauliflower and cabbage belonging to the species *B. oleracea* (CC, 2*n* = 18), and Chinese cabbage, bok choy, pak choy, Japanese mustard spinach (Komatsuna) and mizuna greens belonging to the species *B. rapa* (AA, 2*n* = 20), as well as rutabaga (*Brassicas. napus* var. *napobrassica*) which is used for human consumption and as fodder. This genus also includes the oilseed crop species *B. napus* (AACC, 2*n* = 38), *B. juncea* (AABB, 2*n* = 37) and *B. rapa*; these three species collectively supply about 12% of the total edible oil in the world, which is the second-largest oilseed crop in the world after soybean (USDA 2021). The popularity of the *Brassica* vegetables and oils has increased worldwide in the last decades making this genus as one of the most economically important one of the plant kingdom Plantae.

From a nutritional perspective, *Brassica* vegetables are one of the best sources of vitamin C, fibre, folate, calcium, and certain phytochemicals (Fahey, 2015); ancient Romans and Greeks have also recorded the dietary benefits of cabbages and cauliflowers (Nijhuis, 1995). Among the different chemical components of *Brassica* vegetables, glucosinolates are known for their fungicidal, bactericidal, nematocidal and allelopathic properties (for review, see Vig et al., 2009). Many *Brassica* vegetables have been found to be rich in glucoraphanin and sulforaphane, which possess cancer-chemoprotector and antioxidant properties (Gu et al., 2012). *Brassica* vegetables have been modified for specific nutrition, such as orange cauliflower as a source of β-carotene (Jeffery et al., 2003; Lu et al., 2006). *Brassica* seed oil, especially canola oil is considered as one of the most economical and healthiest edible oil. The genus *Brassica* contains species which evolved from not only neotetraploidization but also mesohexaploidization; therefore, it is suitable for studying genome evolution as well as the structure, function and regulation of complex genomes (for review, see Friedt et al., 2018). The relationships between the species of the genus *Brassica* and the evolution will be discussed in the following sections.

1.1.2 Rapeseed/canola

Rapeseed is the traditional name for the *B. napus* oilseed crops; this includes the cultivars whose oil is used for industrial or edible purposes. Among them, canola quality type is grown most extensively, and its oil is used as edible oil around the world. Canola seed contains about 38 to 44% oil (for review, see Przybylski & Mag, 2011) which is almost free from erucic fatty acid, and a gram of its seed meal contain <30 μmol glucosinolates (GSLs) (for review, see Gupta, 2016). The name "canola" was registered by the Western Canadian Oilseed Crushers in 1978 for the cultivars with low erucic acid in oil and low GSLs in seed meal. Erucic acid is a long carbon-chain mono-unsaturated fatty acid. Consumption of oil containing this fatty acid can result fatty deposits in the heart and skeletal muscles and cause cardiac damage (Charlton et al., 1975; Christophersen & Bremer, 1972). GSLs are Sulfur-rich compound and are secondary plant metabolites produced by most *Brassica* species. The use of traditional rapeseed meal, which contains about 38% protein, as an animal feed has been limited due to the presence high content of GSLs. The degradation products of GSLs are toxic which can interfere with iodine uptake and reduce weight gain in monogastric animals (for review, see Przybylski et al., 2005). To reduce erucic acid and GSLs content, breeders introduced the low erucic acid genes from German spring forage rape cv. Liho (for review, see Przybylski & Mag, 2011) and the low GSL genes from Polish forage cv. Bronowski (Olsen & Sorensen, 1980), and the first low erucic acid and low GSLs or ‘double low’ or canola quality spring type cv. Tower was released in 1974 (for review, see Friedt et al., 2018). Eventually, canola became a standard of *Brassica* oilseed cultivars. Today, canola has become an important crop in the farms due to its economic and ecological value. Canada was the largest exporter of canola during the period of 2017 to 2020. In 2019, Canada produced 20.3 million metric tons of canola seed, and this accounted about one-third of the total world production (USDA 2020).

The fatty acid composition of canola oil is suitable for use as edible oil. Based on Europe and North America's nutritional recommendations, saturated fat intake should be reduced to less than 10% of the total energy in the diet (for review, see Brouwer, 2020). The total saturated fatty acids in canola oil are about 7%; the contents of omega-6 and omega-3 fatty acids and their ratio in canola oil is beneficial for human nutrition as they reduce plasma and LDL cholesterol levels in blood (for review, see Agnihotri et al., 2007 and Przybylski et al., 2005).

In many European and North American countries including Canada, wheat, barley, maize

and canola are mostly cultivated on farms. Among them, canola is a non-cereal dicot crop that can reduce the negative effects of monoculture of using monocot crops. The rotation of wheat and canola or barley and canola can increase canola yield by 0.20-0.36 t ha^{-1} (Harker et al., 2015). The use of this broad-leaf crop in rotation can also improve soil-health for the following cereal crop. Canola straw has been found to be suitable for the production of biofuels (Herrmann et al., 2013). The use of biofuels can reduce at least 35% of greenhouse gas emissions compared to the use of fossil fuels. All these economical, nutritional and ecological values make this crop suitable for growing on a farm.

1.1.3 *Brassica* genomes, their evolution and relationship

The genus *Brassica* includes 15 species of which six are well-known. This includes three monogenomic species and three amphidiploids. The most well-known model for the relationship between these species has been described by Nagaharu U (1935) as the "Triangle of U". He designated the three genomes of the diploid species as "A" for *B. rapa* (2*n* = 20, AA), "B" for *B. nigra* (2*n* = 16, BB) and "C" for *B. oleracea* (2*n* = 18, CC). According to U (1935), the three amphidiploids resulted from the three monogenomic species through interspecific hybridization. According to this model, *B. juncea* (2*n* = 36, AABB) resulted from crossing between *B. rapa* and *B. nigra*; *B. napus* (2*n* = 38, AACC) resulted from *B. rapa* and *B. oleracea*, and *B. carinata* (2*n* = 34, BBCC) resulted from *B. nigra* and *B. oleracea*. Later, cytogenetic analysis and genome sequencing also confirmed the relationship between these species (Mason et al., 2010; Chalhoub et al. 2014). Furthermore, *B. napus* (Rahman, 2001, 2005), *B. juncea* (Axelsson et al., 2000; Hasan & Rahman, 2018) and *B. carinata* (Jourdan & Salazar, 1993) have been successfully resynthesized through crossing their diploid parental species.

Our understanding of the relationships between the *Brassica* genomes and their evolution has been extended from the knowledge of the sequences of *B. rapa* (Wang et al., 2011), *B. napus* and other *Brassica* crop genomes (Yang et al., 2016a, Liu et al., 2014; Wang et al., 2011, Chalhoub et al., 2014; Rousseau-Gueutin et al., 2021) as well as sequence of the *Arabidopsis thaliana* (2*n* = 10) genome (Bevan et al., 2001). Comparative genomic analysis of *B. rapa* and *A. thaliana* showed that both *Arabidopsis* and the diploid *Brassica* species evolved from a diploid ancestor 'translocated Proto-Calepine Karyotype' (tPCK) with *n* = 7 (Figure 1.1). The evolution experienced a whole genome hexaplication event which occurred about 23.3 million years ago

before the divergence of *Brassica* and *Arabidopsis* ancestors (Cheng et al., 2013). This genome evolution apparently occurred in two steps. In the first step, two tPCK genomes merged together resulting an allotetraploid ($2n = 4x = 28$), and in the second step, this tetraploid merged with a third tPCK genome and resulted the ancestral hexaploid (Cheng et al., 2017; for review, see Cheng et al., 2015a; Cheng et al., 2015b). Eventually, this hexaploid became the last common ancestor of the genus *Brassica*. After several rounds of genome reorganization including chromosome translocation, fission, fusion and inter/intra-chromosomal recombination, the *B. nigra* genomes evolved about 7.5 million years ago (Mya) (Mun et al., 2009; Navabi et al., 2013), while *B. oleracea* evolved about 4.6 Mya and diverged from *B.rapa* (Liu et al., 2014). However, a recent study shows the divergence of *B. nigra* could be traced back to 11.5 Mya; *B. oleracea* and *B. rapa* diverged from each other on 6.8 Mya (Perumal et al., 2020). It indicated genus *Brassica* may have a longer evolutionary history than reflected by our current understanding. Some other theories could also explain this evolution process; however, it is commonly accepted that the *Brassica* species evolved from a hexaploid ancestor (Yang et al., 2016b). Evolution of the allotetraploid species *B. napus*, *B. juncea* and *B. carinata* through interspecific hybridization between the three diploid species occurred about a million years ago (Chalhoub et al., 2014; Waminal et al., 2016). A diagram of this evolutionary pathway is shown as figure 1.1.

Interestingly, with the sequencing of additional members of the *Brassicaceae* family, evidence supports the evolution of *R. sativus*, *S. alba*, *C. amplexicaulis*, *S. parvula* and *T. salsuginea* from the tPCK hexaploid ancestor (Cheng et al., 2013; Wang & Kole, 2015). For example, three copies of the two tPCK-specific block associations V/K/L/Wa/Q/X and O/P/W/R were observed on linkage groups of *R. sativus* (for review, see Cheng et al., 2015a). Phylogenetic study based on chloroplast genomes also showed that *R. sativus* is more closely related to *B. rapa* and *B. oleracea* than *B. nigra* (Seol et al., 2017).

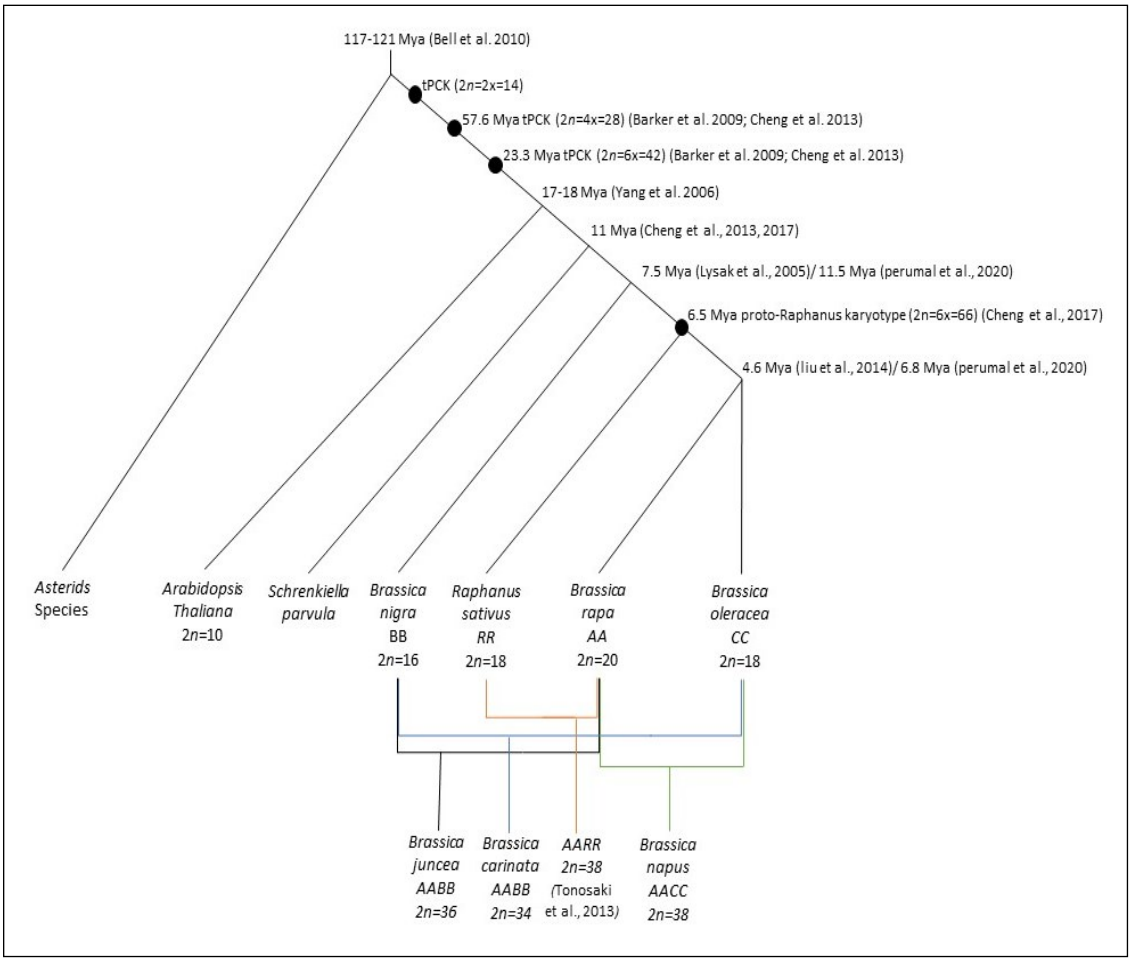


Figure 1.1 Diagram of the evolutionary pathway of the *Brassica* species from the common ancestor (tPCK). The black spots mark the potential genome duplication event. Divergence time are suggested by different authors are also provided. Mya: million years ago.

1.1.4 Relationship of *Brassica* with *Arabidopsis thaliana*

Arabidopsis thaliana is a popular model for studying different plant traits due to its short life cycle, simple genome, amenability to genetic transformation, and availability of its genome sequence (Kaul et al., 2014). This species belongs to the *Brassicaceae* family, and the last common ancestor of *Brassica* and *A. thaliana* existed about 14.5-20.4 million years ago (Hong et al., 2006; Yang et al., 1999, for review, see Yang et al., 2016b). Molecular marker-based phylogenetic and comparative genome analysis showed that *A. thaliana* genome is similar to tPCK, the hypothetical diploid progenitor of *Brassica,* and the genomic blocks of *Arabidopsis* covered about 90% of the *Brassica* genome (Lagercrantz & Lydiate, 1996; Li et al., 2003). Comparative analysis of *B. napus* and *A. thaliana* mitochondrial genomes showed more than

99% similarity between these two species for protein-coding regions and 83% similarity for RNA-editing regions, which further supports the close evolutionary relationship between *Brassica* and *Arabidopsis* (Handa, 2003). *A. thaliana* has also been used as a model host plant to investigate a several canola diseases (Mithen & Magrath, 1992; Dickman & Mitra, 1992; Alix et al., 2007).

1.2 *Plasmodiophora brassicae*

1.2.1 life cycle of *P. brassicae*

Plasmodiophora brassicae belongs to the Kingdom Protista and phylum *Plasmodiophoromycota*, classified as a protist (Alexopoulos et al., 1996). As an obligate parasite, *P. brassicae* cannot complete its life cycle without host plants which are not limited to *Brassica* species (Ludwig-Müller et al., 1999). The life cycle of *P. brassicae* can be divided into three stages, dormant stage, primary infection stage and secondary infection stage (Kageyama & Asano, 2009). *P. brassicae* survives in soil as resting spores (Ayers, 1944). The spores are capable of surviving in soil up to 20 years (Wallenhammar, 1996), and can escape traditional pathogen management practices such as crop rotations and bait crops. Control of this disease using chemicals, such as modification of soil pH, and addition of flusulfamide, chlorothalonil or quintozene in soil was also not effective for management of this disease in canola (Hwang et al., 2014; Kowata-Dresch & May-De Mio, 2012; Takahashi, 1994; Wallenhammar, 1996). A two-year crop rotation was found not to be effective to decrease the level of the pathogen to an acceptable level for cultivating canola (Peng et al., 2014b).

Germination of the resting spores produces zoospores which can move using their flagella and infect root hair or epidermal root cells of either host or non-host plant and become primary plasmodia, which is considered to be the primary infection stage (Ayers, 1944). The primary plasmodia are cleaved into multiple zoosporangia which produces secondary zoospores (Kageyama & Asano, 2009). In case of life cycle (Figure 1.2) of *P. brassicae* where non-host plants are included, its life cycle would be interrupted at this stage; however, using a bait crop the inoculum load cannot be reduced significantly (Ahmed et al., 2011; Friberg et al., 2005). The secondary zoospores penetrate the root cortical tissue and develop into secondary plasmodia and this stage is termed as the secondary infection stage. The secondary plasmodia divide and differentiate into resting spores, and clubbed root is formed due to this secondary infection

(Ludwig-Müller et al., 1999). The majority of the susceptible cultivars show the characteristic clubs on the roots (Kageyama & Asano, 2009; Naiki et al., 1984). These clubs are the feeding sites which provide sugar and other nutrients to *P. brassicae* for its growth (Evans & Scholes, 1995), and help the pathogen to overcome the plant immunity (for review, see Malinowski et al., 2019). The resting spores are released to the soil by disintegrating the roots and its next life cycle begins (Kageyama & Asano, 2009; for review, see Ikegami et al., 1982 and Malinowski et al., 2019).

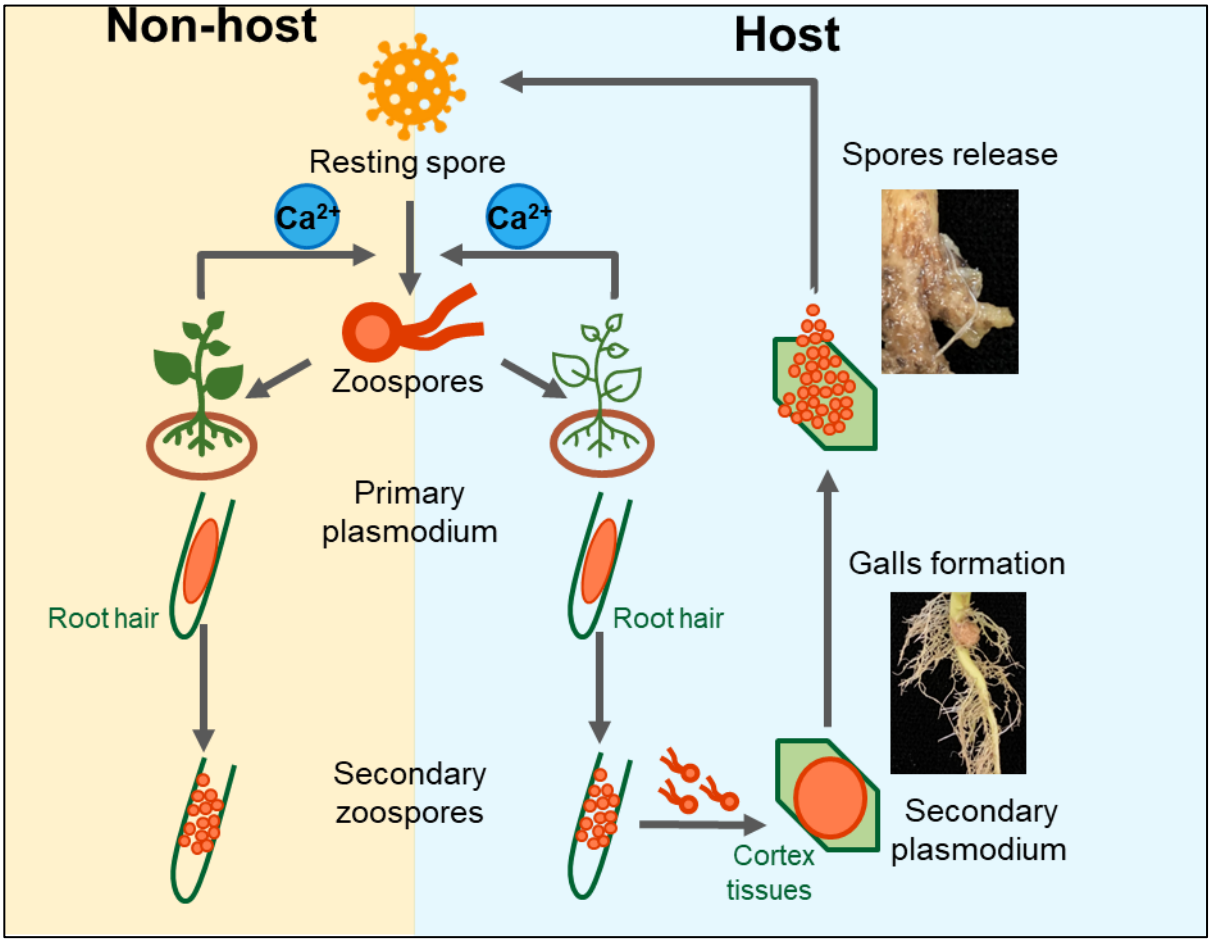


Figure 1.2 The life cycle of *Plasmodiophora brassicae* in host and non-host plants.

1.2.2 Clubroot disease impact in canola

The typical symptom of *P. brassicae* infection is the formation of characteristic galls or "clubs" on the roots of susceptible hosts. These galls can interfere with uptake of water and nutrients and disintegrate the roots and kill the plant. About 10-15% yield loss occurs every year in *Brassica*

crops worldwide (Dixon, 2009). Among the field crops, *B. napus* or canola, has been seriously threatened by this disease, especially in Canada, Europe and China. A high incidence of clubroot has been reported in Germany, United Kingdom, Sweden and Poland (for review, see Botero et al., 2019 and Diederichsen et al., 2009; Wallenhammar, 1996). For example, in Poland, a total of 431 fields during the period of 2013 to 2019 were investigated for the presence of *P. brassicae*, and 61.9% (267/431) of the fields tested positive for this pathogen (Czubatka-Bieńkowska et al., 2020). Clubroot has been found in 28 provinces in China where nine provinces were seriously affected by this disease (Chai et al., 2014). The yield loss of canola due to this disease can be about 10-30% (Tewari et al., 2005; Wang et al., 2008), however, this loss can be up to about 56% in case of heavy infestation (Ren et al., 2012). In Canada, clubroot disease in canola fields was first detected in 2003 (Tewari et al., 2005) since then, it has been spreading dramatically in Alberta, Saskatchewan and Manitoba (Dokken-Bouchard et al., 2010; Manitoba Agriculture, 2013; for review, see Rempel et al., 2014). The number of *P. brassicae* infested fields in Alberta has increased from 12 to 3353 during the period of 2003 to 2019; currently, it is the most severely infested province (Strelkov et al., 2020). The spread of this disease and its impact on canola have been annually monitored in Canada, especially in Alberta and Manitoba, since 2005 and it is reported in the "Canadian Plant Disease Survey", where hundreds of new fields were found to be infected every year (https://phytopath.ca/publication/cpds/).

1.2.3 Pathotypes of *P. brassicae*

Due to long longevity of the resting spores and change in population structure of *P. brassicae* in soil, the development of clubroot resistant cultivars has been considered to be the most economic among measures to combat this disease in canola (for review, see Rahman et al., 2014). Like many other plant pathogens, mutation and other genetic changes such as crossover in meiosis also occur in *P. brassicae* and this results in the evolution of new pathotypes or races which can overcome the plant immunity. As the pathogen races vary in their virulence, differentiating the races or pathotypes for their pathogenic diversity is important while developing clubroot resistant canola cultivars. Initially, pathotype designation was carried out using field pathogen populations and a set of host genotypes or host differential set (for review, see Crute et al., 1980) but later, single-spore-derived isolates were used (Jones et al., 1982; Xue et al., 2008). The European Clubroot Differential (ECD) set, and Williams system were used by different researchers for

several years for pathotype designation; however, none of them were capable of distinguishing all *P. brassicae* strains or pathotypes (Buczacki et al., 1975; Kuginuki et al., 1999; Somé et al., 1996; Strelkov et al., 2007; Williams, 1966). Therefore, a few additional host differential sets were developed by adding additional hosts to the ECD and Williams sets, and they include Somé's set which is mainly used in European countries, Canadian Clubroot Differential set (CCD) for Canada, and Sinitic Clubroot Differential set (SCD) for China (Pang et al., 2020; Somé et al., 1996; Strelkov et al., 2018).

By using the CCD set (Table 1.1), which was developed in 2017, 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, 8P were identified in soil samples collected in 2014, 2015 and 2016 in Canada (Strelkov et al., 2016, 2018). Later, nine new pathotypes, *viz.* 2C, 6D, 8D, 9A, 9B, 9C, 11A, 13A, 13B were identified in Western Canada (Hollman et al., 2021), five pathotypes, *viz*. 5A, 6C, 8A, 8B, 8C were identified in the Peace Country Region (Strelkov et al., 2020), and five pathotypes, *viz*. 2A, 4A, 6A, 6B, 7A were identified within the single-spore isolates (Askarian et al., 2021). Interestingly, these 19 (9+5+5) novel pathotypes were isolated from soil samples collected between 2017 and 2018. Based on this, it can be projected that additional new pathotypes will evolve in the coming years. Precise designation of a pathotype using clubroot differential set is time-consuming and also may not be highly inefficient. Therefore, a new pathotyping platform based on molecular markers need to be developed. Indeed, a probe-based qPCR assay has been developed for the pathotype 5X; the development of similar markers for other pathotypes will benefit the *Brassica* breeding and research community (for review, see: Tso et al., 2021; Zhou et al., 2018).

Table 1.1 A summary of 36 Plasmodiophora brassicae pathotypes have so far been identified in Canada based on Canadian Clubroot Differential (CCD) set. This table was constructed based on Askarian et al. (2021), Hollman et al. (2021) and Strelkov et al. (2016, 2018, 2020). A plus (+) sign denotes a susceptible host reaction, while a minus (−) sign denotes a resistant reaction. European Clubroot Differential (ECD) 02 = *Brassica rapa* ssp. *rapifera* line AAbbCC; ECD 05 = *B. rapa* var. *pekinensis* cv. 'Granaat'; ECD 06 = *B. napus* cv. 'Nevin'; ECD 08 = *B. napus* 'Giant Rape' selection; ECD 09 = *B. napus* New Zealand resistant rape; ECD 10 = *B. napus* var. *napobrassica* cv. 'Wilhemsburger'; ECD 11 = *Brassica oleracea* var. *capitata* cv. 'Badger Shipper'; ECD 13 = *B. oleracea* var. *capitata* cv. 'Jersey Queen'; Brutor = *B. napus* cv. 'Brutor'; Laurentian = *B. napus* var. n*apobrassica* cv. 'Laurentian'; Westar = *B. napus* cv. 'Westar'; and 45H29 = *B. napus* cv. '45H29'

	2A	**2B**	**2C**	**2F**	**3A**	**3D**	**3H**	**3O**	**4A**	**5A**	**5C**	**5G**	**5I**	**5K**	**5L**	**5X**	**6A**	**6B**	**6C**	**6D**	**6M**
ECD02	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ECD05	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
ECD06	+	+	-	+	+	+	+	-	+	-	+	-	+	-	-	-	-	-	-	-	+
ECD08	+	+	-	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	-	+
ECD09	+	+	-	+	+	+	+	+	+	+	+	-	+	-	-	-	-	-	-	-	+
ECD10	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-
ECD11	+	+	+	+	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-
ECD13	+	+	+	+	+	+	+	-	+	-	-	-	-	-	-	-	+	+	+	+	+
Brutor	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+
Laurentian	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-
Mendel	-	+	-	-	+	-	-	-	+	-	-	-	-	-	-	+	-	+	-	-	-
Westar	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+
45H29	+	+	-	-	+	+	-	+	+	-	+	+	-	+	-	+	+	+	-	-	-

	7A	8A	8B	8C	8D	8E	8J	8N	8P	9A	9B	9C	11A	13A	13B
ECD02	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ECD05	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
ECD06	+	-	+	-	-	+	-	+	+	+	-	+	+	-	-
ECD08	-	+	+	+	+	+	+	+	+	+	+	+	+	+	-
ECD09	-	-	-	+	+	+	-	+	+	-	-	-	+	-	-
ECD10	-	-	-	-	-	-	-	-	-	+	+	+	+	-	-
ECD11	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ECD13	+	-	-	-	-	-	-	-	-	-	-	-	+	+	+
Brutor	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Laurentian	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Mendel	-	-	-	-	-	-	-	-	+	+	-	-	+	-	-
Westar	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+
45H29	-	-	-	+	-	+	+	-	+	+	-	+	+	-	-

1.3 Clubroot resistance

1.3.1 Sources of clubroot resistance in *Brassica*

The devastating effect of clubroot disease and the difficulty of the management of this disease, as mentioned earlier, underlines the need for developing clubroot resistant cultivars. Screening of a large number of *Brassica* accessions showed that *B. rapa* including the European turnips (*B. rapa var. rapifera*) and wild turnip (*B. rapa var. oleifera*) carry resistance to this disease (Hasan et al., 2012; Peng et al., 2014a; Ramzi et al., 2018; Walker, 1939), and clubroot resistance in the A genome often controlled by single Mendelian gene (e.g. Chu et al., 2014; Kato et al., 2013; for review, see Hasan et al. 2021). Clubroot resistance has also been found in *B. oleracea,* such as kale (*B. oleracea* var. *acephala*), cabbage (*B. oleracea* var. *capitata*), cauliflower (*B. oleracea* var. *botrytis*) and broccoli (*B. oleracea* var. *italica*) (Crisp et al., 1989; Hasan et al., 2012; Peng et al., 2014a); however, resistance in the C genome is often a quantitative trait and controlled by multiple loci (e.g. Nagaoka et al., 2010; Rocherieux et al., 2004; for review, see Hasan et al. 2021a). Clubroot resistance can also be found in the B genome of black mustard (*B. nigra*) (Hasan et al., 2012; Peng et al., 2014a) however, introgression of this resistance into the AC genome of *B. napus* might be challenging as compared to introgression of resistance from *B. rapa* and *B. oleracea* (Hasan et al., 2021b) Several clubroot resistant *B. rapa* germplasm have been used in the breeding of Chinese cabbage, oilseed *B. napus* and rutabaga (Bradshaw et al., 1997; Frauen, 1999; Hirai et al., 2004).

Beside the diploid *Brassica* species, clubroot resistance has also been found in amphidiploid species such as *Brassica juncea* (Peng et al., 2014a), and in several rutabaga/swede (*Brassica napus var. napobrassica*) accessions (Fredua-Agyeman et al., 2019; for review see Crute et al., 1980; Hasan et al., 2021a; Piao et al., 2009 and Spaner, 2002). Genetic study of clubroot resistance in several accessions, such as ‘New Zealand’, ‘Giant rape’, ‘Wilhelmsburger’, ‘York’ and ‘Brookfield’, showed that major dominant gene often confers clubroot resistance in rutabaga (Ayers & Lelacheur, 1972; Johnston, 1970; Landry et al., 1992). Several of the resistance in rutabaga might have been derived from turnip (*B. rapa*) (Bradshaw et al., 1997). Nevertheless, of the different sources of clubroot resistance available in *Brassica* species, clubroot resistance from rutabaga can be introduced into oilseed *B. napus* relatively easily (for review, see Rahman et al. 2014; Hasan and Rahman 2016).

1.3.2 The molecular basis of clubroot resistance

Plants rely on innate immunity to defend against infection by pathogens, such as *P. brassicae*. Once pathogen penetrates the mechanical barrier of the roots, it is detected by the cells pattern recognition receptors (PRRs), which are located on the surface of the plasma membrane and recognizes the conserved pathogen-associated-molecular-patterns (PAMPs) thereby activating pattern-triggered immunity (PTI) (for review, see De-Lorenzo et al., 2011 and Jones & Dangl, 2006). PAMPs, such as flagellin and peptidoglycans, are highly conserved in pathogens (Felix et al., 1999; Gust et al., 2007) and, therefore, PTI can provide robust, long-term resistance. The contribution of PTI to clubroot resistance increases as the evolutionary distance increase between host and non-host plant (Schulze-Lefert & Panstruga, 2011). It implies that PTI provides a consistent and complete resistance in the non-host plant but not in host plants. A second level immunity, known as effector triggered immunity (ETI), is needed in the case of host plants. This immunity is activated by disease resistance (R) proteins that recognize the pathogens' effector molecules when PTI is suppressed (for review, see Jones & Dangl, 2006). Most of the R proteins are characterized by a nucleotide-binding site (NBS), a leucine-rich repeats (LRRs) domain on C-terminal, and a TIR-domain or a coiled-coil (CC) structure on N-terminal, which are abbreviated as TIR-NBS-LRR (TNL) or CC-NBS-LRR (CNL) proteins (for review, see Dangl & Jones, 2001). According to Thordal-Christensen (2020), some of the pathogens' effector molecules could suppress ETI, while those effectors become new signals to activate another ETI pathway (for review, see Thordal-Christensen, 2020). As a result, plants have hundreds of R genes as intracellular ETI receptors, and the number is increasing as pathogens evolve via mutations (Sarris et al., 2016).

Recently, Hasan et al. (2021a) reviewed the clubroot resistance loci of the A- genome of *B. rapa* and C-genome of *B. oleracea*, and listed a total of 24 and 36 loci, respectively, in these two genomes (Figure 1.3). Clusters of clubroot resistance loci seems to be present on A03 and A08 chromosomes where, respectively, 11 and five loci reported to date. NBS-LRR gene cluster also exist on chromosome 1 and 5 of *A. thaliana* which mainly resulted from tandem and segmental gene duplication events (for review, see Leister, 2004). It is well known that the *Brassica* genomes share evolutionary relationship with the *Arabidopsis* (Lagercrantz & Lydiate, 1996; Li et al., 2003; Handa, 2003). In this regard, it is also probable that the clubroot resistance gene cluster of A03 and A08 chromosomes of the *Brassica* A genome resulted from tandem

duplications.

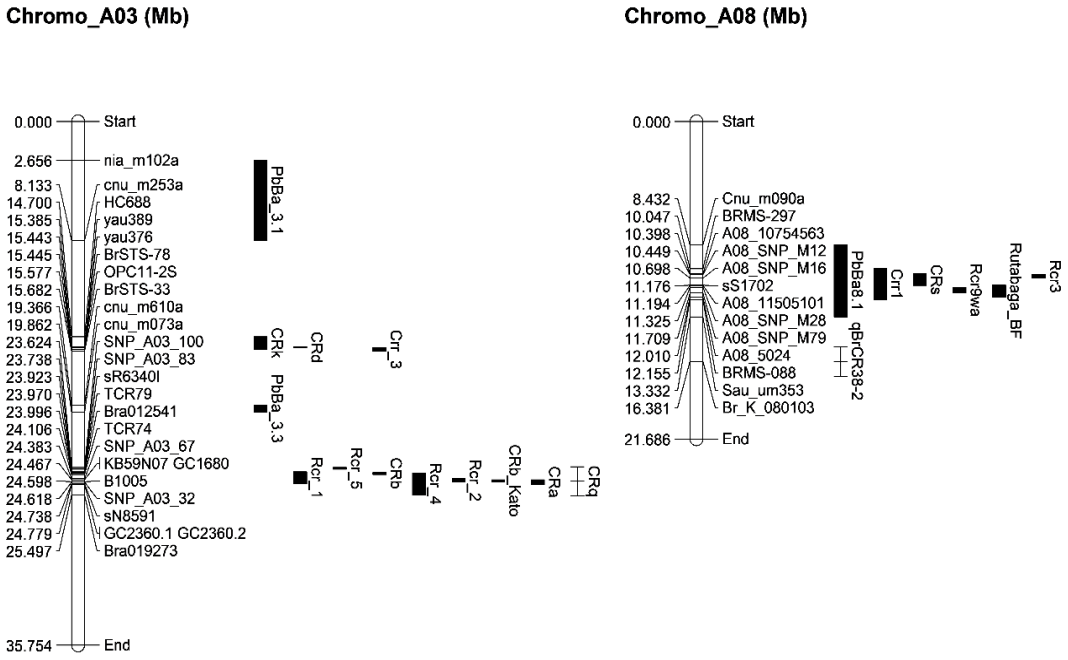


Figure 1.3 Clubroot resistance loci reported on chromosome A03 and A08 of *Brassica rapa* and their position in *Brassica napus* genome. The positions in the *B. napus* genome estimated based on flanking and co-segregating maker sequences. Physical map distances are in mega base pair (Mb) on the left side and marker names are on the right side; the maps are developed based on the reference genome Brassica_napus_v3.1 (http://cruciferseq.ca/?q=node/6).

1.4. Next-generation sequencing (NGS) techniques

With the availability of genome sequence information and affordable high-throughput and accurate next-generation sequencing (NGS) platforms, the discovery of the molecular basis of traits is increasing consistently thereby assisting genomics-based crop breeding. NGS techniques generate massive amounts of sequencing data from the whole genome; this is a much advanced technique of sequencing as compared to traditional Sanger sequencing which sequences a single DNA fragment at each time (Mertens & Sanger, 1985). Over the years, the NGS techniques have become more affordable with the price decreasing from $5000/Mb sequence data in 2010 to $0.008/Mb in recent years (for review, see Varshney et al., 2018) making NGS a readily available tool for sequencing the genomes of the crop plants for gene discovery, complex trait analysis by genome-wide association study (GWAS), global gene expression analysis, genomic

selection, and predictive breeding (for review, see Scheben et al., 2017).

NGS techniques can be divided into two groups based on their mechanism: sequencing by ligation (SBL) and sequencing by synthesis (SBS) (for review, see: Goodwin et al., 2016). To date, three main NGS platforms, *viz.* Pyro sequencing from Roche, Solexa sequencing from Illumina, and SOLID sequencing from Applied Biosystems, have dominated the market; these platforms are based on SBS chemistry (for review, see Hu et al., 2021 and Slatko et al., 2018). The common mechanism in these platforms is that variant nucleotides give unique colours or ionic concentrations when added to the nascent DNA strands by a DNA dependent DNA polymerase. A monitor records the colours or the ionic concentrations to reveal the sequence of the DNA template based on the nucleotide incorporated. The details of the mechanism and the variance of these three platforms have been reviewed by (for review, see Niedringhaus et al., 2011 and Zhou et al., 2010. Among them, the Solexa sequencing platform from Illumina has been used most extensively due to lower costs and higher throughput as compared to the other two platforms. The major limitations of this platform is that it generates reads of 150 or 300 nt in length (for review, see Hu et al., 2021 and Slatko et al., 2018). This type of shorter reads can pose problems while assembling the sequence reads as well as for gene annotation (for review, see Pop & Salzberg, 2008 and Hu et al., 2021). However, this platform works well for research with organisms for which genome sequence is already available, such as research involving whole genome resequencing (WGRS) approaches, where the sequence reads are assembled against an existing reference sequence.

Although NGS platforms have been available since 2001, its use in *B. napus* research has been limited due to complexity of the *Brassica* genomes. This impediment has been resolved to some extent with the release of the whole genome sequence of *B. rapa* (Wang et al., 2011), *B. oleracea* (Liu et al., 2014) and *B. napus* (Chalhoub et al., 2014; Rousseau-Gueutin et al., 2021). By using NGS technologies and the *Brassica* reference genomes, several transcriptomics- and whole genome sequence-based studies have been carried out to identify the genes involved in clubroot resistance as well as to unveil the mechanism of this resistance (e.g. Hasan et al., 2021c; Summanwar et al., 2021; Wang et al., 2019; for review, see Zhou et al., 2021).

1.4.1 Transcriptome sequencing

Transcriptome sequencing, also known as RNA-Seq, is commonly used in studies for rapid

profiling and deep investigation of transcriptomes. This is similar to genome sequencing and involves the use of a high-throughput deep-sequencing technology; however, transcriptome sequencing can also quantity expression level of a transcript in a specific tissue at a specific developmental stage or under a specific environmental condition. Thus, this technology is useful for deducing and quantifying the transcriptomes like the microarray technology; however, transcriptome sequencing is also capable of detecting novel transcripts (for review, see Wang et al., 2010).

Transcriptome refers to a set of transcripts generated under a specific condition, and this includes messenger RNA (mRNA) and different types of noncoding RNA (ncRNA), such as miRNAs, siRNAs, the long ncRNAs. Most of the transcriptomics-based studies were focused on mRNA, for which several RNA-seq kits, such as e.g. Illumina TruSeq, are available. However, in the recent years, ncRNA sequencing has also been exploited to identify the genes involved in disease resistance in *Brassica* including clubroot resistance (Joshi et al., 2016; Joshua et al., 2013; Park et al., 2019; Summanwar et al., 2019).

1.4.2 Identification of genes involved in clubroot resistance using transcriptome sequencing

Two different approaches can be employed to utilize RNA-seq data for identification of clubroot resistance genes. The first approach is bulked segregant RNA-seq (BSR-Seq). This is similar to bulk segregant analysis (BSA) based on molecular marker data (for review, see Hill, 1998); however, uses the RNA-Seq data. In this approach the RNA-Seq data is used for SNP discovery; following this, a comparison is made for SNP allele frequency between the resistant and susceptible bulks which have been used for RNA-Seq. The polymorphic SNP variant of the transcripts between the bulks reveals the loci or QTL to be associated with resistance phenotype. Following this approach, physical position of the clubroot resistance loci *Rcr1*, *Rcr2* and *Rcr5* of the chromosome A03 of *B. rapa* (Huang et al., 2017, 2019; Yu et al., 2016)*; Rcr3* and *Rcr9*^{wa} loci of A08 of *B. rapa* (Karim et al., 2020); *BniB015819* gene in *Rcr6* locus of the chromosome B03 of *B. nigra* (Chang et al., 2019); and *Bo7g108760* and *Bo7g109000* genes of C07 of *B. oleracea* (Dakouri et al., 2018) were identified. The approach of QTL-Seq is also a type of bulk segregation analysis which uses the whole-genome re-sequencing data. Thus, both BSR-Seq and QTL-Seq uses next-generation sequencing of the samples to locate the clubroot resistance loci. The BSR-Seq technically does not require a reference genome like QTL-Seq; although, it always

better to use reference genome to locate the loci (for review, see Zhou et al., 2021). However, QTL-Seq approach is based on gDNA which includes non-coding regions as well, while BSR-Seq is based on cDNA data derived from transcriptomes. The polymorphism or markers identified following BSR-Seq approach are from transcripts; therefore, this approach provides a more accurate genomic position of the loci (Liu et al., 2012; for review, see Lv et al., 2020). It is still unclear whether mutation rate is high or low in the coding region (for review, see Kern & Hahn, 2018; Nei & Kumar, 2000 and Subramanian & Kumar, 2003). However, it is well accepted that only a small fraction of the genomic DNA codes for protein (Dunham et al., 2012; for review, see Li & Liu, 2019); therefore, the amount of polymorphism could be found in the coding region is limited. To overcome this limitation while working with BSR-Seq, a large population need to be genotyped by polymorphic markers developed based on RNA-Seq data to fine map the locus. For example, Yu et al. (2016) used 1587 plants to map the locus *Rcr1,* Huang et al. (2017) used 675 plants to map *Rcr2,* Huang et al. (2019) used 824 plants to map *Rcr5*, and Karim et al. (2020) used 349 (240+109) plants to map *Rcr3* and *Rcr9*^{wa}. BSR-Seq has also been used to fine map the loci identified through linkage association or genome-wide association study (GWAS). For example, the clubroot resistance locus *Rcr1* (also designated as *Rpb1*) has been mapped at 0.54 cM and 0.77 cM away from the SSR markers sN8591 and sR3401, respectively, of A03 (Chu et al., 2013); subsequently, flanking SNP markers located at 1.07 MB to 0.24 MB away from the gene were developed using BSR-Seq approach, and two genes, *Bra019409* and *Bra019410*, were also identified in this genomic region (Chu et al., 2014; Yu et al., 2016).

The second approach is digital gene expression (DGE) based on global transcriptome analyses (Hanriot et al., 2008). The DGE principle was developed in 1997, which is based on counts of cognate sequence tags detected in individuals differing for a phenotypic trait (Audic and Claverie, 1997). While using NGS technology in DGE, the use of cognate sequence tags is not needed (Mortazavi et al., 2008); this technology provides a digital measure of the prevalence of transcripts from the known and unknown genes. Following this approach, the mechanism and modulation of clubroot resistance could be revealed. For example, by using near-isogenic lines for clubroot resistance, (Chen et al., 2016) detected different genes to be involved in host resistance including the genes involved in effector-triggered immunity. Genes involved in host resistance has also been identified using different types of populations in *Brassica* species (Chu

et al., 2014; Chen et al., 2016; Jia et al., 2017; Mei et al., 2019; Wang et al., 2019) or different species (Zhang et al., 2016). While comparing gene expression at different time points after inoculation such as 7, 14 and 21 days after inoculation (DAI), expression pattern of different genes can be investigated; based on this, salicylic acid (SA)-mediated host response was found to play an important role in triggering immune response (Galindo-González et al., 2020). Similar studies have also been carried out by different researchers with differentially expressed genes to understand the molecular basis of host resistance (Chen et al., 2016; Galindo-González et al., 2020; Mei et al., 2019; Zhang et al., 2016); however, functional characterization of the suggestive genes yet to be carried out. Summanwar et al. (2021) compared the differentially expressed genes identified using clubroot resistant lines carrying resistance at A03 or at A08 locus and their susceptible counterparts; besides identifying differentially expressed genes, 17 simple sequence repeats polymorphisms had been found in these genes, which could be used for SSR markers designing.

1.4.3 Noncoding RNAs and their function under stress

RNA-seq is primarily used to identify the expressed mRNAs; recently, it has been also used to identify the noncoding RNAs (ncRNA) affecting plant traits. The first ncRNA *micf*, a micRNA appearing to block the translation of the *ompF,* was identified in 1983 (Mizuno & Chou, 1984); however, their importance gained attention since 1994 with processing of the human genome sequences (Lander et al., 2001). This sequencing project showed that about 93% of the human chromosomal regions can be transcribed; however, only 1.2% codes protein (Dunham et al., 2012) suggesting that most of the RNAs do not translate to into protein.

Recently, ncRNA has received much attention with the discovery of their involvement in gene regulation (for review, see Morillon, 2018). They are usually categorized based on their size into small noncoding RNA (sncRNA) (less 200 nt), long noncoding RNA (lncRNA) (200 nt-50,000 nt) and very long intergenic RNA (vlincRNA) (over 50,000 nt) (Caron et al., 2018; for review, see Zampetaki et al., 2018); however, this classification is not fixed, which can be modified based on other characteristics of these noncoding RNAs (for review, see Boivin et al., 2019). Among them, lncRNA has received greater attention due to their regulatory functions during biotic or abiotic stress (for review, see Lee, 2012 and Shafiq et al., 2016). Most of the ncRNAs are similar to mRNAs, which are transcribed by RNA pol II and carry a 3' poly-A tail as

well as a 5' cap (Beaulieu et al., 2012; Derrien et al., 2012); however, some of the ncRNAs are transcribed by *RNA pol IV* or *pol IV* (Onodera et al., 2005, Wierzbick et al., 2008).

Since the discovery of the first functional plant lncRNA *Enod40* as a lncRNA decoy (Crespi et al., 1994), several functional lncRNA have been identified in plants and they are archived in publicly available databases, such as The Arabidopsis Information Resource (TAIR), Plant long noncoding RNA database (PLncDB), and Plant ncRNA database (PNRD) (Jin et al., 2013; Lamesch et al., 2012; Yi et al., 2015). lncRNAs do not encode proteins however, they can modify the structure, expression and transcription pattern of proteins involved in biotic or abiotic stress resistance (for review, see: Borah et al., 2018; Sun et al., 2018; Urquiaga et al., 2021 and Zhang et al., 2020). Several lncRNAs playing a role in resistance to different plant pathogens, such as *Phytophthora infestans* (late blight disease) (Cui et al., 2017), *fusarium* (fusarium head blight) (Huang et al., 2016), *Pectobacterium carotovorum* (bacterial soft rots) (Kwenda et al., 2016), *Verticillium dahliae* (verticillium wilt) (Zhang et al., 2018), *Blumeria graminis* (powdery mildew) (Xin et al., 2011), *Sclerotinia sclerotiorum* (white mold) (Joshi et al., 2016) and *P. brassicae* (clubroot) (Summanwar et al., 2019) have been identified. However, no reports are available to date demonstrating the possibility of developing a disease resistant plant through introgression of a lncRNA - their involvement in disease resistance has been mostly explained based on theoretical knowledge (Joshi et al., 2016; Xin et al., 2011; for review, see Ahmed et al., 2020).

Chapter 2. Genetic and molecular analysis reveals that two major loci and their interaction confer clubroot resistance in canola introgressed from rutabaga

2.1 Introduction

Canola (*Brassica napus* L.) oil contains a low amount of saturated fatty acids (<7%) and a high content of polyunsaturated omega-3 (10%) and omega-6 (20%) fatty acids (Przybylski et al., 2005); therefore, this oil is considered one of the best vegetable edible oils in the world. This oilseed crop plays an important role in crop rotation with cereals and other crops (Friedt et al., 2018; Harker et al., 2015) and the inclusion of canola in cropping systems is an important component of integrated crop management for successful crop agriculture.

Among the different biotic stresses affecting canola production, clubroot disease caused by *Plasmodiophora brassicae* (Woronin 1878) is one of the most important one. This disease causes an annual yield loss of about 15 % in *Brassica* crops worldwide (Dixon, 2009). The infection by *P. brassicae* results in the formation of characteristic clubs or galls on the roots of susceptible plants, which hinders water and nutrient uptake from soil resulting in wilting, yellowing, stunted growth of the plants, and ultimately reduced crop yield. Clubroot disease in Canadian canola fields was first identified in 2003 (Tewari et al., 2005). Since then, this disease has rapidly spread on the prairies where more than 3,000 infested fields have been identified in 2020; therefore, it is a serious threat to canola production in Canada (Canadian Plant Disease Survey, 2020). Once the pathogen is established in a field, it is difficult to eradicate it due to the extended longevity of its resting spores in soil (>15 years) (Wallenhammar, 1996).

Germplasm carrying clubroot resistance in canola is limited; however, resistance to a wide range of pathotypes can be found in its diploid progenitor species *Brassica rapa* (Hasan et al., 2012; Liu et al., 2018; Rahman et al., 2014) and *Brassica oleracea* (Farid et al., 2020). Resistance to multiple pathotypes has also been reported in rutabaga (*B. napus* var. *napobrassica*) (Hasan et al., 2012; Hasan & Rahman, 2016). Several researchers introgressed clubroot resistance into canola from *B. rapa* (Hasan et al. 2021b; Hirani et al., 2016; Liu et al., 2018) as well as from the European canola cv. Mendel (Rahman et al. 2012) and rutabaga cv. Brookfield (Hasan et al., 2012; Spaner, 2002) (for details, see Rahman et al. 2014). However, some of the resistances have become ineffective due to a change in pathogen population structure in soil and the evolution of new pathotypes (Strelkov et al., 2016). Currently, a total of 36

pathotypes have been reported in Canada; however, not all of them are equally virulent on canola. Some of them, such as pathotype 3A and 3D are most prevalent and virulent (Askarian et al., 2021; Hollman et al., 2021; Strelkov et al., 2016, 2018, 2020); and these pathotypes can overcome some of the resistances currently available in canola (Strelkov et al. 2018).

Conventional breeding can be employed for introgression of clubroot resistance into canola from different sources including its allied species (Hasan et al., 2021a, 2021b; Hasan & Rahman, 2016; Hirani et al., 2016; Liu et al., 2018; for review, see Rahman et al., 2014). However, introgression of resistance into canola is always easy when it is from the same species, such as from rutabaga (Hasan et al., 2021a; Hasan & Rahman, 2016). Several rutabaga accessions have been reported to carry resistance to different pathotypes (Fredua-Agyeman et al., 2020; Hasan et al., 2012) and the rutabaga cv. Polycross carries excellent resistance to clubroot including the recently evolved virulent pathotypes. This cultivar has been bred using a broad-based interbreeding population with selection under field conditions for resistance to multiple pathotypes (Spaner 2002). In fact, the clubroot resistance of rutabaga has been introgressed from European turnips (Bradshaw et al., 1997), and the European turnips carry multiple major clubroot resistance genes (for review, see Hasan et al., 2021c). Therefore, it has been hypothesized that the genetic control of clubroot resistance in rutabaga cultivars may vary widely. Following traditional breeding, clubroot resistance from the cv. Polycross has been introgressed into spring canola, where the resistant canola lines showed resistance to multiple pathotypes including the newly evolved pathotypes 3A and 3D. This resistance was used in this study to understand the genetic and molecular basis of this resistance.

Genetic analysis and molecular mapping of a trait can be carried out by using different segregating populations, such as F_2, backcross, recombinant inbred lines (RILs) and doubled haploids (DHs). Among these, DH population has been widely used by different researchers' due to their complete homozygosity which makes the DH population suitable for use in genetic research including molecular mapping of the traits (for review, see Weyen, 2021). Although the dominance effect of a quantitative trait locus (QTL) cannot be detected using a DH or RIL population, this type of population can be used to identify digenic epistatic effects (Kebede et al. 2012). Furthermore, a DH line, being immortal, can be used in studies for resistance to multiple pathotypes in replicated trials, which is not possible by using a F_2 or backcross population.

With the availability of *Brassica* genome sequence information (Chalhoub et al., 2014; Rousseau-Gueutin et al., 2021) it is possible to compare the sequences of a mapping population partitioned based on contrasting traits (e.g. resistant vs. susceptible) to identify single nucleotide polymorphism (SNP) variants that are associated with a trait. This approach has been applied in different crops including rice (Takagi et al., 2013), groundnut (Pandey et al., 2017), *Brassica juncea* (Zhang et al., 2018) and in *B. napus* (Yao et al., 2017; Hasan et al., 2021a).

The objectives of this research were to understand the genetic and molecular basis of resistance to *P. brassicae* pathotypes 3H, 3A and 3D in a spring *B. napus* canola population carrying clubroot resistance introgressed from the rutabaga cv. Polycross, mapping of this resistance as well as identification of the genes involved in this resistance. We have achieved this using a DH population and using the approach of whole-genome resequencing (WGRS) and linkage analysis, and expression analysis of the genes from the QTL regions.

2.2 Materials and Methods

2.2.1 Plant material

A DH population of 90 lines derived from F_1 of two spring canola lines 1CA1446.476-A1296 and A04-73NA were used in this study. 1CA1446.476-A1296 is a F_9 generation inbred line carrying resistance to multiple *P. brassicae* pathotypes including pathotype 3H, 3A and 3D; this line was developed following backcross and pedigree breeding method. In this, the clubroot resistant rutabaga cv. Polycross was crossed with a clubroot susceptible spring canola cv. Hi-Q and the F_1 was crossed with another clubroot susceptible canola line A03-73NA. The resultant 3-way F_1 [(Polycross × Hi-Q) × A03-74NA] was backcrossed to A04-73NA, and the backcross population was subjected to pedigree breeding from where the line 1CA1446.476-A1296 was developed. With this pedigree, it was expected that the DH population will be close to the susceptible canola line A04-73NA while the individual DH line will be differing for clubroot resistance. Use of this type of population was expected to reduce the background genetic difference for the other genes or DNA sequences and would allow mapping of the traits with high confidence.

2.2.2 Resistance test to *P. brassicae* pathotypes

Single spore derived isolates of *P. brassicae*, classified as pathotypes 3H, 3A, 3D based on Canadian clubroot differential (CCD) set (Strelkov et al., 2018) were used for inoculation. These isolates were obtained from Dr. Stephen Strelkov, Department of Agricultural, Food and Nutritional Sciences, University of Alberta, in the form of galls. Resting spore suspensions (inoculum) from the galls were prepared following the protocol described by Strelkov et al., (2007) and the concentration of the suspension was adjusted to 1×10^7 to 1×10^8 resting spores/mL.

The DH lines and their parents were seeded in a greenhouse (20-22/15 °C day/night, 16 h photoperiod, light intensity of 130 µmol m^{-2} s^{-1}) in 9 × 8 cells trays (cell size: 4 cm × 4 cm × 5 cm, L × W × D) filled with Sunshine Professional growing mix (Sun Gro Horticulture Canada Ltd, Seba Beach, Canada) for resistance to pathotype 3H, 3A and 3D. For testing resistance to each of the pathotypes, the experiment was carried out in three replications where each replication included eight plants of each DH line. The cv. Hi-Q was used as susceptible control. The seedlings were inoculated at 7-days after germination by pipetting 1 mL of spore suspension at the base of the seedling, and the inoculation was repeated on the following day to ensure successful infection. The plants were scored for disease severity at 50 days after inoculation. For this, the roots were washed with tap water and examined for galls and were rated on a 0–3 scale (Kuginuki et al., 1999), where 0 = no galls, 1 = a few small galls on the lateral roots, 2 = moderate galls, and 3 = severe galls. By using the resistance scores, disease severity index (DSI) was calculated for using the following formula,

$$\text{DSI}(\%) = \frac{\sum(n_0 \times 0 + n_1 \times 1 + n_2 \times 2 + n_3 \times 3)}{N \times 3} \times 100\%$$

where, *n* is the number of plants in a class, 0, 1, 2, and 3 is the disease symptom severity classes, and *N* is the total number of plants.

2.2.3 Statistical analysis of phenotypic data

Chi-square ($\chi 2$) test of segregation for resistance to a pathotype was carried out using the following formula: $\chi 2 = \sum(O - E)^2/E$, where O and E are the observed and expected numbers, respectively; the *p*-value was calculated using CHISQ.DIST.RT function in Excel. Principal component analysis (PCA) of the DH population using DSI data for resistance to the three

pathotypes (3H, 3A and 3D) was carried out using prcomp package in R (Kassambara, 2017), and the biplots were generated using Excel. Coefficient of correlation was calculated using the CORREL function in Excel.

2.2.4 DNA isolation, library preparation and whole-genome resequencing (WGRS)

Leaf samples of all DH lines and their parents at the age of two weeks after germination were collected and genomic DNA was extracted using SDS/NaCl DNA method (Kotchoni et al. 2009). Based on resistance phenotype of the DH lines (Supplementary Table S2.1), the following six DNA samples were prepared (Supplementary Table S2.2) for WGRS:

RP = Resistant parent 1CA1446.476-A1209; resistant to all three pathotypes (DSI 1.5 to 4.2%);

SP = Susceptible parent A04-73NA; susceptible to all pathotypes (3H, 3A and 3D) with DSI 100%;

RB = Bulk of 15 lines resistant to pathotypes 3H (DSI 5.5%), 3A (DSI 5.2%) and 3D (DSI 8.6%);

SB-3H = Bulk of three lines resistant to pathotypes 3A (DSI 8.8%) and 3D (DSI 12%), but partially susceptible to pathotype 3H (DSI 51.2%);

SB-3D = Bulk of five lines partially susceptible to pathotype 3D (DSI 37.4%) but resistant to pathotypes 3H (DSI 12.8%) and 3A (DSI 8.3%);

SB = Bulk of 15 lines susceptible to all pathotypes (3H, 3A and 3D) with DSI 100%.

A DSI of less than 25% is considered as resistant while a DSI of more than 70% is considered as susceptible (Dakouri et al., 2021).

The six DNA samples were sent to Genewiz NGS Laboratory (Montréal, Québec/Canada) for library preparation and paired-end (2×150 bp) sequencing using the Illumina HiSeq 4000 platform. All samples were sequenced 30X in two lanes to increase the depth of coverage. Sequence reads were trimmed using Trimmomatic v0.36 to a minimum read length of 35 bp, and the trimmed reads of the RP were aligned to two *B. napus* reference genomes, “Brassica_napus_v3.1” (http://cruciferseq.ca/?q=node/6) kindly provided by Dr. Isobel Parkin, AAFC, Saskatoon, SK, and “Brassica_napus_v4.1” (https://www.genoscope.cns.fr/brassicanapus/data/) (Chalhoub et al., 2014) using BWA v 0.7.12; these two reference genomes hereafter will be referred to as v3.1 and v4.1, respectively. After

aligning, the RP sequence was processed to identify the exact position of the mapped reads, and they were filtered using the mpileup function of Samtools v.1.3.1 (Li et al., 2009). The nucleotides of the RP and the two reference genomes were individually called from the mpileup output using VarScan v2.3.9 with the following settings: minimum coverage 10, the minimum number of reads 7, minimum variant frequency 10%, and the highest *p*-value 0.05. By using this, two consensus sequences of the RP based on the two reference genomes were generated using VCFtools v 0.1.15; these two consensus genomes of the RP hereafter will be referred to as resistant parent assembly RPA-v3.1 and RPA-v4.1, respectively. The mapping population was DH; therefore, any heterozygous SNP variants were ignored in this analysis. Trimmed reads of the other four bulks, RB, SB-3H, SB-3D and SB, were individually aligned to the two RPA's and SNPs were called for these bulks. All SNPs with read depth ≥10, meanwhile alternative allele present in the bulks were used for further analysis.

2.2.5 ΔSNP-index calculation and statistical analysis

The SNP indexes of each of the four bulks were calculated using the following formula (Abe et al., 2012):

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

The following calculations were made for ΔSNP-index of RB, SB-3H and SB-3D: SNP index of RB – SNP index of SB, SNP index of SB-3H – SNP index of SB, and SNP index of SB-3D – SNP index of SB. A ΔSNP index value of –1 for an SNP locus indicates that an allele is present in both RB and RP while an alternate allele is present in SB. The tricube-smoothed ΔSNP-index plot (Takagi et al., 2013), based on G' analysis (Magwene et al., 2011), was carried out using QTLseqr package (Mansfeld & Grumet, 2018) in R to visualize the distribution of the ΔSNP indices along the chromosomes. For this analysis, the read depth was set 11 to 400, replications 10,000, and window size 10 Mb. The output of the QTL regions was obtained using getQTLTable function in QTLseqr package with threshold of "alpha = 0.01". Physical positions of the SNPs, simple-sequence repeats (SSRs) or QTL reported in this article, if not otherwise stated, are based on the "Brassica_napus_v3.1" (http://cruciferseq.ca/?q=node/6) reference genome. However, in many cases, their physical position based on "Brassica_napus_v4.1" (https://www.genoscope.cns.fr/brassicanapus/data/) reference genome also given in brackets.

2.2.6 Allele-specific PCR (AS) primers design and genotyping

Based on WGRS data, SNP allele-specific (AS) primers from 85 SNP loci of A03 and A08 chromosomes observed to be associated with clubroot resistance were designed. This included primers of 37 loci from 10.27 to 16.25 Mb (v3.1) (9.27 to 14.98 Mb, v4.1) region of A03, and 48 loci from 6.22 to 14.54 Mb (v3.1) (6.09 to 12.37 Mb, v4.1) region of A08 (Supplementary Table S2.3 a, b). AS primers of 21 loci of A03 and 40 loci of A08 carried one mismatch nucleotide within the three bases from the 3'end (Liu et al., 2012) and the remaining primers carried two mismatches (You et al., 2008). In addition to the above-mentioned AS primers, AS primers from nine SNP loci of A08 reported to be associated with clubroot resistance (Karim et al., 2020; Suwabe et al., 2006; Zhu et al., 2019) were also used for genotyping. Electrophoretic separation of the polymerase chain reaction (PCR) amplicons was carried out on 2.0% agarose gel, stained with SYBR Green 1, and imaged with a Typhoon FLA 7000 laser scanner (GE_Healthcare _2010). In addition to the above-mentioned AS markers from 94 SNP loci, SSR markers from 55 loci spanning 8.01 to 20.45 Mb (v3.1) (7.23 to 19.02 Mb, v4.1) region of A03 and of 125 loci spanning 9.59 to 13.68 Mb (v3.1) (8.02 to 11.70 Mb, v4.1) region of A08 were also used for genotyping (Supplementary Table S2.3 c, d). For this, the PCR amplicons were detected using ABI sequencer 3730 (Applied Biosystems, Foster City, CA, USA), and fragment size was standardized with GeneScan^{TM} 600LIZ^{TM} size standard v2.0 with orange fluorophores. Among the published markers, OPC11-1S, OPC11-2S, BrSTS-33, BrSTS-54, BrSTS_61, HC688 were reported to be linked with clubroot resistance locus *Crr3*/*CRk*; Yau376 and Yau389 linked with *CRd*; BRMS-088 and A90_A08_SNP_M12 linked with *Crr1*; A90_A08_SNP_M28 and M79 linked with *Rcr9*; A90_A08_SNP_M16 linked with *Rcr3*; and Br_K_080103 linked with *qBrCR38_2* (Hirai et al., 2004; Karim et al., 2020; Pang et al., 2018; Saito et al., 2006; Sakamoto et al., 2008; Suwabe et al., 2006; Zhu et al., 2019).

For genotyping work, initially, the parental lines were genotyped with the above-mentioned 94 SNP-AS and 180 SSR markers. Based on genotyping results and physical position of the markers, 15 (six of A03 and nine of A08) SNP-AS and 17 (10 of A03 and seven of A08) SSR markers were used for genotyping the whole DH population (n = 90), using agarose gel or capillary electrophoresis.

2.2.7 Linkage map construction and QTL mapping

Linkage map of the A03 and A08 chromosomes, which were identified based on WGRS data to be carrying clubroot resistance, were constructed using genotyping data of the DH population and the software program QTL IciMapping version 4.2 (Meng et al., 2015). For this, recombination frequency estimated by using Kosambi mapping function (Kosambi, 2016) was used. Physical map of these two chromosomes was also constructed based on position of the markers. QTL mapping of the loci conferring resistance to different pathotypes and the estimates of their additive effects was carried out using the same software program with a walking speed of 0.4 cM and stepwise regression probability of 0.004; the threshold for LOD (p-value < 0.05) was obtained from 1,000 permutations to declare the location of the QTL. The additive effect and the percent phenotypic variation explained by the QTL were obtained from inclusive composite interval mapping model (ICIM) (Li et al., 2008). Digenic epistasis analysis for additive × additive effect was carried out using QTL IciMapping version 4.2 (Meng et al., 2015) with a walking speed of 1.0 cM and stepwise regression probability of 0.004; the threshold for LOD (p-value < 0.05) was obtained from 1,000 permutations.

2.2.8 Genotyping the mapping population with multiple markers in a PCR reaction

In addition to genotyping the mapping population with a single marker in a PCR reaction for construction of linkage maps, genotyping the population was also carried out using two markers, one from each QTL, in a PCR reaction for linkage analysis of the markers with resistance. For this, PCR amplification and electrophoretic separation of the PCR products was carried out following the same procedures described above.

2.2.9 Principal component analysis (PCA) of the DH population using genotypic data

The genetic distance matrix of the DH lines (n = 90) was calculated by using genotypic data of 15 SNP-AS and 17 SSR markers, which were used to construct the linkage map. Analysis was carried out using the software program TASSEL v5.2.73 (Bradbury et al., 2007), and the biplot was constructed in Excel using the PCA 1 and PCA 2 values obtained from TASSEL.

2.2.10 Bioinformatic analysis for putative clubroot resistance genes in the QTL regions

In Brassicaceae Database (BRAD), 244 genes of *B. rapa* are listed to be related to disease resistance (http://brassicadb.cn/#/ResistanceGene/). We downloaded the sequence of these genes and identified their homologues in *B. napus* by using the "Blat" function of Genoscope Brassica Napus Genome Browser (https://www.genoscope.cns.fr/blat-server/cgi-bin/colza/webBlat). The two QTL that we detected in the present study are located at 13.27-14.24 Mb (v4.1) region of A03 and 8.43-9.26 Mb (v4.1) region of A08; therefore, we filtered the homologous *B. napus* genes located in the 12-15 Mb (v4.1) region of A03 and 8-11 Mb (v4.1) region of A08 and identified a total of 39 genes from these two genomic regions. To further confirm putative function of these genes, we searched the annotation database in NCBI (*Brassica napus* Annotation Release 101), *i.e.,* accession NC_027759.2 and NC_027764.2, for these 39 genes. However, we were not able to detect any functional domain in 12 of these genes using SMART (http://smart.embl-heidelberg.de/), and therefore, they were eliminated from further study. Thus, a total of 27 putative genes, 13 from A03 and 14 from A08, presumably affecting clubroot resistance were selected and used for expression analysis. Of the 27 genes, 13 carry leucine-rich repeats (LRR) where some of them carry one or several nucleotide binding site (NBS) and/or TIR and/or coiled-coil (CC) domains, while the remaining 14 genes were considered to be involved in clubroot resistance based on their annotated function from NCBI (Supplementary Table S2.4). In case of the chromosome A03, nine genes were located within the QTL region and four were located at 0.5 Mb up- or down-stream of this region. For A08, 13 genes were located in the QTL region, and one was located 0.5 Mb downstream of this region.

2.2.11 Expression analysis with quantitative reverse transcription PCR (qRT-PCR)

The genetic distance matrix based on PCA showed that the DH population could be divided into four groups: DHs carrying (I) RP alleles on both A03 and A08, (II) RP alleles on A03 but SP alleles on A08, (III) RP alleles on A08 but SP alleles on A03, and (IV) SP alleles on both A03 and A08 (Supplementary Figure S2.1). Of the four groups, three groups of DH lines belonging to II, III and IV were used for qRT-PCR; 20 lines from each group were used for this analysis. For this, the three groups of DH lines were seeded in a greenhouse in three replications where each replication for each group included 120 plants (6 plants × 20 lines). Half of these plants (3 × 20) were inoculated with *P. brassicae* pathotype 3A to study expression of the putative clubroot

resistance genes identified in the QTL regions after infection, while the other half (3 × 20) was mock-inoculated with water. Root tissue of the 20 lines of each of the groups were collected at 7- and 14-days post inoculation (dpi) and were bulked. This constituted a total of 36 bulks: 3 groups × 2 time points × 2 treatments × 3 replications. Total RNA of the bulks was isolated using TRIzol reagent (Invitrogen, USA), and cDNA was synthesized using RevertAid First Strand cDNA Synthesis Kit (ThermoFisher Scientific) following manufacturer's instructions. qRT-PCR was carried out on a StepOne Plus real-time PCR system (Life Technologies, Burlington, Canada) using FASTSYBR Green mix from (ThermoFisher Scientific, USA) Applied Biosciences. Two technical replicates of each biological replicate were analyzed.

Gene-specific primer pairs for 27 putative genes were designed using Primer Express v3.0.1 (Life Technologies, ON, Canada). Constitutively expressed housekeeping gene Ubiquitin-Conjugating Enzyme 10 (*UBC10*) from *B. napus* was used as endogenous control. The relative expression level of each gene was calculated using the $2^{-\Delta\Delta Ct}$ method (Livak & Schmittgen, 2001). Data was statistically analysed with a generalized linear model with two factor-ANOVA (PROC MIXED LSMEANS STATEMENT of Statistical Analysis System software, SAS Institute INC 2000). The significance was declared at *p*-value of 0.05.

2.3 Results

2.3.1 Inheritance of clubroot resistance

Of the 90 DH lines tested, 48 were resistant (DSI ≤25%) to all three pathotypes while 18 lines were highly susceptible (DSI≥95%); DSI for the remaining lines varied from 26.4% to 94.4% (Supplementary Table S2.1). The coefficient of correlation for DSI between the replications was ≥ 0.9 (Supplementary Table S2.5), indicating reliability and consistency of scoring data from three replications. The distribution of disease severity in response to infection by pathotypes 3H, 3A and 3D could be divided into two major classes: resistant lines with DSI ≤ 25% and susceptible or moderately susceptible lines with DSI > 50% (Figure 2.1 a,b,c); lines with DSI between 25-50% could be considered as modestly resistant. Of the whole DH population, 65, 64 and 65 lines were resistant or modestly resistant, while 24, 24 and 23 lines were susceptible to pathotype 3H, 3A and 3D, respectively. This distribution followed a 3:1 segregation ($\chi^2 = 0.18$, $p = 0.91$; $\chi^2 = 0.24$, $p = 0.89$; $\chi^2 = 0.06$, $p = 0.97$), suggesting that resistance to each of the three pathotype is controlled by two independent gene loci. Coefficients of correlation for resistance

(DSI values) to pathotype 3H vs. 3A, 3A vs. 3D, 3H vs. 3D were 0.98, 0.97 and 0.97, respectively (Supplementary Table S2.5) indicating that these two loci play an important role in the control of resistance to these pathotypes. This could also be confirmed by PCA analysis of the DH population for resistance to pathotype 3H, 3A and 3D which resulted two distinct clusters– one cluster included the lines showing resistance to these pathotypes while the other cluster included the susceptible lines (Figure 2.1d).

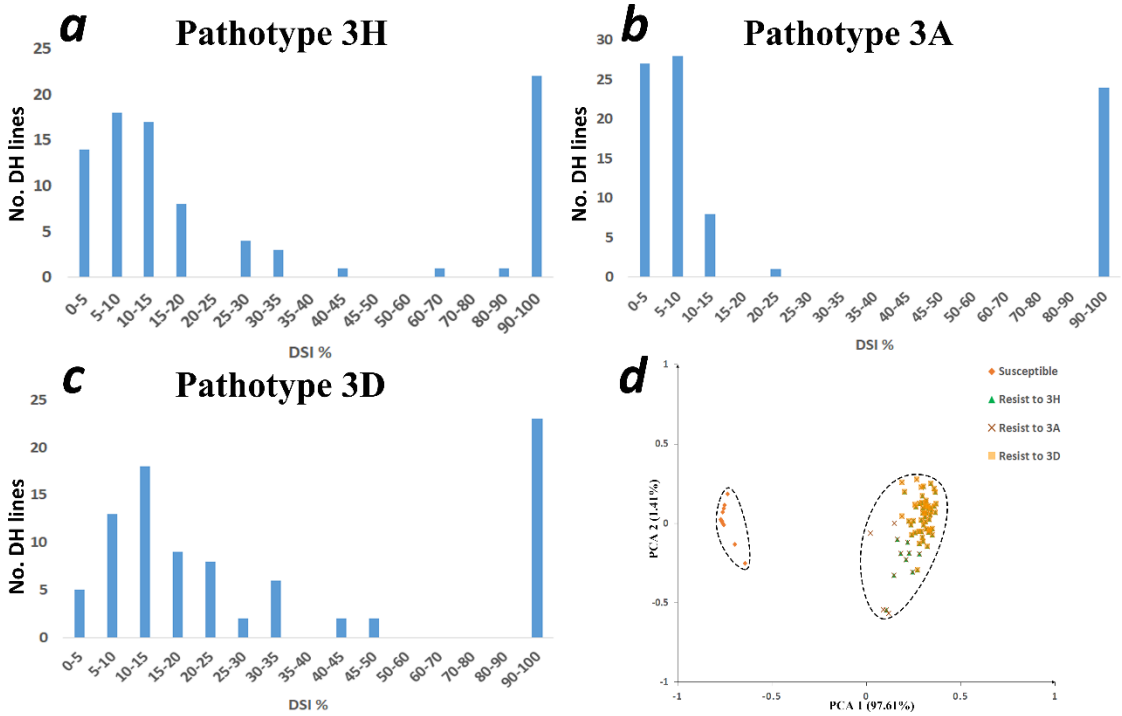


Figure 2.1 (a-c) Frequency distribution and (d) principal component analysis (PCA) of the *Brassica napus* doubled haploid (DH) lines (n = 90) for disease severity index (DSI %) obtained by inoculating the lines with *Plasmodiophora brassicae* pathotypes 3H, 3A and 3D. In case of PCA figure, lines with DSI more than 50% for resistance to all three pathotypes are showed with red-diamond; lines with DSI less than 25% for resistance to pathotype 3H, 3A and 3D are indicated with yellow-square, green-triangle and claret-cross, respectively; the PCA placed the population into two groups.

2.3.2 Whole-genome re-sequencing (WGRS) analysis

The WGRS data were generated for six bulks, *i.e.*, RP, SP, RB, SB, SB-3D, and SB-3H (Supplementary Table S2.2) using two reference genomes v3.1 and v4.1. After SNP calling, 609,531 SNPs based on v3.1 and 496,583 SNPs based on v4.1 were identified between the RP and SP (Table 2.1). Identification of the SNPs with ΔSNP-index of –1 was carried out for RB,

SB-3H and SB-3D from the following comparisons: RB vs. SB, SB-3H vs. SB and SB-3D vs. SB. Based on this, of the total number of SNPs detected, 1.89%, 1.36% and 0.91% SNPs based on v3.1 and 2.16%, 1.55% and 1.00% SNPs based on v4.1 were found to have ΔSNP-index of –1 in the RB, SB-3H and SB-3D samples comparing with SB, respectively. Thus, the use of two *B. napus* reference genomes yielded similar results.

As mentioned in the materials and methods, RB carries resistance to all three pathotypes (3H, 3A and 3D) while SB is susceptible to all pathotypes, SB-3H carries complete resistance to 3A and 3D but partial susceptibility (DSI 51.2%) to 3H, and SB-3D carries complete resistance to 3H and 3A but partial susceptibility (DSI 37.4%) to 3D. Based on this, the RB vs. SB comparison was expected to disclose the SNPs associated with resistance to all three pathotypes; the SB-3H vs. SB comparison was expected to disclose the SNPs associated with resistance to pathotype 3A and 3D; and the SB-3D vs. SB comparison was expected to disclose the SNPs associated with resistance to pathotype 3H and 3A. The RB vs. SB and SB-3D vs. SB comparison's identified SNPs with ΔSNP-index of –1 mostly from chromosome A03 and A08, while the SB-3H vs. SB comparison identified SNPs mostly from A03 (Supplementary Table S2.6 a,b; Supplementary Figure S2.2). Analysis of data using QTLseqr identified two QTL on A03 and A08 from RB vs. SB and SB-3D vs. SB comparisons (Figure 2.2a, c), and one QTL on A3 from SB-3H vs. SB comparisons (Figure 2.2b). In case of the A03, the SNPs with ΔSNP-index of –1 based on RB vs. SB and SB-3D vs. SB are located at 14.20-20.95 Mb (v4.1: 12.76-19.54 Mb) and 15.71-17.90 Mb (v4.1: 14.44-16.62 Mb), respectively. In case of A08, the SNPs with ΔSNP-index of –1 based on RB vs. SB, SB-3H vs. SB and SB-3D vs. SB are located at 9.96-13.51 Mb (v4.1: 8.48-11.70 Mb), 10.02-16.70 Mb (v4.1: 8.43-14.24 Mb) and 10.18-13.51 Mb (v4.1: 9.56-11.62 Mb), respectively. The genomic regions identified using the above-mentioned comparisons overlapped; therefore, we have considered them as one QTL located on A03 at 14.20-20.95 Mb (v4.1: 12.76-19.54 Mb) and the other QTL on A08 at 9.96-16.70 Mb (v4.1: 8.43-14.24 Mb) (Figure 2.2d).

Thus, the distribution of the SNPs with ΔSNP-index = –1 estimated based on RB vs. SB comparison identified two QTL on A03 and A08 (Figure 2.2a) and the same two QTL could also be detected based on the SB-3D vs. SB comparison (Figure 2.2c). However, the A03 QTL could not be detected based on the SB-3H and SB comparison (Figure 2.2b). The distribution of the

SNPs with ΔSNP-index =-1 calculated based on reference genome v4.1 gave similar results (Supplementary Figure S2.3).

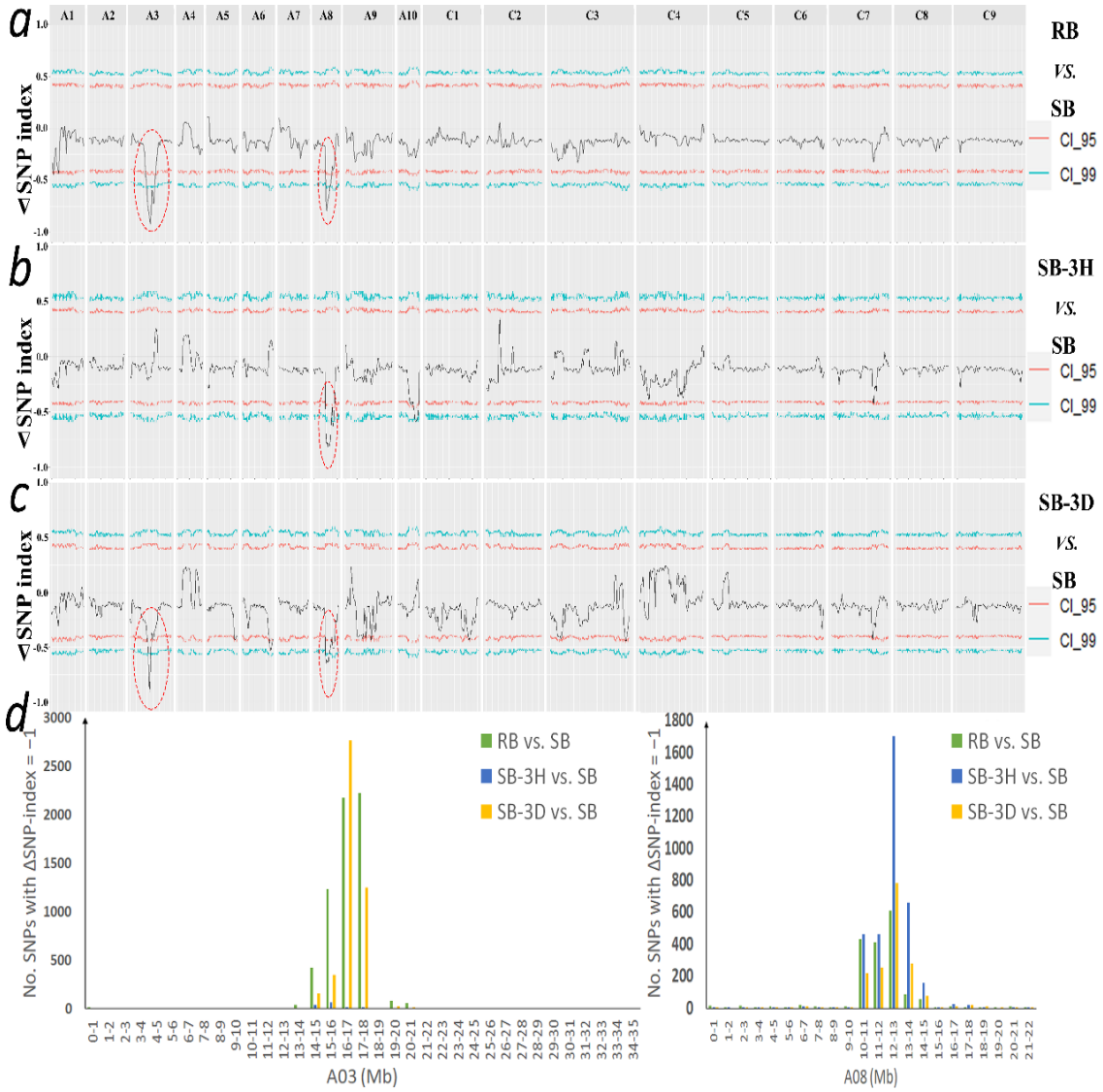


Figure 2.2 (a-c) The ΔSNP-index plots of *Brassica napus* chromosomes developed using three bulks of doubled haploid (DH) lines : RB = Bulk resistant to pathotype 3H, 3A and 3D; SB = bulk susceptible to pathotype 3H, 3A and 3D; SB-3H = bulk partially susceptible to 3H (DSI 51.2%) but resistant to 3A and 3D; and SB-3D = bulk partially susceptible to 3D (DSI 37.4%) but resistant to 3H and 3A (Supplementary Table S2.3). Analysis carried out using whole genome resequencing (WGRS) data based on the *B. napus* reference genome Brassica_napus_v3.1 (http://cruciferseq.ca/?q=node/6). Putative quantitative trait loci (QTL) for clubroot resistance are encircled. (d) Frequency distribution of the SNPs with ΔSNP-index = -1 positioned (Mb) on the A03 and A08 chromosomes; the ΔSNP-indices were calculated by comparing the above-mentioned bulks.

2.3.3 Marker development and mapping of the candidate region

To further confirm the above-mentioned two QTL and to develop markers associated with resistance, 92 markers from A03 and 182 markers from A08 were tested on the two parents, where 39 makers of A03 and 42 of A08 were found polymorphic; no amplification product could be detected for 16 markers of A03 and 33 of A08 (Supplementary Table S2.3 a,b,c,d). The whole mapping population was genotyped using 16 polymorphic markers of A03 and 16 of A08, and partial linkage map of these two chromosomes were constructed (Figure 2.3). These markers were selected based on their physical position and by avoiding multiple markers from the same position.

QTL mapping identified a peak located between the markers SNP_5241 and SSR_yau106 with LOD score of 14.1 for resistance to pathotype 3D, and a peak between the markers SSR_yau301 and SSR_yau106 with LOD of 11.3 to 11.8 for resistance to pathotype 3H and 3A on chromosome A03 (Figure 2.3; Table 2.2). The flanking markers for these two genomic regions are 6.4 and 1.6 cM apart, respectively, and these three QTL regions were overlapping; therefore, we considered them as one QTL conferring resistance to all three pathotypes. Based on physical position of the flanking markers, the A03 QTL was located at 14.41-15.44 Mb (v4.1: 13.27-14.24 Mb), *i.e.,* within one Mb region (Figure 2.3). In case of chromosome A08, three overlapping QTL peaks spanning a region 20 cM downstream of the marker SNP_5269 (0.0 cM) to the marker SSR_3319 (79.9 cM) was identified with LOD score of 13.93, 14.16 and 14.90 for resistance to 3H, 3A and 3D, respectively (Figure 2.3). As these peaks were overlapping, we considered them as one QTL conferring resistance to these three pathotypes. This is a large genomic region between the flanking markers; therefore, we tested 120 additional markers from this genomic region to saturate the linkage map; however, none of them could increase the resolution of the map. The two flanking markers are located at 7.85-11.09 Mb (v4.1: 7.48 - 9.26 Mb) region of A08, *i.e.,* at about the 2-3 Mb region. We further compared the genomic regions identified based on ΔSNP-index of –1 (Figure 2.2) and the genomic regions identified by using the linkage map and physical position of the QTL markers (Figure 2.3). Based on this, the A03 QTL region could be deduced to 14.41-15.44 Mb (v4.1: 13.27-14.24 Mb) and the A08 QTL to 9.96-11.09 Mb (v4.1: 8.43-9.26 Mb), *i.e.,* we could narrow down the two QTL regions, which spanning about 0.97-1.03 Mb on A03 and about 0.83-1.13 Mb on A08.

QTL analysis showed that the A03 QTL exerted an additive effect of 22.03%, 22.95% and 23.03% DSI (%) and explained 29.95%, 28.29% and 34.80% of the total phenotypic variance for resistance to pathotypes 3H, 3A and 3D, respectively; in this case, the allele contributing to reduced DSI derived from the cv. Polycross. Similarly, the A08 QTL exerted an addictive effect of 24.91%, 27.13% and 24.10% DSI and explained 38.81%, 40.01% and 38.58% of the total phenotypic variance for resistance to these three pathotypes (Table 2.2); in this case also the allele contributing to reduced DSI derived from the cv. Polycross. A digenic epistasis analysis showed that the above-mentioned two QTL jointly explained 89.86%, 93.10% and 89.12% of the total phenotypic variance for resistance to these three pathotypes (Table 2.3). This is greater than the phenotypic variance explained by the individual QTL; therefore, the existence of epistatic effect of these two QTL was evident. The epistasis analysis showed that these two QTL exerted an additive × additive interaction effects of 21.09%, 23.93% and 18.93% reduced DSI for resistance to 3H, 3A and 3D, respectively (Table 2.3). In addition to this, epistasis analysis also identified an epistatic QTL on A08 (80 cM), which is located at about 0.5 cM downstream of the main-effect A08 QTL (20.0-79.6 cM). This QTL individually exerted almost no additive effect (2.33% DSI); however, this epistatic QTL in interaction with the A03 QTL reduced 21.08% DSI for resistance to pathotype 3H. Similarly, an 81 cM genomic region of A08 with almost no additive effect (2.49% DSI) reduced 20.44% DSI in interaction with the A03 QTL for resistance to pathotype 3D. A similar epistatic locus was also detected on A03 (35 cM) at about 20 cM downstream of the major-effect A03 QTL (8.4-14.8 cM). This epistatic A03 QTL, individually, did not exert any strong effect (1.62% DSI); however, this locus in interaction with the main-effect A08 QTL reduced 23.25% DSI for resistance to pathotype 3A (Table 2.3).

Table 2.1 Summary whole-genome resequencing (WGRS) data of *Brassica napus* resistant (RP) and susceptible (SP) parents, and different bulks of doubled haploid (DH) lines based on resistance to *Plasmodiophora brassicae* pathotype 3H, 3A, and 3D. Resistance in this population is derived from rutabaga cv. Polycross.

Sample [1]	Yield (Gb)	Ref[2]	Total no. reads	Mapped (%)	Ref[3]	Total no. reads	Mapped (%)
RP	44.68	v3.1	298,429,735	96.99	v4.1	289,243,824	97.78
SP	39.19	v3.1	257,250,455	99.24	v4.1	253,000,371	99.45
RB	51.69	v3.1	340,380,884	98.97	v4.1	334,611,303	99.27
SB	45.13	v3.1	295,858,814	99.01	v4.1	291,834,701	99.22
SB-3H	45.13	v3.1	296,596,373	99.08	v4.1	291,986,981	99.36
SB-3D	46.34	v3.1	304,815,342	86.59	v4.1	299,288,402	86.73

[1] RB = Bulk resistant to pathotype 3H, 3A and 3D; SB = bulk susceptible to pathotype 3H, 3A and 3D; SB-3H = bulk partially susceptible to 3H (DSI 51.2%) but resistant to 3A and 3D; and SB-3D = bulk partially susceptible to 3D (DSI 37.4%) but resistant to 3H and 3A

[2] Reference genome, Brassica_napus_v3.1 (v3.1) is an unpublished whole genome sequence of *B. napus* line DH12075, which was derived from a cross between the French spring-type cv. Cresor and the Canadian spring-type cv. Westar (kindly provided by Dr. Isobel Parkin, AAFC).

[3] Reference genome, Brassica_napus_v4.1 (v4.1) is a published whole genome sequence of the French winter *B. napus* cv. Darmor-bzh.

Table 2.2 Quantitative trait loci (QTL) for resistance to *Plasmodiophora brassicae* pathotypes 3H, 3A and 3D detected in a *Brassica napus* doubled haploid (DH) population (n = 90) carrying clubroot resistance of rutabaga cv. Polycross; analysis carried out using inclusive composite interval mapping (ICIM-Add) method.

Pathotypes	Chromosome	Marker interval	Peak of QTL (cM)	Confidence interval (cM)[1]	LOD[2]	PVE (%)[3]	ADD[4]
3H	A03	SSR_yau301 and SSR_yau106	14.8	14.0 - 14.8	11.80	29.95	-22.03
3H	A08	SNP_5269 and SSR_3319	77.6	21.2 - 79.6	13.93	38.81	-24.91
3A	A03	SSR_yau301 and SSR_yau106	14.8	14.0 - 14.8	11.25	28.29	-22.95
3A	A08	SNP_5269 and SSR_3319	77.6	20.0 - 79.6	14.16	40.01	-27.12
3D	A03	SNP_5241 and SSR_yau106	14.8	8.4 - 14.8	14.10	34.80	-23.03
3D	A08	SNP_5269 and SSR_3319	78.0	23.6 - 79.6	14.90	38.57	-24.10

[1]Confidence interval: the region with LOD >3.0. [2] LOD: logarithm of the odds score. [3] PVE (%): Phenotypic variation explained by the QTL. [4]ADD: Estimated additive effect of the QTL; negative sign indicate cv. Polycross allele reduces disease severity index (DSI).

Table 2.3 Epistatic effect of the quantitative trait loci (QTL) of the chromosomes A03 and A08 for resistance to *Plasmodiophora brassicae* pathotypes 3H, 3A and 3D detected in a *Brassica napus* doubled haploid (DH) population (*n = 90*) carrying clubroot resistance of rutabaga cv. Polycross; analysis carried out using inclusive composite interval mapping (ICIM-Epi) method.

Patho-types	Marker interval of A03 QTL	Peak of A03 QTL (cM)	Marker interval of A08 QTL	Peak of A08 QTL (cM)	LOD[1]	PVE (%)[2]	Add[3] A03 QTL	Add[3] A08 QTL	A × A[4]
3H	SSR_yau301 - SSR_yau106	14	SNP_3319 - SNP_5039	80	27.17	57.03	-21.88	2.33	21.08
3H	SSR_yau301 - SSR_yau106	14	SNP_5269 - SSR_3319	79	25.83	89.86	-21.89	-20.57	21.09
3A	SSR_BnGMS417 - SSR_sau_um398	35	SNP_5269 - SSR_3319	60	42.57	68.79	-1.62	-23.92	23.25
3A	SNP_5241 - SSR_yau106	14	SNP_5269 - SSR_3319	79	34.58	93.10	-22.86	-22.40	23.93
3D	SNP_5242 - SSR_yau301	9	SNP_3319 - SNP_5039	81	25.32	58.07	-21.26	2.49	20.44
3D	SNP_5241 - SSR_yau106	14	SNP_5269 - SSR_3319	79	22.30	89.12	-22.89	-19.65	18.93

[1] LOD: logarithm of the odds score. [2] PVE (%): Phenotypic variation explained by the QTL. [3]Add: Estimated additive effect of the QTL; negative sign indicate cv. Polycross allele reduces disease severity index (DSI). [4] A × A: Estimated additive by additive interaction effect of two QTL reducing DSI.

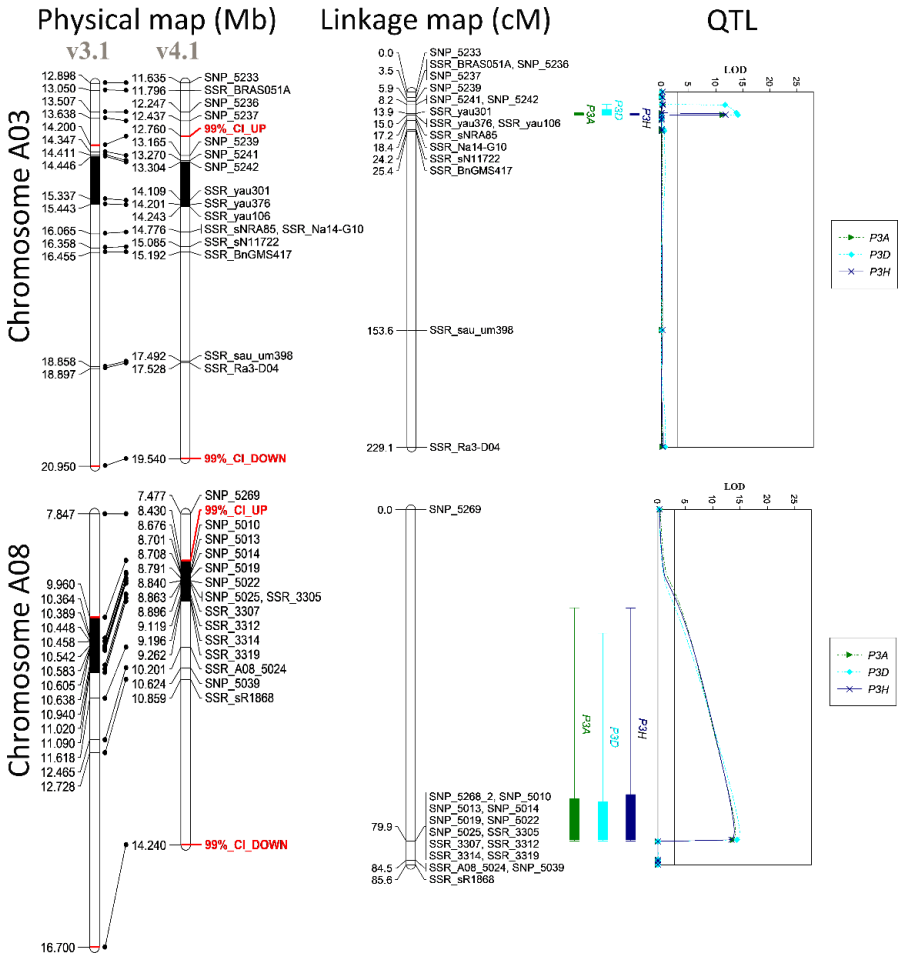


Figure 2.3 Linkage (middle) and physical (left) maps of a part of A03 and A08 chromosomes of *Brassica napus* carrying QTL for clubroot resistance. The linkage maps are constructed using SSR and SNP markers. The marker distances are in centimorgan (cM) which are shown on the left side, and markers names are shown on the right side. The QTL positions for resistance to *Plasmodiophora brassicae* pathotype 3H, 3A and 3D are shown with green-triangle, blue-diamond and purple-cross, respectively. The black box shows the location of the QTL with LOD score. Physical map distances are in megabase pair (Mb) on the left side and marker names are on the right side. The maps are developed based on two reference genomes, Brassica_napus_v3.1 (http://cruciferseq.ca/?q=node/6) (v3.1) (left side) and Brassica_napus_v4.1(Chalhoub et al., 2014) (v4.1) (right side). The boundaries of the putative QTL are labelled with red lines.

2.3.4 Genotyping the mapping population with two markers in a PCR reaction

Five markers (SNP_5241, SNP_5242, SSR_yau301, SSR_yau376 and SSR_yau106) from A03 and 12 markers (SNP_5268_2, SNP_5010, SNP_5013, SNP_5014, SNP_5019, SNP_5022, SNP_5025, SSR_3305, SSR_3307, SSR_3312, SSR_3314, SSR_3319) from A08, which were

used for construction of the genetic linkage maps were analysed for co-segregation with resistance; however, no strong linkage association could be found when a single marker was taken into account. However, analysis of marker data jointly from the two QTL regions showed a better co-segregation of the marker genotypes and the resistance phenotype. To further confirm this, we used the following four combinations of SNP markers from the A03 and A08 QTL in a PCR reaction: SNP_5242 (A03) and SNP_5013 (A08), SNP_5242 (A03) and SNP_5014 (A08), SNP_5242 (A03) and SNP_5019 (A08), and SNP_5242 (A03) and SNP_5025 (A08), and genotyped the whole mapping population which showed excellent co-segregation with resistance to pathotypes 3H, 3A and 3D. For example, while using the SNP_5242 marker of A03 and SNP_5019 marker of A08 in a single PCR reaction, the genotypic data showed that all resistant DH lines carried the RP allele of, at least, one of the two loci, while the susceptible lines are lacking the resistance alleles (Figure 2.4). To investigate whether the A03 and A08 QTL contribute in a similar way to resistance to the different pathotypes, we divided resistant lines (DSI 0-25%) into two groups, lines with DSI less than 5% (highly resistant) and lines with DSI 5-25% (resistant), and further partitioned each group based on marker genotype. The genotype composition of the highly resistant and resistant groups varied widely while comparing for resistance to pathotype 3H or 3A or 3D (Figure 2.5). For example, in case of resistance to pathotype 3H, about 93% (50+43) of the highly resistant lines carried RP allele of A03, while only 49% (24+25) of the resistant lines carried this allele. Similar differences between these two groups of lines could also be observed for resistance to other pathotypes (Supplementary Table S2.7). This suggests that the two genomic regions of A03 and A08 were contributing to resistance to these three pathotypes in a different manner.

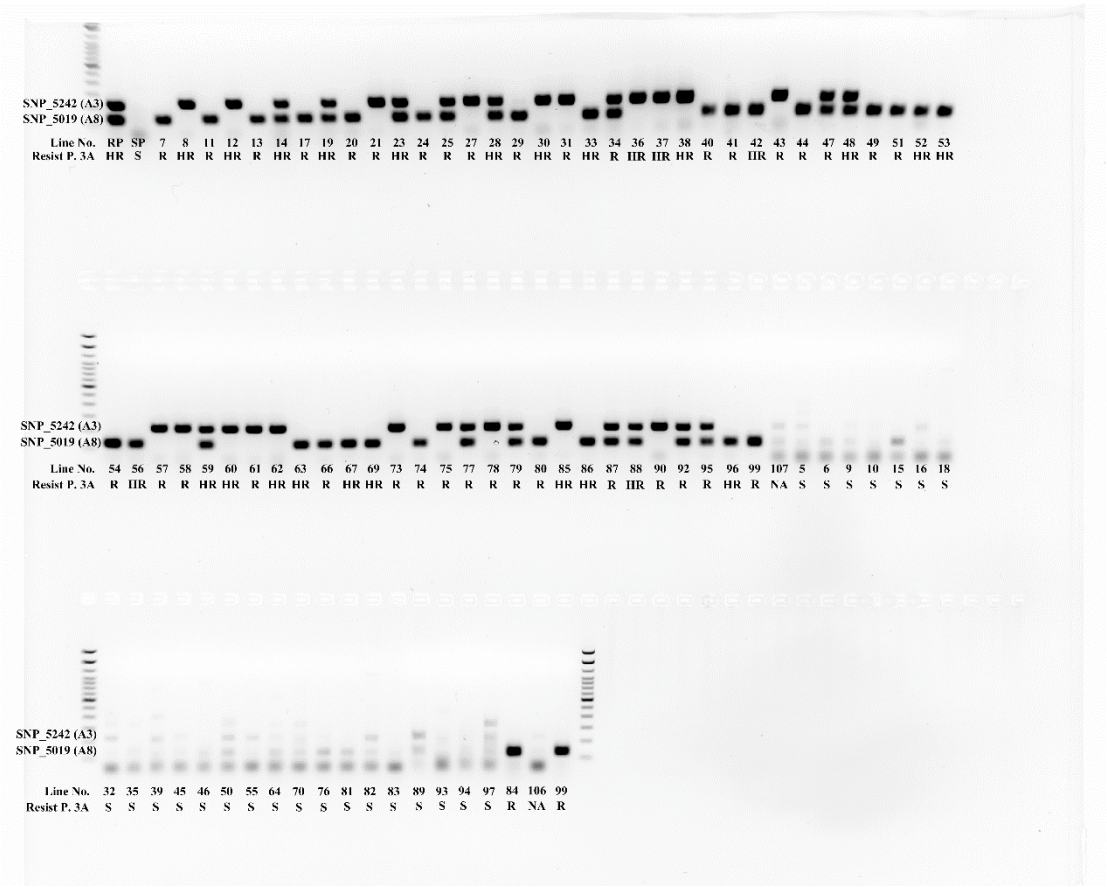


Figure 2.4 Gel image of 90 *Brassica napus* doubled haploid (DH) lines genotyped with SNP-allele specific markers; Invitrogen™ 100 bp DNA Ladder (ThermoFisher) also included in the gel image.The band around 228 bp is the SNP marker (SNP_5242) allele of A03 QTL co-segregating with clubroot resistance, while the band around 129 bp is the SNP marker (SNP_5019) allele of A08 QTL co-segregating with clubroot resistance. Clubroot resistance phenotype of the lines in response to *Plasmodiophora brassicae* pathotype 3A showed at the bottom, where HR = highly resistant, DSI <5%; R = resistant, DSI 5-25%; S = susceptible, DSI >90%. Additional information of the markers and DH lines can be found in Supplementary Tables S2.1 and S2.4, respectively.

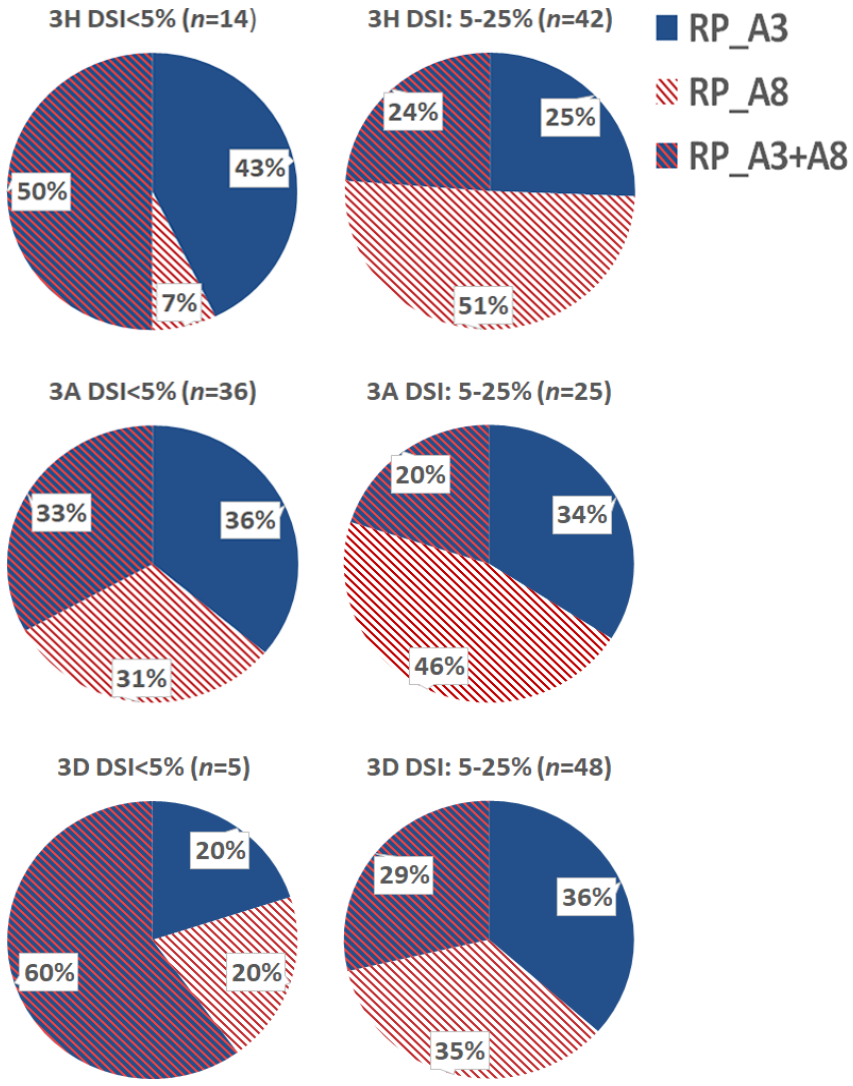


Figure 2.5 Pie chart for marker genotype of the *Brassica napus* doubled haploid (DH) lines highly resistant (DSI <5%) or resistant (DSI 5-25%) to *Plasmodiophora brassicae* pathotypes 3H, 3A and 3D. Slices of the pie chart covered by blue background denote the proportion of the DH lines carrying resistant parent (RP) allele of the chromosome A03, red strip denotes the lines carrying RP allele of the A08, and overlap region (blue background and red stripe) denotes the lines carrying RP alleles of both chromosomes. The numbers within the slice indicate the proportion of the total number of lines; the total number of lines are showed at the top of the charts with '*n*'.

2.3.5 Identification of putative candidate genes in the QTL regions and their expression analysis with qRT-PCR

To investigate the putative candidate genes from the A03 and A08 QTL regions, we carried out PCA analysis of the DH population based on marker genotype data (Supplementary Figure S2.1). This analysis divided the population into four clusters. Cluster I included the DH lines

carrying the RP alleles of both A03 and A08; Cluster II included the lines carrying the RP allele of A03 but SP allele of A08; Cluster III included the lines carrying the RP allele of A08 but SP allele of A03; and Cluster IV included the lines carrying SP alleles of both A03 and A08.

Through bioinformatics analysis of the A03 and A08 QTL regions, we identified 13 candidate genes from A03 and 14 from A08 QTL regions. Of the total 27 genes, 13 encode proteins consisting of TIR/CC-NBS-LRR (TNL) or truncated TNL domain; six possess transport functions with multiple transmembrane regions; and the remaining eight genes were annotated to possess a stress-induced structure, such as zinc-finger, chitinase or KIN2 (Supplementary Table S2.4). All 27 genes were used for expression analysis on a subset of DH lines belonging to Cluster II, III and IV, and they were designated as Bulk-A03, Bulk-A08, and Bulk-Sus, respectively. Of the 27 genes, data of the six genes of A03 and six of A08 which differential expression which could be detected through qRT-PCR are presented in Figure 2.6, and data for the remaining genes are presented in Supplementary Figure S2.4.

None of the TNL or stress-induced genes of the A03 QTL were significantly up- or down-regulated in Bulk-A03 as compared to Bulk-A08 and Bulk-Sus. However, expression of two sugar transporter genes, viz. *BnaA03g29290D* and *BnaA03g29310D,* were significantly reduced at 7 and 14 DAI in Bulk-A03 as compared to Bulk-A08 and Bulk-Sus (Figure 2.6). These two genes are homoeologous (identity 97.95% and 97.26%) to the *sugar transport protein 6* (*STP-6*) genes of *B. rapa* and *B. napus*. Interestingly, of the remaining four genes from the A03 QTL, *BnaA03g28890D*, *BnaA03g27910D*, *BnaA03g29130D* and *BnaA03g29580D*, two genes (*BnaA03g28890D* and *BnaA03g27910D*) were upregulated at 7 DAI and two (*BnaA03g29130D* and *BnaA03g29580D*) were upregulated at 14 DAI in Bulk-A08. The gene *BnaA03g28890D* encodes a putative defensin-like protein with transmembrane domain, *BnaA03g27910D* is a stress-induced gene, while the other two are TNL genes. The level of upregulation of *BnaA03g28890D* in A08-Bulk was very high, 11.83-fold at 7 DAI when Bulk-A03 showed only 2.85-fold change and Bulk-Sus showed 1.88-fold change. Similarly, a 7.02-fold upregulation of the TNL gene *BnaA03g29580D* was found at 14 DAI in A08-Bulk when the level of upregulation of this gene in Bulk-A03 was only 1.83-fold and in Bulk-Sus it was 1.72-fold (Figure 2.6).

In case of the A08 QTL genes, *BnaA08g09220D* and *BnaA08g10090D* were upregulated at 7 DAI, while four genes, viz. *BnaA08g10100D*, *BnaA08g09330D*, *BnaA08g10470D* and

BnaA08g10540D, were upregulated at 14 DAI in Bulk-A08 as compared to Bulk-A03 and Bulk-Sus. Among the six genes, *BnaA08g09220D*, *BnaA08g10100D* and *BnaA08g10540D* are TNL genes, *BnaA08g09330D* and *BnaA08g10470D* are stress-induced genes, and *BnaA08g10090D* is a transport protein gene encoding Detoxification-39.

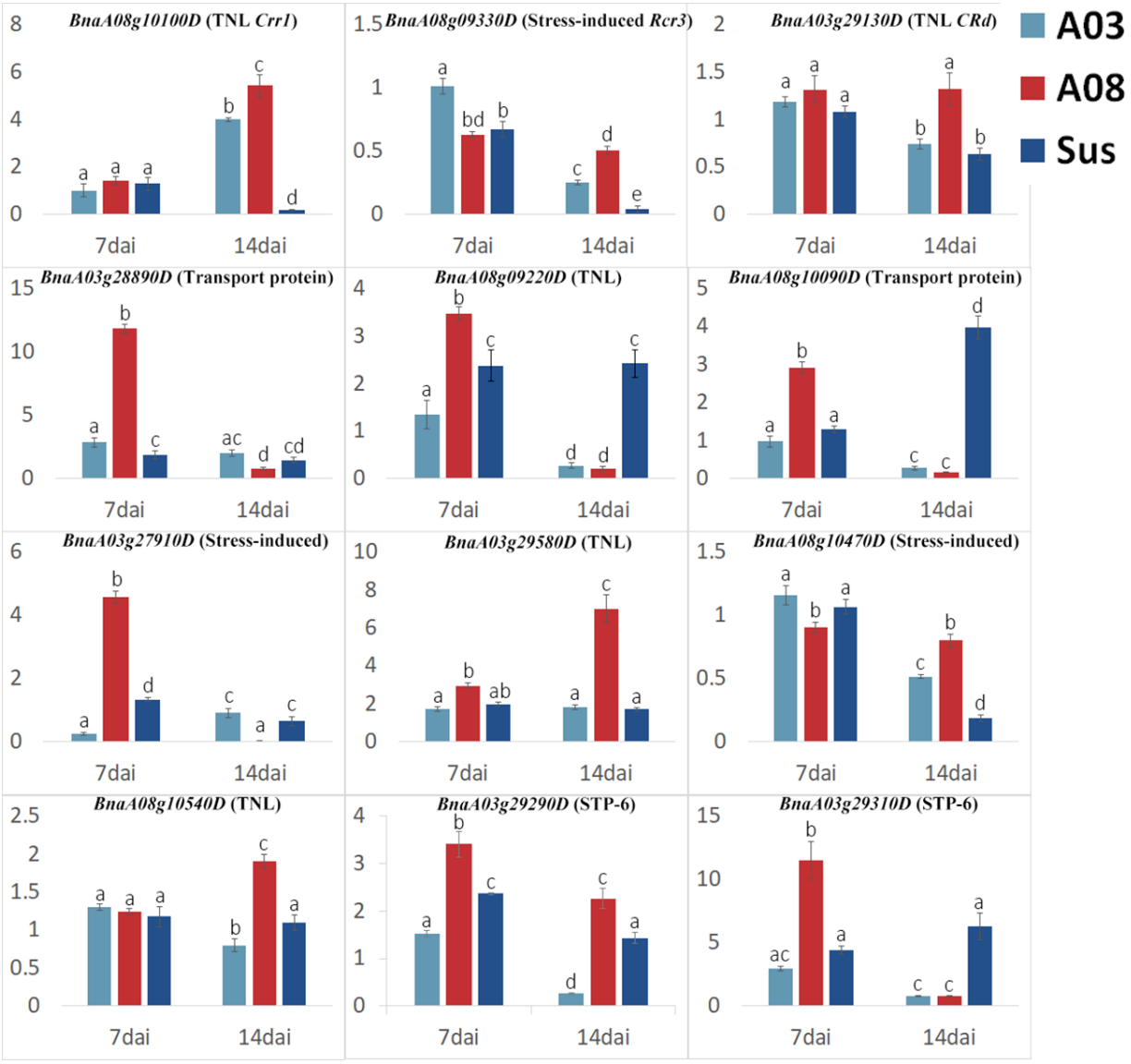


Figure 2.6 Expression analysis of 12 selected genes in roots of the *Brassica napus* doubled haploid (DH) lines carrying resistance on A03 (pale blue) or A08 (red), and the susceptible lines (dark blue) at 7 and 14 days after inoculation (dai). The name of the genes and the reasons for inclusion is shown at the top of figures. Expression of the mock-inoculated plants was set at 1 for each time point. Lines on the top of the bars show standard error, and the bars with same letters indicate that expression was not significantly different (p<0.05).

2.4 Discussion

In this study, by using a DH population carrying clubroot resistance of the rutabaga cv. Polycross, we identified two major loci on A03 (14.41-15.44 Mb of v3.1; 13.27-14.24 Mb of v4.1) and A08 (9.96-11.09 Mb of v3.1; 8.43-9.26 Mb of v4.1) conferring resistance to *P. brassicae* pathotypes 3A, 3D and 3H. By using clubroot resistance of the rutabaga cv. Brookfield, Hasan et al. (2021a) and Hasan and Rahman (2016) previously identified a major locus at 11.18-11.62 Mb (9.35-10.20 Mb of v4.1) position of A08 conferring resistance to pathotype 2, 3, 5, 6 and 8, and a small effect QTL at 15.59-16.10 Mb (14.29-14.82 Mb of v4.1) of A03 conferring resistance to pathotype 3 only. Recently, by using a *B. napus* DH population where resistance was derived from *B. rapa* turnip cv. Debra (ECD01), Yu et al. (2022) also reported a QTL at 9.8-13.5 Mb of A03 ($Rcr10^{ECD01}$) and a QTL at 9.3-11.5 Mb of A08 ($Rcr9^{ECD01}$) conferring resistance to pathotypes 3H, 3A, 3D and 5X. The A03 and A08 QTL that we detected in this study are located at the same genomic region as reported by Hasan et al. (2021a) and Yu et al. (2022); however, the A03 QTL that we detected in this study exerted a major effect on clubroot resistance and also showed a significant interaction with the A08 QTL with additive × additive effect of about 20% reduced DSI. This indicates that different allelic forms of the A03 locus exists, or multiple clubroot resistance genes are present in the same genomic region. The evidence for non-allelic interaction of the clubroot resistance was also evident from the analysis of the resistant DH lines (DSI <25%) partitioned into highly resistant and resistant groups and marker genotype of these groups. A variable effect of the two QTL on the level of resistance as well as on resistance to different pathotypes was observed indicating that these two loci were not behaving in a similar way in regards to resistance to these pathotypes. Expression analysis of the genes from these two QTL regions also provided evidence that non-Mendelian gene effects, such as trans-regulation of genes, is involved in clubroot resistance. Of the six genes of A03 exhibiting differential expression, four of them showed significantly greater expression in Bulk-A08 carrying clubroot resistance at A08 locus as compared to their expression in Bulk-A03. Li et al. (2013) also provided evidence that some of the QTL transcripts can act as a trans-acting regulator for the expression of other genes.

To date, several clubroot resistance loci have been identified on seven of the ten A genome chromosomes, where majority of the loci or genes are located on A03 and A08. At least eight loci, viz. *CRa/CRb*, *CRd*, *CRk*, *Rcr1*, *Rcr2*, *Rcr4*, *Rcr5* and *Crr3*, have been reported on

A03 (Chu et al., 2014; Hatakeyama et al., 2017; Hirai et al., 2004; Huang et al., 2017, 2019; Pang et al., 2018; Piao et al., 2004; Saito et al., 2006; Sakamoto et al., 2008; Ueno et al., 2012; Yu et al., 2016, 2017; for review, see Hasan et al., 2021c), where five of them are located about 8 Mb downstream of the A03 QTL that we detected in this study, while the loci *Crr3*, *CRk* and *CRd* are observed to reside in the same genomic region. The flanking or co-segregating markers of *Crr3* and *CRk* were either monomorphic or did not produce any amplification product in our mapping population However, based on search of the BrSTS-78 and BrSTS-33 marker sequences of *Crr3* (Hirai et al., 2004; Saito et al., 2006) in the reference genomes v3.1 and v4.1 using the software SnapGene, we could position this locus at 15.44-15.68 Mb of A03 (14.20-14.37 Mb of v4.1). Previously, the *CRk* has been reported in the same region of *Crr3* (Werner et al., 2008); our aforementioned analysis is consistent with that report. The marker yau376 of *CRd* mapped in the A03 QTL region (Figure 2.3); sequence of this marker as well as yau389 from the same locus could be positioned at 15.39-15.44 Mb (14.16-14.20 Mb of v4.1) region. Thus, these three loci found to be overlapping with the A03 QTL that we detected in the present study; therefore, it is possible that all are the same locus or multiple QTL is located in this genomic region.

Five clubroot resistance loci were detected on A08 and this includes *Rcr3*, *qBrCR38*, *Crr1*, *Rcr9* and *CRs* (Hasan & Rahman, 2016; Hatakeyama et al., 2013; Karim et al., 2020; Laila et al., 2019; Zhu et al., 2019). Among them, *qBrCR38* is located at least 10 Mb downstream of the A08 QTL that we identified in this study; the remaining four loci were observed to be located in the same genomic region of the A08 QTL. However, none of the published markers from these loci were polymorphic in our DH population. Following a similar approach, as mentioned above, we could localize the *Rcr9* based on its flanking markers M28 and M79 at 11.33-11.71 Mb (9.49-9.90 Mb of v4.1), *Rcr3* based on its flanking markers M12 and M16 at 10.45-10.70 Mb (8.70-8.95 Mb of v4.1), and *CRs* based on the co-segregating markers A08_10754563 and A08_11505101 at 10.40-11.19 Mb (8.68-9.37 Mb of v4.1) positions of A08. The resistance gene *Bra020861* identified in the *Crr1* region (Hatakeyama et al., 2013) is orthologous to *BnaA08g10100D* that we identified in the A08 QTL region. Thus, it is highly likely that the A08 QTL that we detected in the present study is the same as *Crr1* or *Rcr3* or *Rcr9*, or all three are the same locus. However, based on non-allelic interaction analysis, it appears that more than one gene is located in the A08 QTL region. This locus alone does not exert significant effect on

clubroot resistance; however, this in interaction with A03 QTL contribute to the resistance to pathotypes 3H and 3D (Table 2.3).

The two QTL of A03 and A08 could be Identified based on ΔSNP-index from RB vs. SB and SB-3D vs. SB comparisons; the SB-3H vs. SB comparison disclosed only the A08 QTL (Figure 2.2). However, the A03 QTL for resistance to 3H could be detected based on QTL mapping and using the whole population (Figure 2.3). The SB-3H bulk was partially resistant to pathotype 3H (DSI 51.2%; Supplementary Table S2.2), but the mapping population used for QTL mapping included all DH lines with DSI ranging from 0.0 to 100%. The lack of identification of the A03 QTL based on SB-3H and SB comparison indicates that the resistance allele of A03 QTL is essentially needed for high resistance to pathotype 3H. This was also evident from comparison of the highly resistant (DSI <5%) and resistant (DSI 5-25%) lines for marker genotypes, where 93% (50+43) of the highly resistant lines carried the A03 resistance allele while only 49% resistant lines carried this allele (Figure 2.5).

We carried out WGRS by using two reference genomes (v3.1 and v4.1) to ensure the accuracy of the sequencing results and QTL positions. The Brassica_napus_v4.1 (v4.1) is the sequence of the French winter *B. napus* cv. Darmor-bzh (Chalhoub et al. 2014); this is one of the most complete and well-annotated published reference genomes. The second reference genome is of the DH line DH12075 derived from cross between the French spring-type cv. Cresor and Canadian spring-type cv. Westar (Brassica_napus_v3.1); this sequence has not been published yet. Using these two reference genomes we obtained comparable results demonstrating the reliability of our findings. Among the two genomes, the *B. napus* cv. Darmor-bzh genome is well annotated; therefore, we used this for the development of markers linked to resistance, as well as for bioinformatics analysis for the identification of putative candidate genes from the QTL regions. Based on the SNPs identified through WGRS to be associated with resistance, we developed AS PCR markers for use in agarose gel-based genotyping. The PCR-based markers are user-friendly and cost-effective for genotyping large breeding populations (Lefever et al. 2019).

Of the three bulks, Bulk-A03, Bulk-A08 and Bulk-Sus, which we used for expression analysis of the genes from the QTL regions, Bulk-A03 carried resistance allele only of the A03 QTL, Bulk-A08 carried resistance allele only of the A08 QTL, and Bulk-Sus lack resistance alleles of both QTL. In this regard, expression analysis of the genes in Bulk-A03 and Bulk-A08

relative to Bulk-Sus was expected to confirm the genes involved in clubroot resistance. In case of Bulk-A03, no differential expression was found for the A03 QTL genes related to effector-triggered immunity (ETI) — especially the TNL genes. However, the expression of two sugar transport genes from this QTL was significantly downregulated in this bulk. Sucrose is mainly synthesized in leaves (source organs); transportation of sugar to infection site or pathogen is needed for gall formation (Li et al., 2018; Walerowski et al., 2018). Several sugar transport genes, such as *sucrose transporters* (SUC) *SUC2/SUC1*, *sugars will eventually be exported transporters* (*SWEET*) *SWEET11/SWEET12*, and *sugar transporter proteins* (STP) *STP4b/STP12*, have been reported to play an essential role in gall development (Li et al. 2018; Walerowski et al., 2018; Zhang et al., 2019a; Zhang et al., 2019b; for review, see Malinowski et al., 2019). In this regard, the downregulation of the *STP-6* like genes, *BnaA03g29290D* and *BnaA03g29310D*, as observed in this study (Figure 2.6) appear to contribute to the clubroot resistance conferred by the A03 QTL; however, functional validation of these genes, e.g. through *Arabidopsis* transformation, will be needed to further confirm this.

Results from expression analysis of the genes from the A08 QTL largely agree with our current knowledge of clubroot resistance. Expression of three TNL genes of A08, viz. *BnaA08g10100D, BnaA08g09220D* and *BnaA08g10540D*, were significantly upregulated at either 7 or 14 dai in Bulk-A08 (Figure 2.6). The TNL genes are known to detect or bind to the pathogen effectors and trigger defence response in ETI (for review, see Jones and Dangl 2006 and Hasan et al. 2021c). Among the three TNL genes, *BnaA08g10100D* is the orthologous of *Bra020861* (87.3% identity) located in *Crr1* (Hatakeyama et al., 2013); however, the involvement of the other two genes, *i.e., BnaA08g09220D* and *BnaA08g10540D* or their homologous genes, in clubroot resistances has not been reported previously. The stress-induced gene *BnaA08g09330D*, upregulated at 14 dai (Figure 2.6), is an orthologous of *Bra020951* (98.6% identity) located in *Rcr3* (Karim et al., 2020); this gene is annotated as glycosyl hydrolase family protein with chitinase insertion domain. Chitinase is known to be involved in Pattern-triggered immunity (PTI) and ETI; the ETI and PTI play a crucial role in clubroot resistance (Chandrashekar et al., 2015; Chen et al., 2018). The other stress-induced gene *BnaA08g10470D*, annotated as pentatricopeptide repeat-containing protein (PPR), and the transport protein gene *BnaA08g10090D*, annotated as detoxification protein, may play a role in clubroot resistance; however, detailed information of these two genes cannot be found in

literature. Two stress-induced genes, one of A03 QTL (*BnaA03g28240D*) and one of A08 QTL (*BnaA08g10610D*), showed significantly greater expression in Bulk-Sus as compared to the two bulks carrying resistance gene (Supplementary Figure S2.4), which apparently resulted from the stress induced in the plants due to pathogen infection. Knowledge of this type of genes could be important for developing a clubroot resistant canola following other approaches, such as gene editing or antisense transformation.

Two TNL genes of the A03 QTL, viz. *BnaA03g29130D* and *BnaA03g29580D*, were found upregulated in the Bulk-A08. Among them, *BnaA03g29130D* is an orthologue of the clubroot resistance gene *Bra001160* (80.8% identity) located in *CRd* (Pang et al., 2018); however, involvement of *BnaA03g29580D* or its homologs in clubroot resistances has not been reported previously. In addition to this, a stress-induced protein KIN2 gene (*BnaA03g27910D*), which is involved in abscisic acid signal transduction pathway, and thus, related to ETI (for review, see Asselbergh et al., 2008; Ton et al., 2009), and a putative defensin-like protein with transmembrane domain gene (*BnaA03g28890D*) of A03 QTL were also upregulated in Bulk-A08 at 7 DAI. Expression of A03 QTL genes in Bulk-A08, which lacking the resistance allele of the A03 QTL, demonstrates that epistatic effect play a role in clubroot resistance in Brassica.

It has been generally accepted that plant innate immunity, PTI and ETI, relies on signals from the infection site (for review, see Jones and Dangl 2006). TNL genes are usually considered as disease resistance or 'R' genes as they monitor the pathogen effector and activate the ETI. Expression of five TNL genes was upregulated in Bulk-A08; however, three of them were from the A08 QTL while two from A03 QTL. According to the "iceberg model" of plant effector-triggered immunity proposed by Thordal-Christensen (2020), some of the TNL genes can become incapable of activating the ETI due to suppression by pathogen effectors. Based on this, one or multiple copies of the three TNL genes of A08 QTL, viz. *BnaA08g10100D, BnaA08g09220D* and *BnaA08g10540D*, might have been involved in activation of the ETI, while the expression of both TNL genes of A03 QTL, viz. *BnaA03g29130D* and *BnaA03g29580D*, might have been inhibited by the suppressor in Bulk-A03. Several other researchers also found that the upregulated TNL genes are not limited to the QTL region only. For example, the CR QTL of the *B. napus* cv. ZHE-226 is located at about 8.43-13.33 Mb region of A08 (Zhan et al., 2015); however, transcriptome analysis showed that several TNL genes from almost all 19 chromosomes were upregulated in ZHE-226 (Mei et al., 2019). This indicate that a complex

interaction of genes regulates clubroot resistance in Brassica, and this might be another reason for increased expression of the TNL genes of A03 QTL in Bulk-A08.

In summary, the results from this study demonstrated that two major QTL located at 14.41-15.44 Mb of A03 and at 9.96-11.09 Mb of A08 chromosomes and their interaction confer resistance to *P. brassicae* pathotypes 3H, 3A and 3D in the rutabaga cv. Polycross which has been introgressed into *B. napus* canola. This has been confirmed not only by whole-genome resequencing of the resistant and susceptible lines and QTL mapping of the traits using a DH population, but also through expression analysis of the genes from the two QTL regions. Based on expression analysis, it was also evident that downregulation of the *STP-6* like genes *BnaA03g29290D* and *BnaA03g29310D* play an important role in the resistance conferred by the A03 QTL, while upregulation of the TNL genes *BnaA08g10100D, BnaA08g09220D* and *BnaA08g10540D* could be the major determinant of the resistance conferred by the A08 QTL. The SNP allele-specific PCR-based markers that we developed in this study could be used in breeding for the development of clubroot resistant cultivars including pyramiding of the resistance genes; these markers can be detected by agarose gel electrophoresis, and thus, are user-friendly.

Chapter 3. Characterization of the putative clubroot resistance genes of rutabaga by gain-of-function analysis in *Arabidopsis thaliana*

3.1 Introduction

The *Brassica napus* ssp. *napobrassica* rutabaga cvs. Brookfield and Polycross were bred and developed in Newfoundland (Spaner, 2002); therefore, it is expected that they share pedigree relationships. The cv. Brookfield carries resistance to *Plasmodiophora brassicae* pathotypes 2, 3, 5, 6 and 8 and this resistance has been introgressed into *Brassica napus* canola (Hasan & Rahman, 2016). Molecular mapping of this resistance identified a major locus at 9.35-10.20 Mb on chromosome A08 conferring resistance to the above-mentioned five pathotypes, and a small effect QTL at 14.29-14.82 Mb of A03 conferring resistance to pathotype 3 (Hasan et al., 2021a).

Summanwar et al. (2019, 2021) carried out transcriptome analysis of the clubroot resistant *B. napus* lines carrying the resistance of the cv. Brookfield to understand the molecular basis of this resistance. This study identified eight long noncoding RNAs (lncRNAs) from the chromosome A08 which expression could be detected in resistant plants but not in the susceptible plants. Among them, the *LNC_001163* located at 12,360,844 - 12,364,795 bp position of A08 was hypothesized to play a role in clubroot resistance. Based on sequence of this lncRNA and its secondary structure, it was annotated as a putative endogenous target mimic for a microRNA *bna-miR824* that located at the upstream of the *LNC_001163*, at about 12.36 Mb position of A08 (Summanwar et al., 2019). The *bna-miR824* has been reported to inhibit the expression of *Agamous-like 16* (*AGL16*) gene (Kutter et al., 2007; Xu et al., 2012); it is probable that the *LNC_001163* binds with *bna-miR824* and prevents the inhibiting of *AGL16*. In this regard, an increasing expression of *LNC_001163* would result an increasing expression of *AGL16*, and this increased expression of *AGL16* results increasing incidence of stomata in plants (Kutter et al., 2007). Furthermore, *bna-miR824* has been reported to be a stress responsive gene that can be upregulated by biotic and abiotic stress (Jian et al., 2018; Szaker et al., 2019). Summanwar et al. (2019) also found three lncRNAs that do not locate but target chromosome A08, which only expressed in the resistant plants. Among them, *LNC_003848* locates at 25,154,652 - 25,156,299 bp on chromosome C08, which targets and shows a positive expression correlation to gene *BnaA08g06670D*. The gene *BnaA08g06670D* is located at 6.67 Mb position of A08 and encodes *Transparent Testa 12* (*TT12*) protein (Summanwar et al., 2019); while the

TT12 protein is involved in flavonoid and plant hormone biosynthesis (Peer et al., 2001). Flavonoids are known to play an important role in protecting the plants from biotic and abiotic stresses (for review, see Ferreyra et al., 2012 and Treutter, 2005). In this regard, it can be hypothesized that the *LNC_003848* play a role in resistance to *P. brassicae* through regulating the flavonoid biosynthesis pathway.

Through a comparative analysis of transcriptomes of the *B. napus* lines carrying clubroot resistance of the rutabaga cv. Brookfield and lines carrying clubroot resistance of the European winter canola cv. Mendel, which carry clubroot resistance gene at 22.87-23.23 Mb position of the chromosome A03 (Fredua-Agyeman & Rahman, 2016), Summanwar et al. (2021) identified the mRNA gene *BnaA08g13940D* in the lines carrying Brookfield-resistance and *BnaA03g48910D* in the lines carrying Mendel-resistance. These two genes encode *cytokinin response factor 4* (*CRF4*) protein and found to be upregulated in resistant plants and located around the genomic regions carrying clubroot resistance region. The *CRF* family proteins known to be involved in plant development and response to abiotic stress; in particular, *CRF4* is involved in the response of cold and nitrogen (for review, see Hallmark & Rashotte, 2019). Based on this, it was hypothesized that the *BnaA08g13940D* gene, which is located at 11,929,169 - 11,930,155 bp of A08 might play a fundamental role in *P. brassicae* resistance (Summanwar et al., 2021). This research also developed several SSR markers based on differentially expressed genes, where the marker developed based on *BnaA08g03250D* showed genetic linkage association with clubroot resistance; this gene encodes dihydroxyacetone kinase and is located at 2,668,059 - 2,671,885 bp of A08.

The objective of this research was to investigate the role of the above-mentioned putative genes, *LNC_001163, LNC_003848, BnaA08g13940D* and *BnaA08g03250D*, for resistance to clubroot disease through over-expression in *A. thaliana*.

3.2 Materials and Methods

3.2.1 Plant material and growth conditions

Genomic DNA of the clubroot resistant *B. napus* var. *napobrassica* rutabaga cv. Brookfield (Spaner, 2002) was used as the template to amplify the targeted genes, and the wild-type *Arabidopsis thaliana* ecotype Columbia (Col-0) was used for overexpression of these genes. The seeds of wild-type Col-0 and the transformed lines were surface sterilized with 70% ethanol and

0.8% bleach and rinsed eight times with sterile water and were sown on MS medium (2.25 g/L Murashige and Skoog medium + 0.8% Phytoblend) with 1.5% sucrose (50 mg/L); the selection markers Kanamycin was added to the growth media when transgenic seeds were grown. After plating, the seeds were treated at 4 ℃ in dark for 3 days and the plates were transferred to a growth chamber set at 24 ℃ with 6h day/ 8h night photoperiod with light intensity of 200 μmol m^{-2} s^{-1}. After one to two weeks, seedlings were transplanted to 4 × 8 cells trays (cell size: 6 cm × 6 cm × 7.5 cm, L × W × D) filled with Sunshine Professional growing mix (Sun Gro Horticulture Canada Ltd, Seba Beach, Canada).

3.2.2 DNA extraction and cDNA synthesis

Plasmids with or without putative clubroot resistance genes were extracted from *Escherichia coli*with QIAprep® Spin Miniprep Kit (QIAGEN), and plant DNA was extracted with Phire Plant Direct PCR Kit (Thermo Fisher Scientific); RNA of the plants was extracted with RNeasy® Plant Mini Kit (QIAGEN). Following the manufacturer's instructions, the cDNA was synthesized using RevertAid First Strand cDNA Synthesis Kit (ThermoFisher Scientific).

3.2.3 Cloned primer design for putative clubroot resistant genes

To clone the correct sequences of lncRNAs and mRNAs, the cloned primers needed to flank the full-length sequence of genes; therefore, the forward and reverse primers were starting from the first and last bp of the genes. The sequences of *LNC_001163* and *LNC_003848* were provided by Dr. Summanwar, based on her RNA-seq result (Summanwar et al., 2019). The sequences of *BnaA08g03250D* and *BnaA08g13940D* were from the DH12075_v3.1 sequence (Chalhoub et al., 2014) in Genoscope Brassica Napus Genome Browser. In addition, all physical positions noted in this chapter is based on this genome reference (Chalhoub et al., 2014). To avoid superfluous region, the reference sequences were trimmed to the longest open reading frame, which were figured using NCBI ORF finder (https://www.ncbi.nlm.nih.gov/orffinder/). As a result, the forward and reverse primers were starting from the start and stop codon.

3.2.4 Full-length sequence isolation and alignment with reference genome

The full-length sequences of the long noncoding RNAs, *LNC_001163*, *LNC_003848*, and the

coding sequence of the mRNAs of *BnaA08g03250D* and *BnaA08g13940D* were amplified from the cDNAs of the rutabaga cv. Brookfield using the primers listed in Supplementary Table S3.1. In addition to this, the longest open reading frames of these mRNA were obtained using Open Reading Frame Finder (https://www.ncbi.nlm.nih.gov/orffinder/). The PCR amplified insertion sequences were aligned with their reference sequence (mentioned above), using Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/). While all physical positions noted in this chapter have been re-localized based on reference genomes (Chalhoub et al., 2014).

3.2.5 Vector construction and plant transformation

The correct insertion sequences were cloned at eight bp downstream of the Cauliflower Mosaic Virus 35S (*CaMV35S*) promoter and upstream of the kanamycin resistance gene of the binary vector *pCM2300*. The binary vector was kindly provided by Dr. Enrico Scarpella, Department of Biological Science, University of Alberta. The constructs *pCM2300: CaMV35S:: LNC_001163*, *pCM2300: CaMV35S:: LNC_003848*, *pCM2300: CaMV35S:: BnaA08g03250D*, and *pCM2300: CaMV35S:: BnaA08g13940D* and the empty vector *pCM2300* were prepared and amplified in XL10-Gold* ultracompetent *Escherichia coli* cell (Agilent Technologies, Inc.). Then, these amplified constructs were introduced into *Agrobacterium tumefaciens* strain GV3101 by electroporation and were used to transform *A. thaliana* Col-0 by floral dip method (Clough & Bent, 1998).

3.2.6 Development of homozygous transgenic lines

The *A. thaliana* used for floral dipping was considered as T_0 generation. The seeds harvested from the T_0 plants after floral dipping were seeded on 0.5 × MS Kanamycin plate. A week after germination, seedlings with green cotyledon and leaves were moved to soil and they were considered as T_1 plants. Thirty-two seeds harvested from each T_1 generation plants were seeded on Kanamycin plates, where 24 T_2 seedlings were green and eight bleached (3:1 ratio); all 24 T_2 seedlings were moved to soil. The T_2 plants were self-pollinated and the harvested T_3 seeds were seeded on Kanamycin plates. The seedlings of the T_3 seedlings plates without any bleached seedlings were moved to soil. PCR test was carried out on this T_3 plants to ensure the success of transformation. The seeds of the self-pollinated T_3 plants were harvested as T_4 seeds and were

homozygous for the transgene.

3.2.7 Expression analysis of the transgenic lines by qRT-PCR

A total of 162 plants transgenic and wild type plants were grown in a growth chamber. This included 27 homozygous T_4 transgenic plants derived from three individual T_3 plants (9 × 3) of each of the five insertion lines (*LNC_001163*, *LNC_003848*, *BnaA08g03250D*, *BnaA08g13940D* and empty vector), and 27 plants of the wild type Col-0. Two weeks after germination, the first true leaf of each plant was collected for RNA extraction, which followed synthesis of cDNA and qRT-PCR. The qRT-PCR was carried out on a StepOne Plus real-time PCR system (Life Technologies, Burlington, Canada) using FASTSYBR Green mix from (ThermoFisher Scientific, USA). As endogenous control, a constitutively expressed housekeeping gene *glyceraldehyde-3-phosphate dehydrogenase* (*GAPDH*) was used. The relative expression level of each gene was calculated using the $2^{-\Delta\Delta Ct}$ method (Livak & Schmittgen, 2001). The fold of gene expression in the transgenic lines and in the vector control lines were calculated based on the expression of the wild type Col-0.

3.2.8 Evaluation of homozygous transgenic lines for clubroot resistance

Single spore derived isolate of *P. brassicae*, classified as pathotypes 3H based on CCD set (Strelkov et al., 2018), were used for inoculation. These isolates were obtained from Dr. Stephen Strelkov, Department of Agricultural, Food and Nutritional Sciences, University of Alberta, in the form of galls. Resting spore suspensions (inoculum) from the galls were prepared following the protocol described by Strelkov et al. (2007), and the concentration of the suspension was adjusted to 1×10^7 to 1×10^8 resting spores/mL. Forty-five homozygous T_4 plants derived from five individual T_3 plants (including those three used in expression analysis) (9 × 5 = 45) of each of the five insertion lines (*LNC_001163*, *LNC_003848*, *BnaA08g03250D*, *BnaA08g13940D*, vector control) and 45 plants of the wild type Col-0), i.e. a total of 270 plants (9 × 5 × 6) were seeded in a greenhouse (20-22/15 °C day/night, 16 h photoperiod, light intensity of 130 µmol m^{-2} s^{-1}) in 9 × 8 cells trays (cell size: 4 cm × 4 cm × 5 cm, L × W × D) filled with Sunshine Professional growing mix (Sun Gro Horticulture Canada Ltd) for resistance to pathotype 3H. The seedlings were inoculated two weeks after germination by pipetting 1 mL of spore suspension at

the base of the seedling, and the inoculation was repeated on the following day to ensure successful infection. The plants were scored for disease severity at 30 days after inoculation. For this, the roots were washed with tap water and examined for galls and were rated on a 0–3 scale, where 0 = no galls, 1 = a few small galls on the lateral roots, 2 = moderate galls, and 3 = severe galls (for review, see Ludwig-Müller et al., 2017) (Figure 3.1).

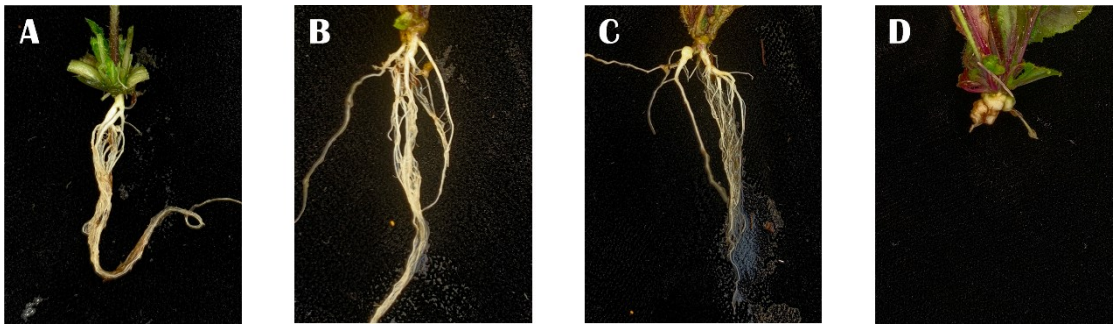


Figure 3.1 Clubroot disease severity symptom in *Arabidopsis thaliana* after inoculation with *Plasmodiophora brassicae* pathotype 3H. (A) Un-inoculated Col-0 root showing no disease symptom (disease score 0); roots are soft and carry no gall, (B) inoculated transgenic plant with disease score 1; a few small galls on the lateral roots, (C) inoculated transgenic plant with disease score 2; moderate galls on roots, (D) from inoculated Col-0 with disease score of 0; large galls on the roots.

3.3 Results

3.3.1 Vectors construction and development of transgenic lines

Due to the binary vector *pCM2300* lacking a robust constitutive promoter upstream of the multiple cloning sites (MCS), in the first step, a Cauliflower Mosaic Virus 35S (*CaMV35S*) promoter was inserted at the beginning of the MCS region. The length of this insertion was 802 bp, whose sequence was 100% (802/802) identical to the *CaMV35S* promoter sequence in NCBI (Supplementary Figure S3.1a). In the plasmid, the full transcriptive region of the lncRNAs *LNC_001163* or *LNC_003848* or the longest coding region of the *BnaA08g03250D* or *BnaA08g13940D* was cloned 8 bp downstream of the *CaMV35S* promoter. A total of four binary plasmids were constructed, which were *pCM2300: CaMV35S:: LNC_001163*, *pCM2300: CaMV35S:: LNC_003848*, *pCM2300: CaMV35S:: BnaA08g03250D* and *pCM2300: CaMV35S:: BnaA08g13940D*, respectively. Among them, the length of *LNC_001163* was 1976 bp; this carried a 73-bp insertion, a 2-bp deletion and a 1-bp variant, and thus had 96.3% (1903/1976) identity with the reference sequence in the *B. napus* genome (Chalhoub et al. 2014)

(Supplementary Figure S3.1b). The length of *LNC_003848* was 648 bp, which carried two 1-bp variants and had 99.7% (646/648) identity with the reference sequence (Supplementary Figure S3.1c). The length of gene *BnaA08g03250D* was 1785 bp, which carried three 1-bp variants and had 99.8% (1782/1785) identity with the reference sequence (Supplementary Figure S3.1d). The length of gene *BnaA08g13940D* was 987 bp, which carried two 1-bp variants and had 99.8% (985/987) identity with the reference sequence (Supplementary Figure S3.1e). Those variations existed in the genome of rutabaga cv. Brookfield; however, they did not cause nonsense or frameshift mutation. The genes *LNC_001163*, *LNC_003848*, *BnaA08g03250D* and *BnaA08g13940D* were cloned into binary vector *pCM2300* with *CaMV35S* promoter. Comparison of the above-mentioned lncRNA and gene sequences of rutabaga cv. Brookfield with the *B. napus* reference sequence (Chalhoub et al. 2014) showed that the variations found in rutabaga are not expected to cause nonsense or frameshift mutation.

The four gene constructs and the empty vector *pCM2300::CaMV35S* were introduced to *A. thaliana* Col-0, and homozygotic transgenic lines were developed. Genotyping T_3 plants confirmed the presence of these genes in the transgenic lines; however, not in the Col-0 wild-type line (Figure 3.2). A total of 36 (9 × 4) homozygous T_4 lines from four independent *LNC_001163* transgenic events, 27 (9 × 3) homozygous T_4 lines from three independent *LNC_003848* transgenic events, 36 (9 × 4) homozygous T_4 lines from four independent *BnaA08g03250D* transgenic events, nine (9 × 1) homozygous T_4 lines from one *BnaA08g13940D* transgenic event, and 18 (9 × 2) homozygous T_4 lines from two independent *pCM2300* vector control transgenic events were developed.

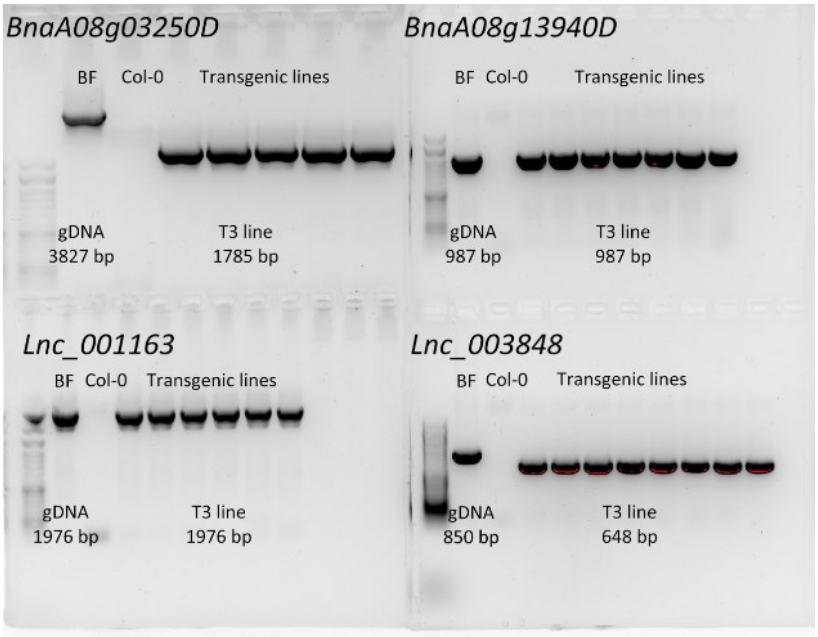


Figure 3.2 Genotyping of T_3 generation transgenic *Arabidopsis thaliana* lines using cloning primer. The gDNA of rutabaga cv. Brookfield was used as positive control and wild type col-0 as negative control. The band sizes are shown on the gel image.

3.3.2 Expression analysis of the genes in the transgenic plants

To verify that the genes were not only introduced into *A. thaliana* but they also expressed in the transgenic lines, expression analysis was carried out on three independent transgenic lines for each gene (Figure 3.3). The transgenic lines carrying *LNC_001163* showed 7567.0-, 3065.7- and 250369.2-folds greater expression as compared to wild-type; while the vector control lines showed only 36.5-, 0.8- and 10.2-folds greater expression. In case of *LNC_003848*, the transgenic lines showed 2927.5-, 1139.5- and 15839.2-folds greater expression as compared to wild-type; in this case, the vector control showed only 28.2-, 1.0- and 1.7-folds greater expression. A high-level expression was also for the transgenic plants carrying *BnaA08g03250D* and *BnaA08g13940D*. In case of *BnaA08g03250D*, the transgenic lines showed 196453.9-, 819688.0- and 75727.3-folds greater expression as compared to wild type, while vector control lines showed only 2.7-, 3.4- and 1.5-folds greater expression. Similarly, expression in transgenic *BnaA08g13940D* lines were 9226.4-, 208255.0- and 141994.4-folds greater than the wild type; in this case the vector control lines also showed almost no expression (1.8-, 0.7- and 0.0-folds).

3.3.3 Evaluation of the transgenic lines for clubroot resistance

All *A. thaliana* transgenic T_4 lines including the vector control and wild type were evaluated for resistance to *P. brassicae* pathotype 3H. The transgenic plants often received score of 1 or 2 (Figure 3.1B,C) while the wild type and vector control plants often received disease score of 2 or 3 (Figure 3.1C,D); The mean clubroot disease score of the transgenic lines carrying *LNC_001163*, *LNC_003848*, *BnaA08g03250D*, *BnaA08g13940D* and vector control were 1.89, 2.36, 2.07, 2.22 and 2.57 respectively. None of these mean values were significantly different from mean disease score of 2.46 of the wild type (Table 3.1, Figure 3.4).

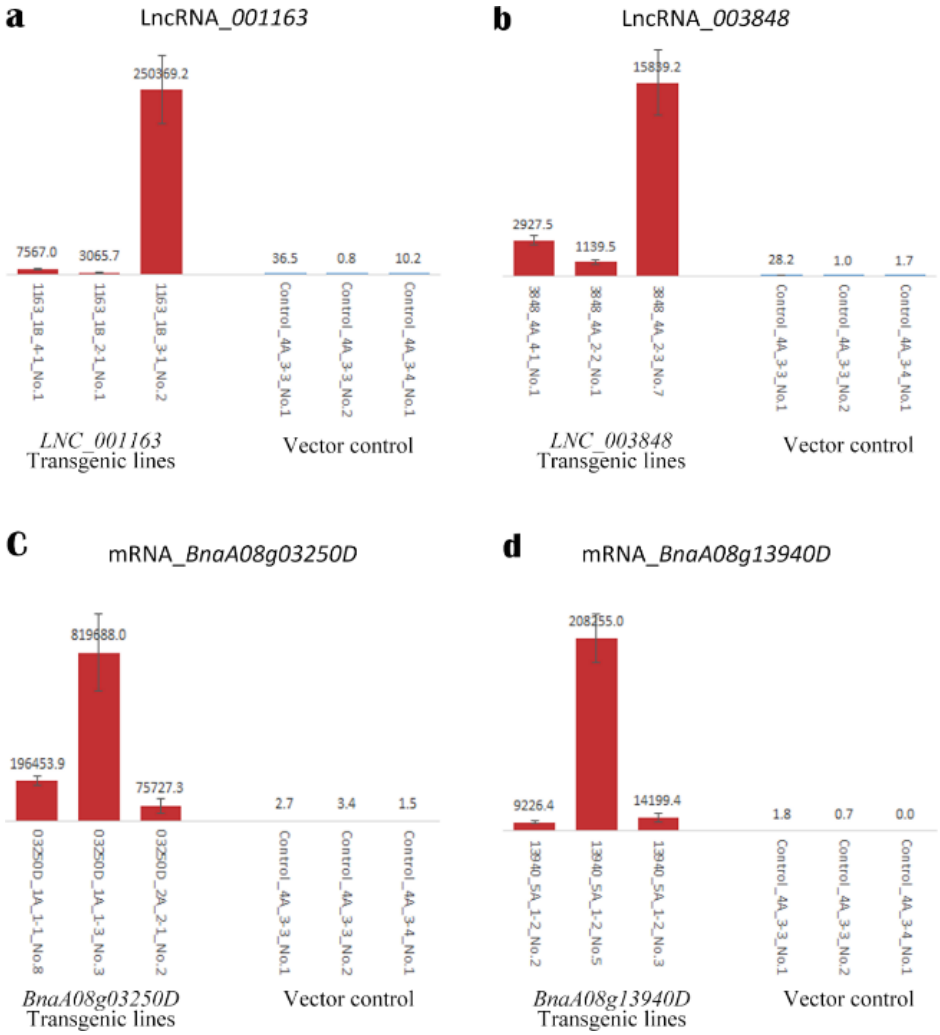


Figure 3.3 Expression of the two lncRNAs and two mRNA genes in transgenic *Arabidopsis thaliana* T_4 lines and in vector control line. Expression of in wild-type col-0 plants was set 1. Standard error bars shown on the histograms.

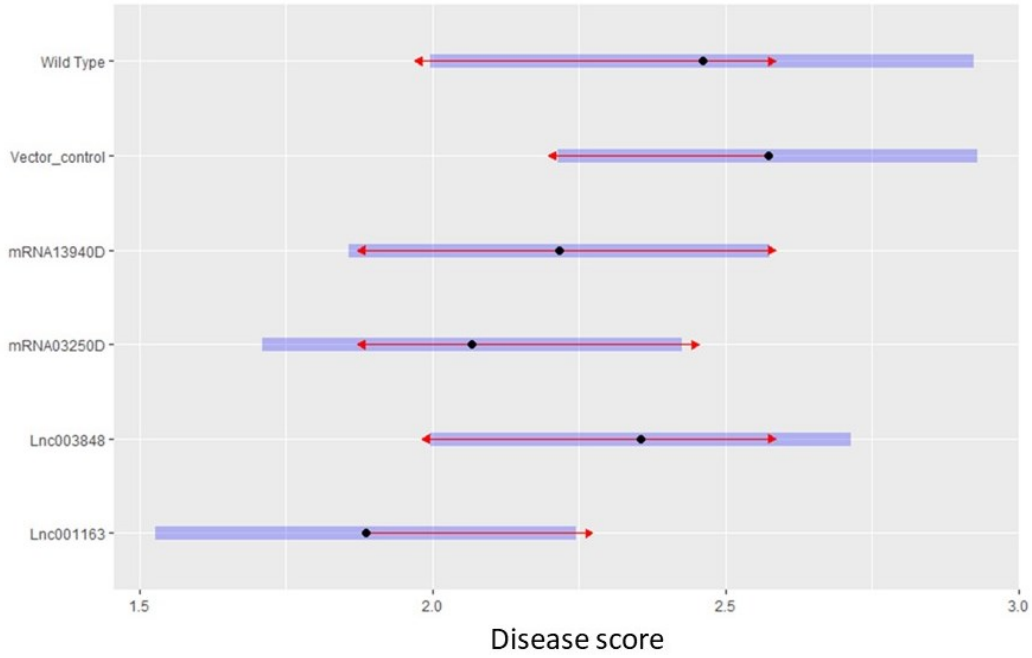


Figure 3.4 Comparisons of clubroot disease score between the transgenic *Arabidopsis thaliana* T_4 lines carrying lncRNAs and mRNA genes, and vector control and wild type lines obtained by inoculating the lines with *Plasmodiophora brassicae* pathotype 3H. The black spots show the mean value, the purple bars show the range with 99% confident interval, and the red arrows show standard error, which only visualize in the interval between smallest and largest mean disease score (1.89-2.57).

Table 3.1 Comparisons of clubroot disease score between the transgenic *Arabidopsis thaliana* T_4 lines carrying lncRNAs and mRNA genes, vector control to wildtype.
Clubroot disease was evaluated by inoculating the lines with *Plasmodiophora brassicae* pathotypes 3H. The *p*-values calculated with Tukey HSD test.

Line_1	Mean score	Line_2	Mean score	df	t.ratio	p.value
Lnc001163	1.89	Wild Type	2.46	22	-2.034	0.3562
Lnc003848	2.36	Wild Type	2.46	22	-0.368	0.999
mRNA03250D	2.07	Wild Type	2.46	22	-1.389	0.733
mRNA13940D	2.22	Wild Type	2.46	22	-0.864	0.9511
Vector Control	2.57	Wild Type	2.46	22	0.397	0.9985

3.4 Discussion

More than 24 loci involved in clubroot resistance have been reported in the A genome of *Brassica* (for review, see Hasan et al., 2021b); however, only one gene, *Bra020861* of the *Crr1* locus of *B. rapa* located on A08 chromosome has so far been characterized for its role on

clubroot resistance through transformation of *A. thaliana* (Hatakeyama et al., 2013). Through transcriptome analysis of the *B. napus* lines carrying clubroot resistance of the rutabaga cv. Brookfield, Summanwar et al. (2019, 2021) identified two lncRNAs (*LNC_001163* and *LNC_003848*) and two mRNAs (*BnaA08g13940D* and *BnaA08g03250D*) to be involved in clubroot resistance. The major clubroot resistance locus of the cv. Brookfield conferring resistance to *P. brassicae* located at 9.35-10.20 Mb position of the chromosome A08 (Hasan et al., 2021a). On the other hand, the *LNC_001163* and *BnaA08g13940D* are located at 12.36 Mb and 11.92 Mb of A08, *i.e*. adjacent to the clubroot resistance locus. The gene *BnaA08g03250D* and the target of *LNC_003848* are located at 2.67 Mb and 6.67 Mb on A08, *i.e.* upstream of the clubroot resistance locus. Thus, none of these four genes introduced into *Arabidopsis* are from 9.35-10.20 Mb region of A08 of *B. napus*. However, the increased expression of the two lncRNAs and two mRNAs detected by Summanwar et al. (2019, 2021) in the clubroot resistant lines carrying rutabaga-resistance indicate that they might play a role in the regulation of the clubroot resistance gene.

Homologous sequence of *LNC_001163, LNC_003848* and *BnaA08g13940D* could not be found in *A. thaliana*; therefore, they can be considered exogenous genes. The gene *BnaA08g03250D* showed sequence identity of 89% with *A. thaliana* gene *AT1G48430* encoding dihydroxyacetone kinase; therefore, *BnaA08g03250D* can be considered a homologous gene. However, the expression analysis strongly indicated that *LNC_001163, LNC_003848*, *BnaA08g13940D,* and *BnaA08g03250D* were significantly expressed in the transgenic lines.

Although all four lncRNA/genes showed strong expression in the transgenic plants, none of them showed resistance to *P. brassicae*. One of the probable reasons is that, the heterologous expression of the exogenous *B. napus* genes in *A. thaliana* may resulted loss of their function. Several researchers (Bertrand & Sorensen, 2019; for review, see Gomes et al., 2016) has indicated the importance of host system for correct expression of the transgene. The lncRNAs usually does not encode protein, but they modify the structure, expression level and transcriptional pattern of protein (for review, see Borah et al., 2018 and Sun et al., 2018); therefore, the function of lncRNA depends on other genes in the host. For *LNC_001163,* the functional region is a complementary sequence for binding with *bna-miR824* (Kutter et al., 2007). The mature sequence of *Bna-miR824* is "TAGACCATTTGTGAGAAGGGA" (Jian et al., 2018), which also exists in the *At-miR824* of *A. thaliana* (Supplementary Figure S3.2); therefore,

the *LNC_001163* could inhibit *At-miR824* in *A. thaliana*; just as inhibiting *Bna-miR824* in *B. napus*. In case of *LNC_003848*, the target gene of this lncRNA has not been reported in *A. thaliana*. The lack of significant effect of the above-mentioned four lncRNA/genes on clubroot resistance in the transgenic *Arabidopsis* lines also indicate that they might not be directly involved in clubroot resistance. The infection by *P. brassicae* hinder water and nutrient uptake from soil by damaged cells in the roots; therefore, differential expression of the four lncRNA/genes could also have resulted from abiotic stress. Based on bioinformatics, *LNC_001163*, *LNC_003848* and *BnaA08g13940D* have been found to be involved in abiotic and biotic stress (Jian et al., 2018; Szaker et al., 2019; Peer et al., 2001; for review, see Hallmark & Rashotte, 2019).

In conclusion, two lncRNAs (*LNC_001163* and *LNC_003848*) and two mRNAs (*BnaA08g13940D* and *BnaA08g03250D*) from the *B. napus* var. *napobrassica* cv. Brookfield has been heterologously overexpressed in *A. thaliana*; however, none of the homozygous transgenic lines showed resistance to *P. brassicae*. This indicates that these lncRNA/genes may not be directly involved in clubroot resistance. This study also provided evidence that a large majority of the genes identified through transcriptome analysis or similar other approaches may not be directly related to the trait of interest; validation of the differentially expressed genes is essentially needed to achieve the benefits from this type of 'omics' studies.

Chapter 4. Conclusion and general discussion

4.1 General discussion

Canola (*Brassica napus* L.) oil containing almost no erucic fatty acid, low saturated fatty acids and high omega 3 and 6 fatty acids (Przybylski et al., 2005) is recognized one of the best edible oils in the world. It is a non-cereal dicot crop and plays an important role in crop rotation (for review, see Friedt et al., 2018 and Harker et al., 2015). This is the second-largest oilseed crop in the world after soybean, and supply about 12% of the total edible oil in the world (USDA 2021). Among the different threats to *Brassica* crop production, clubroot disease, caused by *Plasmodiophora brassicae*, is one of the most serious one worldwide. This disease has been reported in all five continents in *Brassica* crops. In case of canola, significant yield loss has been reported in the European Union, Canada and China (Botero et al., 2019; Chai et al., 2014; Donald & Porter, 2014; for review, see Rempel et al., 2014) This disease is difficult to control by most cultural practices such as modification of soil pH, the addition of flusulfamide, chlorothalonil or quintozene, crop rotation, and growing bait crop (Cao et al., 2020; Kowata-Dresch & May-De, 2012; for review, see Hwang et al., 2014 and Peng et al., 2014) due to its long longevity of *P. brassicae* in soil (Wallenhammar, 1996) and its ability to infect both host and non-host plants (Ahmed et al., 2011; Friberg et al., 2005). Therefore, the development of clubroot-resistant cultivars has been considered the best strategy for management of this disease (for review, see Hwang et al. 2014).

The European turnip (*B. rapa* subsp. *rapifera*) has been reported to exhibit resistance to clubroot disease and carry multiple dominant genes (Hasan et al., 2012; for review, see Piao et al., 2009 and Hasan et al. 2021a). This resistance has been used to develop clubroot resistant Chinese cabbage (*B. rapa* subsp. *pekinensis*) as well as canola cultivars (for review, see Hasan et al., 2021a and Piao et al., 2009). Several rutabaga or Swede (*B. napus* var. *napobrassica*) accessions including some of the accessions from Maritime provinces of Canada, such as cvs. Brookfield and Polycross, has been reported to carry clubroot resistance (Spaner, 2002). Rutabaga and canola belong to the same species; therefore, this could be an ideal source of resistance for breeding clubroot resistant canola. Following traditional breeding, clubroot resistance from the rutabaga cv. Brookfield has been introgressed into spring canola (for review, see Rahman et al. 2014), where the resistant canola lines showed resistance to multiple

pathotypes, including 2, 3, 5, 6, and 8 (Hasan & Rahman, 2016). To date, several clubroot resistant canola cultivars exhibiting resistance to the above-mentioned pathotypes have been developed and commercialized in Canada. However, new pathotypes of *P. brassicae* can evolve and pathotype population structure can change, which can overcome the available resistances in canola (Strelkov et al. 2016). Therefore, it is important to identify new clubroot resistance sources and introgress the resistance into canola. The rutabaga cv. Polycross found to carry resistance to the recently evolved pathotypes, and this resistance has also been introgressed into canola by the University of Alberta Canola Program. For efficient and knowledge-based use of the resistances in breeding, it is important to understand the genetic and molecular basis of this. By using the resistance of the rutabaga cv. Brookfield, Hasan et al. (2021b) and Hasan & Rahman (2016) identified a minor effect locus at 14.29-14.82 Mb (v4.1) of A03 conferring resistance to pathotype 3, and a major locus at 9.35-10.20 Mb (v4.1) position of A08 conferring resistance to pathotype 2, 3, 5, 6 and 8. In this thesis research, I found the clubroot resistance loci of the rutabaga cv. Polycross are located at a similar position of A03 (13.27-14.24 Mb v4.1) and A08 (8.43-9.26 Mb v4.1); however, these two loci exhibited major effect and non-allelic interaction for resistance to the pathotypes 3H, 3A and 3D. My study also demonstrated that allelic variation for the A genome clubroot resistance loci can occur.

With the availability of genome sequence information and affordable high-throughput next-generation sequencing costs, it is possible to get a better understanding of the molecular basis of resistance as well as develop molecular makers for the resistance. Conventional QTL mapping requires a large population and genotyping the whole population with hundreds of markers; this is a labour-intensive and time-consuming task. In contrast, whole genome re-sequencing (WGRS) using bulk of lines exhibiting extreme phenotype has been found efficient for identification of QTL in a short period of time (Takagi et al., 2013). Another advantage of WGRS is that it can minimize the complexity associated with genotyping. The *Brassica* A genome exhibits homoeology with the C genome (Cheung et al., 2009); therefore, a marker can anneal at multiple homologous regions in *B. napus* generating ambiguous genotyping results. The flip side of the WGRS is that inclusion of a limited number of lines in bulk can generate multiple suggestive QTL (for review, see Nguyen et al., 2019). However, the WGRS can provide a few thousand SNPs from the QTL region, which can be used for genotyping a mapping

population for high-resolution mapping of the trait.

My thesis research contributed to the breeding of clubroot-resistant canola for durable resistance by using the clubroot resistance of the rutabaga cv. Polycross and Brookfield. Firstly, I identified two major resistance loci and developed SNP allele-specific PCR-based markers showing genetic linkage association with resistance. This result can be used not only in breeding canola cultivars carrying resistance of the rutabaga cv. Polycross, but also for pyramiding this resistance with other resistance through marker assisted breeding to increase the durability of resistance in canola cultivars (Wang et al., 2007; Zheng et al., 2022). Secondly, I identified two sugar transport protein encoding genes in loci on chromosome A03, which were downregulated in resistant lines. Three TNL genes and three stress-induced genes in loci on chromosome A08, which were upregulated in resistant lines. The downregulation of sugar transporter genes and upregulation of TNL genes were associated with clubroot resistance. This result makes breeding genetically modified clubroot resistant cultivars becomes possible.

Summanwar et al. (2019, 2021) as well as several researchers (e.g., Joshi et al., 2016; for review, see Zhou et al., 2021) carried out transcriptome analysis to understand host pathogen interaction and to identify the genes to be involved in clubroot resistance. However, confirmation of these genes through functional analysis is largely missing. I investigated two genes from primary and secondary metabolic pathways (*BnaA08g13940D* and *BnaA08g03250D*) and two lncRNAs (*LNC_001163, LNC_003848*) through overexpression in *A. thaliana*. These genes were not TNL or R genes; however, they showed significantly greater expression in the clubroot resistant lines carrying resistance of the rutabaga cv. Brookfield as compared to the susceptible lines (Summanwar et al. 2019, 2021). The mRNA *BnaA08g13940D* encodes *cytokinin response factor 4* protein which is involved in the response of cold and nitrogen stress (for review, see Hallmark & Rashotte, 2019), while *BnaA08g03250D* cosegregates with a SSR marker, which showed genetic linkage association with clubroot resistance. The lncRNA *LNC_001163* results an increasing expression of *AGL16* as well as increasing incidence of stomata in plants (Kutter et al., 2007), and *LNC_003848* is involved in the regulation of flavonoid and plant hormone biosynthesis (Peer et al., 2001). However, none of the homozygous transgenic *A. thaliana* lines carrying the above-mentioned genes/lncRNAs showed resistance to clubroot disease. It indicates these genes/lncRNAs might not be directly involved in clubroot resistance, but involved in other abiotic stress (e.g., drought, salinity) bringing with *P. brassicae* infection. It also provided

evidence that the genes identified through transcriptome analysis may not be directly related to the trait of interest; validation of the differentially expressed genes is essentially needed to achieve the benefits from this type of 'omics' studies.

4.2 Conclusion

The following conclusions can be drawn from the study using clubroot resistance of the rutabaga cv. Polycross:

- The genomic regions of 13.27-14.24 Mb of A03 and 8.43-9.26 Mb (v4.1) of A08 and their interactions confer resistance to *P. brassicae* pathotypes 3H, 3A and 3D.
- WGRS followed by QTL-seq and linkage map-based QTL mapping approach is an efficient way to study the genetic and molecular basis of the resistance.
- SNP allele-specific PCR-based markers showing strong linkage association with clubroot resistance can be developed for used in breeding.
- Downregulation of sugar transporter gene and upregulation of TNL gene is associated with clubroot resistance.
- The putative clubroot resistance genes (*BnaA08g13940D* and *BnaA08g03250D*) and lncRNAs (*LNC_001163* and *LNC_003848*) identified through transcriptome analysis may not be involved in clubroot resistance in *A. thaliana***.**

4.3 Future Prospects

- The clubroot resistance of the rutabaga cv. Polycross can be an excellent source of resistance for use in breeding; this resistance should be used to develop canola cultivars.
- The function of the gene found to be associated with susceptibility (*BnaA03g29290D* and *BnaA03g29310D*) and the TNL genes (*BnaA08g10100D, BnaA08g09220D* and *BnaA08g10540D*) identified by using clubroot resistance of the cv. Polycross need to be characterized through transformation of *A. thaliana*.
- By using the SNP allele-specific markers, the Polycross-resistance can be pyramided with other resistance for the development of canola cultivars for durable resistance to clubroot disease.
- The homozygous *A. thaliana* lines carrying a single gene or lncRNA can be crossed to

develop lines carrying multiple genes to study the effect of the accumulated genes on clubroot resistance.

- Transformation of *B. napus* using the putative clubroot resistance genes (*BnaA08g13940D* and *BnaA08g03250D*) and lncRNAs (*LNC_001163* and *LNC_003848*) will be needed to confirm their role in clubroot resistance.
- A comparative transcriptome analysis of the lines carrying clubroot resistance of the rutabaga cvs. Polycross and Brookfield could be carried to understand the genetic and molecular basis of clubroot resistance.

Bibliography

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., & Terauchi, R. (2012). Genome sequencing reveals agronomically important loci in rice using MutMap. *Nature Biotechnology*, *30*, 174–178. https://doi.org/10.1038/nbt.2095

Agnihotri, A., Prem, D., & Gupta, K. (2007). The chronicles of oil and meal quality improvement in oilseed rape. *Advances in Botanical Research*, *45*, 49–97. Academic Press. https://doi.org/10.1016/S0065-2296(07)45003-0

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

Ahmed, W., Xia, Y., Li, R., Bai, G., Siddique, K. H. M., & Guo, P. (2020). Non-coding RNAs: Functional roles in the regulation of stress response in *Brassica* crops. *Genomics*, *112*, 1419–1424. https://doi.org/10.1016/j.ygeno.2019.08.011

Alexopoulos, C. J., Mims, C. W., & Blackwell, M. (1996). *Introductory mycology.* (Issue Ed.4).

Alix, K., Lariagon, C., Delourme, R., & Manzanares-Dauleux, M. J. (2007). Exploiting natural genetic diversity and mutant resources of *Arabidopsis thaliana* to study the *A. thaliana*-*Plasmodiophora brassicae* interaction. *Plant Breeding*, *126*, 218–221. https://doi.org/10.1111/j.1439-0523.2007.01314.x

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

Asselbergh, B., de Vleesschauwer, D., & Höfte, M. (2008). Global switches and fine-tuning-ABA modulates plant pathogen defense. *Molecular Plant Microbe Interactions*, *21*, 709–719. https://doi.org/10.1094/MPMI-21-6-0709

Audic, S., & Claverie, J.-M. (1997). The significance of digital gene expression profiles. *Genome Research*, *7*, 986–995.

Axelsson, T., Bowman, C. M., Sharpe, A. G., Lydiate, D. J., & Lagercrantz, U. (2000). Amphidiploid *Brassica juncea* contains conserved progenitor genomes. *Genome*, *43*, 679–688. https://doi.org/10.1139/g00-026

Ayers, G. W. (1944). Studies on the life history of the club root organism, *Plasmodiophora brassicae*. *Canadian Journal of Research*, *22*, 143–149.

Ayers, G. W., & Lelacheur, K. E. (1972). Genetics of resistance in rutabaga to two races of *Plasmodiophora brassicae*. *Canadian Journal of Plant Science*, *52*, 897–900.

Barker, M. S., Vogel, H., & Schranz, M. E. (2009). Paleopolyploidy in the *Brassicales*: analyses of the Cleome transcriptome elucidate the history of genome duplications in *Arabidopsis* and other *Brassicales*. *Genome biology and evolution*, *1*, 391–399. https://doi.org/10.1093/gbe/evp040

Beaulieu, Y. B., Kleinman, C. L., Landry-Voyer, A. M., Majewski, J., & Bachand, F. (2012). Polyadenylation-dependent control of long noncoding rna expression by the poly(a)-binding protein nuclear. *PLoS Genetics*, *8*. https://doi.org/10.1371/journal.pgen.1003078

Bertrand, R. L., & Sorensen, J. L. (2019). Lost in translation: challenges with heterologous expression of lichen polyketide synthases. *ChemistrySelect*, *4*, 6473–6483.

Bell, C. D., Soltis, D. E., & Soltis, P. S. (2010). The age and diversification of the angiosperms re-revisited. *American journal of botany*, *97*, 1296–1303. https://doi.org/10.3732/ajb.0900346

Bevan, M., Mayer, K., White, O., Eisen, J. A., Preuss, D., Bureau, T., Salzberg, S. L., & Mewes, H. W. (2001). Sequence and analysis of the *Arabidopsis* genome. *Current Opinion in Plant Biology*, *4*, 105–110. https://doi.org/10.1016/S1369-5266(00)00144-8

Boivin, V., Faucher-Giguère, L., Scott, M., & Abou-Elela, S. (2019). The cellular landscape of mid-size noncoding RNA. *Wiley Interdisciplinary Reviews: RNA*, *10*, 1–17. https://doi.org/10.1002/wrna.1530

Borah, P., Das, A., Milner, M. J., Ali, A., Bentley, A. R., & Pandey, R. (2018). Long non-coding RNAs as endogenous target mimics and exploration of their role in low nutrient stress tolerance in plants. *Genes*, *9*, 1. – 17. https://doi.org/10.3390/genes9090459

Botero, A., García, C., Gossen, B. D., Strelkov, S. E., Todd, C. D., Bonham-Smith, P. C., & Pérez-López, E. (2019). Clubroot disease in Latin America: distribution and management strategies. *Plant Pathology*, *68*, 827–833. https://doi.org/10.1111/ppa.13013

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & 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

Bradshaw, J. E., Gemmell, D. J., & Wilson, R. N. (1997). Transfer of resistance to clubroot (*Plasmodiophora brassicae*) to swedes {*Brassica napus* L. var. *napobrassica peterm*) from *B. rapa*. *Annals of Applied Biology*, *130*, 337–348.

Brouwer, I. A. (2020). The public health rationale for reducing saturated fat intakes: Is a maximum of 10% energy intake a good recommendation? *Nutrition Bulletin*, *45*, 271–280. https://doi.org/10.1111/nbu.12449

Buczacki, S. T., Toxopeus, H., Mattusch, P., Johnston, T. D., Dixon, G. R., & Hobolth, L. A. (1975). Study of physiologic specialization in *Plasmodiophora brassicae*: Proposals for attempted rationalization through an international approach. *Transactions of the British Mycological Society*, *65*, 295–303. https://doi.org/10.1016/s0007-1536(75)80013-1

Canadian Plant Disease Survey 2020 Volume 100: Disease Highlights 2019. (2020). *Canadian Journal of Plant Pathology*, *42*, 1–175. https://doi.org/10.1080/07060661.2020.1752524

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

Caron, M., St-Onge, P., Drouin, S., Richer, C., Sontag, T., Busche, S., Bourque, G., Pastinen, T., & Sinnett, D. (2018). Very long intergenic non-coding RNA transcripts and expression profiles are associated to specific childhood acute *lymphoblastic leukemia* subtypes. *PLoS ONE*, *13*, 1–15. https://doi.org/10.1371/journal.pone.0207250

Chai, A. L., Xie, X. W., Shi, Y. X., & Li, B. J. (2014). Special Issue: Research status of clubroot (*Plasmodiophora brassicae*) on cruciferous crops in China. *Canadian Journal of Plant Pathology*, *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. X. X., Chiquet, J., Belcram, H., Tong, C., Samans, B., Corréa, M., da Silva, C., Just, J., Falentin, C., Koh, C. S., le Clainche, I., Bernard, M., Bento, P., Noel, B., … 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

Chandrashekar, N., Ali, S., Rawat, S., & Grover, A. (2015). Gene expression profiling of *Arabidopsis thaliana* chitinase genes in response to *Alternaria brassicae* challenge Gene expression profiling of *Arabidopsis thaliana* chitinase genes in response to *Alternaria brassicae* challenge. *Indian Phytopath*, *68*, 106–111.

Chang, A., Lamara, M., Wei, Y., Hu, H., Parkin, I. A. P., Gossen, B. D., Peng, G., & Yu, F. (2019). Clubroot resistance gene *Rcr6* in *Brassica nigra* resides in a genomic region homologous to chromosome A08 in *B. rapa*. *BMC Plant Biology*, *19*, 1–11. https://doi.org/10.1186/s12870-019-1844-5

Charlton, K. M., Corner, A. H., Davey, K., Kramer, J. K., Mahadevan, S., & Sauer, F. D. (1975). Cardiac lesions in rats fed rapeseed oils. *Canadian Journal of Comparative Medicine*, *39*, 261–269.

Chen, J., Pang, W., Chen, B., Zhang, C., & Piao, Z. (2016). Transcriptome analysis of *Brassica rapa* near-Isogenic lines carrying clubroot-resistant and -susceptible alleles in response to *Plasmodiophora Brassicae* during early infection. *Frontiers in Plant Science*, *6*, 1–14. https://doi.org/10.3389/fpls.2015.01183

Chen, J., Piao, Y., Liu, Y., Li, X., & Piao, Z. (2018). Genome-wide identification and expression analysis of chitinase gene family in *Brassica rapa* reveals its role in clubroot resistance. *Plant Science*, *270*, 257–267. https://doi.org/10.1016/j.plantsci.2018.02.017

Chen, S. W., Liu, T., Gao, Y., Zhang, C., Peng, S. D., Bai, M. B., Li, S. J., Xu, L., Zhou, X. Y., & Lin, L. B. (2016). Discovery of clubroot-resistant genes in *Brassica napus* by transcriptome sequencing. *Genetics and Molecular Research*, *15*, 1–19. https://doi.org/10.4238/gmr.15038243

Cheng, F., Liang, J., Cai, C., Cai, X., Wu, J., & Wang, X. (2017). Genome sequencing supports a multi-vertex model for *Brassiceae* species. *Current Opinion in Plant Biology*, *36*, 79–87. https://doi.org/10.1016/j.pbi.2017.01.006

Cheng, F., Lysak, M. A., Mandáková, T., & Wang, X. (2015a). *The Common Ancestral Genome of the Brassica Species BT - The Brassica rapa Genome.* Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-47901-8_8

Cheng, F., Mandáková, T., Wu, J., Xie, Q., Lysak, M. A., & Wang, X. (2013). Deciphering the diploid ancestral genome of the mesohexaploid *Brassica rapa*. *Plant Cell*, *25*, 1541–1554. https://doi.org/10.1105/tpc.113.110486

Cheng, F., Wu, J., Liu, B., & Wang, X. (2015b). *Genome evolution after whole genome triplication: the subgenome dominance in Brassica rapa BT - The Brassica rapa genome*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-47901-8_9

Cheung, F., Trick, M., Drou, N., Lim, Y. P., Park, J. Y., Kwon, S. J., Kim, J. A., Scott, R., Chris Pires, J., Paterson, A. H., Town, C., & Bancroft, I. (2009). Comparative analysis between homoeologous genome segments of *Brassica* napus and its progenitor species reveals extensive sequence-level divergence. *Plant Cell*, *21*, 1912–1928. https://doi.org/10.1105/tpc.108.060376

Christophersen, B. O., & Bremer, Jon. (1972). Erucic acid-An inhibitor of fatty acid oxidation in the heart. *Biochimica et Biophysica Acta*, *280*, 506–514.

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

Chu, M., Yu, F., Falk, K. C., Liu, X., Zhang, X., Chang, A., & Peng, G. (2013). Identification of the clubroot resistance gene *Rpb1* and introgression of the resistance into canola breeding lines using a marker-assisted approach. *Acta Horticulturae*, *1005*, 599–606. https://doi.org/10.17660/actahortic.2013.1005.74

Clough, S. J., & Bent, A. F. (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. *The Plant Journal*, *16*, 735–743.

Crespi, M. D., Jurkevitch, E., Poiret, M., D'Aubenton-Carafa, Y., Petrovics, G., Kondorosi, E., & Kondorosi, A. (1994). *Enod40*, a gene expressed during nodule organogenesis, codes for a non-translatable RNA involved in plant growth. *EMBO Journal*, *13*, 5099–5112. https://doi.org/10.1002/j.1460-2075.1994.tb06839.x

Crisp, P., Crute, I. R., Sutherland, R. A., Angell, S. M., Bloor, K., Burgess, H., & Gordon, P. L. (1989). The exploitation of genetic resources of *Brassica oleracea* in breeding for resistance to clubroot (*Plasmodiophora brassicae*). *Euphytica*, *42*, 215–226. https://doi.org/10.1007/BF00034457

Crute, I. R., Gray, A. R., Crisp, P., & Buczacki, S. T. (1980). Variation in *Plasmodiophora brassicae* and resistance to clubroot disease in *Brassicas* and allied crops-a critical review. *Plant Breeding Abstracts*, *50*, 91–104.

Cui, J., Luan, Y., Jiang, N., Bao, H., & Meng, J. (2017). Comparative transcriptome analysis between resistant and susceptible tomato allows the identification of *lncRNA16397* conferring resistance to *Phytophthora infestans* by co-expressing glutaredoxin. *Plant Journal*, *89*, 577–589. https://doi.org/10.1111/tpj.13408

Czubatka-Bieńkowska, A., Kaczmarek, J., Marzec-Schmidt, K., Nieróbca, A., Czajka, A., & Jędryczka, M. (2020). Country-wide qpcr based assessment of *Plasmodiophora brassicae* spread in agricultural soils and recommendations for the cultivation of *brassicaceae* crops in Poland. *Pathogens*, *9*, 1–17. https://doi.org/10.3390/pathogens9121070

Dakouri, A., Lamara, M., Karim, M. M., Wang, J., Chen, Q., Gossen, B. D., Strelkov, S. E., Hwang, S. F., Peng, G., & Yu, F. (2021). Identification of resistance loci against new pathotypes of *Plasmodiophora brassicae* in *Brassica napus* based on genome-wide association mapping. *Scientific Reports*, *11*, 1–11. https://doi.org/10.1038/s41598-021-85836-9

Dakouri, A., Zhang, X., Peng, G., Falk, K. C., Gossen, B. D., Strelkov, S. E., & 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*. *Scientific Reports*, *8*, 1–10. https://doi.org/10.1038/s41598-018-36187-5

Dangl, J. L., & Jones, J. D. G. (2001). Defence responses to infection. *Nature*, *411*.

De-Lorenzo, G., Brutus, A., Savatin, D. V., Sicilia, F., & Cervone, F. (2011). Engineering plant resistance by constructing chimeric receptors that recognize damage-associated molecular patterns (DAMPs). *FEBS Letters*, *585*, 1521–1528. https://doi.org/10.1016/j.febslet.2011.04.043

Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., Guernec, G., Martin, D., Merkel, A., Knowles, D. G., Lagarde, J., Veeravalli, L., Ruan, X., Ruan, Y., Lassmann, T., Carninci, P., Brown, J. B., Lipovich, L., Gonzalez, J. M., … Guigó, R. (2012). The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. *Genome Research*, *22*, 1775–1789. https://doi.org/10.1101/gr.132159.111

Dickman, M. B., & Mitra, A. (1992). *Arabidopsis thaliana* as a model for studying *Sclerotinia sclerotiorum* pathogenesis. *Physiological and Molecular Plant Pathology*, *41*, 255-263. https://doi.org/10.1016/0885-5765(92)90025-Q

Diederichsen, E., Frauen, M., Linders, E. G. A., Hatakeyama, K., & Hirai, M. (2009). Status and perspectives of clubroot resistance breeding in crucifer crops. *Journal of Plant Growth Regulation*, *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. *Journal of Plant Growth Regulation*, *28*, 194–202. https://doi.org/10.1007/s00344-009-9090-y

Dokken-Bouchard, F. L., Bouchard, A. J., Ippolito, J., Peng, G., Strelkov, S. E., Kirkham, C. L., & Kutcher, H. R. (2010). Detection of *Plasmodiophora brassicae* in Saskatchewan, 2008. *Canadian Plant Disease Survey/ Inventaire Des Maladies Des Plantes Au Canada*.

Donald, E. C., & Porter, I. J. (2014). Special Issue: Clubroot in Australia: The history and impact of *Plasmodiophora brassicae* in *Brassica* crops and research efforts directed towards its control. *Canadian Journal of Plant Pathology*, *36*, 66–84. https://doi.org/10.1080/07060661.2013.873482

Dunham, I., Kundaje, A., Aldred, S. F., Collins, P. J., Davis, C. A., Doyle, F., Epstein, C. B., Frietze, S., Harrow, J., Kaul, R., Khatun, J., Lajoie, B. R., Landt, S. G., Lee, B. K., Pauli, F., Rosenbloom, K. R., Sabo, P., Safi, A., Sanyal, A., … Lochovsky, L. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature*, *489*, 57–74. https://doi.org/10.1038/nature11247

Evans, J. L., & Scholes, J. D. (1995). How does clubroot alter the regulation of carbon metabolism in its host? *Aspects of Applied Biology*, *No. 42*, 125–132.

Fahey, J. W. (2015). *Brassica*: Characteristics and properties. *Encyclopedia of food and health* (3rd ed.). Elsevier Ltd. https://doi.org/10.1016/B978-0-12-384947-2.00083-0

Falcone Ferreyra, M. L., Rius, S. P., & Casati, P. (2012). Flavonoids: Biosynthesis, biological functions, and biotechnological applications. *Frontiers in Plant Science*, *3*, 1–15. https://doi.org/10.3389/fpls.2012.00222

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

Felix, G., Duran, J. D., Volko, S., & Boller, T. (1999). Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. *Plant Journal*, *18*, 265–276. https://doi.org/10.1046/j.1365-313X.1999.00265.x

Frauen, M. (1999). A new clubroot resistant variety in winter oilseed rape. *10th International Rapeseed Congress*, 1–3.

Fredua-Agyeman, R., & 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., & Feindel, D. (2019). Identification of *Brassica* accessions resistant to 'old' and 'new' pathotypes of *Plasmodiophora brassicae* from Canada. *Plant Pathology*, *68*, 708–718. https://doi.org/10.1111/ppa.12980

Fredua-Agyeman, R., Yu, Z., Hwang, S. F., & Strelkov, S. E. (2020). Genome-Wide mapping of loci associated with resistance to clubroot in *Brassica napus* ssp. *napobrassica* (Rutabaga) accessions from Nordic Countries. *Frontiers in Plant Science*, *11*, 1–17. https://doi.org/10.3389/fpls.2020.00742

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

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

Galindo-González, L., Manolii, V., Hwang, S. F., & Strelkov, S. E. (2020). Response of *Brassica napus* to *Plasmodiophora brassicae* Involves Salicylic Acid-Mediated Immunity: An RNA-Seq-Based Study. *Frontiers in Plant Science*, *11*, 1–20. https://doi.org/10.3389/fpls.2020.01025

Gomes, A. R., Byregowda, S. M., Veeregowda, B. M., & Balamurugan, V. (2016). *An overview of heterologous expression host systems for the production of recombinant proteins*.

Goodwin, S., McPherson, J. D., & McCombie, W. R. (2016). Coming of age: Ten years of next-generation sequencing technologies. *Nature Reviews Genetics*, *17*, 333–351. https://doi.org/10.1038/nrg.2016.49

Gu, Z., Guo, Q. & Gu, Y. (2012). Factors influencing glucoraphanin and sulforaphane formation in *Brassica* plants: A review. *Journal of Integrative Agriculture*, *11*, 1804–1816. https://doi.org/10.1016/S2095-3119(12)60185-3

Gupta, S. K. (2016). *Brassicas. Breeding Oilseed Crops for Sustainable Production*. https://doi.org/10.1016/B978-0-12-801309-0/00003-3

Gust, A. A., Biswas, R., Lenz, H. D., Rauhut, T., Ranf, S., Kemmerling, B., Götz, F., Glawischnig, E., Lee, J., Felix, G., & Nürnberger, T. (2007). Bacteria-derived peptidoglycans constitute pathogen-associated molecular patterns triggering innate immunity in *Arabidopsis*. *Journal of Biological Chemistry*, *282*, 32338–32348. https://doi.org/10.1074/jbc.M704886200

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

Hallmark, H. T., & Rashotte, A. M. (2019). Review – Cytokinin Response Factors: Responding to more than cytokinin. *Plant Science*, *289*, 110251. https://doi.org/10.1016/j.plantsci.2019.110251

Handa, H. (2003). The complete nucleotide sequence and RNA editing content of the mitochondrial genome of rapeseed (*Brassica napus* L.): Comparative analysis of the mitochondrial genomes of rapeseed and *Arabidopsis thaliana*. *Nucleic Acids Research*, *31*, 5907–5916. https://doi.org/10.1093/nar/gkg795

Hanriot, L., Keime, C., Gay, N., Faure, C., Dossat, C., Wincker, P., Scoté-Blachon, C., Peyron, C., & Gandrillon, O. (2008). A combination of LongSAGE with Solexa sequencing is well suited to explore the depth and the complexity of transcriptome. *BMC Genomics*, *9*, 1–9. https://doi.org/10.1186/1471-2164-9-418

Harker, K. N., O'Donovan, J. T., Turkington, T. K., Blackshaw, R. E., Lupwayi, N. Z., Smith, E. G., Johnson, E. N., Gan, Y., Kutcher, H. R., Dosdall, L. M., & Peng, G. (2015). Canola rotation frequency impacts canola yield and associated pest species. *Canadian Journal of Plant Science*, *95*, 9–20. https://doi.org/10.4141/CJPS-2014-289

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

Hasan, M. J., & Rahman, H. (2018). Resynthesis of *Brassica juncea* for resistance to *Plasmodiophora brassicae* pathotype 3. *Breeding Science*, *68*, 385–391. https://doi.org/10.1270/jsbbs.18010

Hasan, M. J., & 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*, *815*, 805–815.

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

Hasan, M. J., Shaikh, R., Megha, S., Herrmann, D. T., Kebede, B., & Rahman, H. (2021c). Mapping of flowering time, seed quality and clubroot resistance in rutabaga × spring canola populations and their association. *Euphytica*, *217*, 1–21. https://doi.org/10.1007/s10681-021-02889-7

Hasan, M. J., Strelkov, S. E., Howard, R. J., & Rahman, H. (2012). Screening of *Brassica* germplasm for resistance to *Plasmodiophora brassicae* pathotypes prevalent in Canada for broadening diversity in clubroot resistance. *Canadian Journal of Plant Science*, *92*, 501–515. https://doi.org/10.4141/CJPS2010-006

Hatakeyama, K., Niwa, T., Kato, T., Ohara, T., Kakizaki, T., & Matsumoto, S. (2017). The tandem repeated organization of NB-LRR genes in the clubroot-resistant *CRb* locus in *Brassica rapa* L. *Molecular Genetics and Genomics*, *292*, 397–405. https://doi.org/10.1007/s00438-016-1281-1

Hatakeyama, K., Suwabe, K., Tomita, R. N., Kato, T., Nunome, T., Fukuoka, H., & Matsumoto, S. (2013). Identification and Characterization of *Crr1a*, a Gene for Resistance to Clubroot Disease (*Plasmodiophora brassicae Woronin*) in *Brassica rapa* L. *PLoS ONE*, *8*, 1–10. https://doi.org/10.1371/journal.pone.0054745

Herrmann, I. T., Jørgensen, A., Bruun, S., & Hauschild, M. Z. (2013). Potential for optimized production and use of rapeseed biodiesel. Based on a comprehensive real-time LCA case study in Denmark with multiple pathways. *International Journal of Life Cycle Assessment*, *18*, 418–430. https://doi.org/10.1007/s11367-012-0486-8

Hill, W. G. (1998). A note on the theory of artificial selection in finite populations and application to QTL detection by bulk segregant analysis. *Genetical Research*, *72*, 55–58. https://doi.org/10.1017/S0016672398003334

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

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

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

Hong, C. P., Plaha, P., Koo, D. H., Yang, T. J., Choi, S. R., Lee, Y. K., Uhm, T., Bang, J. W., Edwards, D., Bancroft, I., Park, B. S., Lee, J., & Lim, Y. P. (2006). A survey of the *Brassica rapa* genome by BAC-end sequence analysis and comparison with *Arabidopsis thaliana*. *Molecules and Cells*, *22*, 300–307.

Hu, T., Chitnis, N., Monos, D., & Dinh, A. (2021). Next-generation sequencing technologies: An overview. *Human Immunology*, *82*, 801–811. https://doi.org/10.1016/j.humimm.2021.02.012

Huang, Y., Li, L., Smith, K. P., & Muehlbauer, G. J. (2016). Differential transcriptomic responses to *Fusarium graminearum* infection in two barley quantitative trait loci associated with Fusarium head blight resistance. *BMC Genomics*, *17*, 1–16. https://doi.org/10.1186/s12864-016-2716-0

Huang, Z., Peng, G., Gossen, B. D., & Yu, F. (2019). Fine mapping of a clubroot resistance gene from turnip using SNP markers identified from bulked segregant RNA-Seq. *Molecular Breeding*, *39*. 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., & Yu, F. (2017). Fine mapping of a clubroot resistance gene in Chinese cabbage using SNP markers identified from bulked segregant RNA sequencing. *Frontiers in Plant Science*, *8*, 1–9. https://doi.org/10.3389/fpls.2017.01448

Hwang, S. F., Howard, R. J., Strelkov, S. E., Gossen, B. D., & Peng, G. (2014). Special Issue: Management of clubroot (*Plasmodiophora brassicae*) on canola (*Brassica napus*) in western

Canada. *Canadian Journal of Plant Pathology*, *36*, 49–65. https://doi.org/10.1080/07060661.2013.863806

Ikegami, H., Naiki, T., Ito, T., & Imuro, Y. (1982). *Ultrastructural growth process of Plasmodiophora brassicae in infected cells of Chinese cabbage and turnip (studies on the clubroot of cruciferous plants)*.

Jeffery, E. H., Brown, A. F., Kurilich, A. C., Keck, A. S., Matusheski, N., Klein, B. P., & Juvik, J. A. (2003). Variation in content of bioactive components in broccoli. *Journal of Food Composition and Analysis*, *16*, 323–330. https://doi.org/10.1016/S0889-1575(03)00045-0

Jia, H., Wei, X., Yang, Y., Yuan, Y., Wei, F., Zhao, Y., Yang, S., Yao, Q., Wang, Z., Tian, B., & Zhang, X. (2017). Root RNA-seq analysis reveals a distinct transcriptome landscape between clubroot-susceptible and clubroot-resistant Chinese cabbage lines after *Plasmodiophora brassicae* infection. *Plant and Soil*, *421*, 93–105. https://doi.org/10.1007/s11104-017-3432-5

Jian, H., Ma, J., Wei, L., Liu, P., Zhang, A., Yang, B., Li, J., Xu, X., & Liu, L. (2018). Integrated mRNA, sRNA, and degradome sequencing reveal oilseed rape complex responses to *Sclerotinia sclerotiorum* (Lib.) infection. *Scientific Reports*, *8*, 1–17. https://doi.org/10.1038/s41598-018-29365-y

Jin, J., Liu, J., Wang, H., Wong, L., & Chua, N. H. (2013). PLncDB: Plant long non-coding RNA database. *Bioinformatics*, *29*, 1068–1071. https://doi.org/10.1093/bioinformatics/btt107

Johnston, T. D. (1970). A new factor for resistance to club root in *Brassica napas* L. *Plant Pathology*, *19*, 156–158.

Jones, D. R., Ingram, D. S., & Dixon, G. R. (1982). Characterization of isolates derived from single resting spores of *Plasmodiophora brassicae* and studies of their interaction. *Plant Pathology*, *31*, 239–246.

Jones, J.D.G. & Dangl. J. L. (2006). The plant immune system. *Nature*, *444*, 323–329.

Joshi, R. K., Megha, S., Basu, U., Rahman, H., & Kav, N. N. V. (2016). Genome wide identification and functional prediction of long non-coding RNAs responsive to *Sclerotinia sclerotiorum* infection in *Brassica napus*. *PLoS ONE*, *11*, 1–19. https://doi.org/10.1371/journal.pone.0158784

Joshua, Z., Analysis, C., Link, C., Friedman, N., Gnirke, A., & Regev, A. (2013). *Comprehensive comparative analysis of strand-specific Accessed NIH Public Access sequencing methods*. *7*, 709–715. https://doi.org/10.1038/nmeth.1491

Jourdan, P., & Salazar, E. (1993). *Brassica carinata* resynthesized by protoplast fusion. *Theoretical and Applied Genetics*, *86*, 567–572. https://doi.org/10.1007/BF00838710

Kageyama, K., & Asano, T. (2009). Life cycle of *Plasmodiophora brassicae*. *Journal of Plant Growth Regulation*, *28*, 203–211. https://doi.org/10.1007/s00344-009-9101-z

Karim, M. M., Dakouri, A., Zhang, Y., Chen, Q., Peng, G., Strelkov, S. E., Gossen, B. D., & Yu, F. (2020). Two clubroot-resistance genes, *RCR3* and *RCR9wa*, mapped in *Brassica* RAPA using bulk segregant RNA sequencing. *International Journal of Molecular Sciences*, *21*, 1–15. https://doi.org/10.3390/ijms21145033

Kassambara, A. (2017). *Practical guide to principal component methods in R: PCA, M (CA), FAMD, MFA, HCPC*. Sthda.

Kato, T., Hatakeyama, K., Fukino, N., & Matsumoto, S. (2013). Fine mapping of the clubroot resistance gene *CRb* and development of a useful selectable marker in *Brassica rapa*. *Breeding Science*, *63*, 116–124. https://doi.org/10.1270/jsbbs.63.116

Kaul, S., Koo, H. L., Jenkins, J., Rizzo, M., Rooney, T., Tallon, L. J., Feldblyum, T., Nierman, W., Benito, M.-I., Lin, X., Town, C. D., Venter, J. C., & Fraser, C. M. (2014). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Genetic Resources and Crop Evolution*, *61*, 796–815. https://doi.org/10.1134/S1022795411020074

Kebede, B., Cheema, K., Greenshields, D. L., Li, C., Selvaraj, G., & Rahman, H. (2012). Construction of genetic linkage map and mapping of QTL for seed color in *Brassica rapa*. *Genome*, *55*, 813-823. https://doi.org/10.1139/g2012-066

Kern, A. D., & Hahn, M. W. (2018). The neutral theory in light of natural selection. *Molecular Biology and Evolution*, *35*, 1366–1371. https://doi.org/10.1093/molbev/msy092

Kosambi, D. D. (2016). *The estimation of map distances from recombination values. In R*. Springer, New Delhi. https://doi.org/10.1007/978-81-322-3676-4

Kowata-Dresch, L. S., & May-De Mio, L. L. (2012). Clubroot management of highly infested soils. *Crop Protection*, *35*, 47–52. https://doi.org/10.1016/j.cropro.2011.12.012

Kuginuki, Y., Yoshikawa, H., & Hirai, M. (1999). Variation in virulence of *Plasmodiophora brassicae* in Japan tested with clubroot-resistant cultivars of Chinese cabbage (*Brassica*

rapa L. ssp. *pekinensis*). *European Journal of Plant Pathology*, *105*, 327–332. https://doi.org/10.1023/A:1008705413127

Kutter, C., Schöb, H., Stadler, M., Meins, F., & Si-Ammour, A. (2007). MicroRNA-mediated regulation of stomatal development in *Arabidopsis*. *Plant Cell*, *19*, 2417–2429. https://doi.org/10.1105/tpc.107.050377

Kwenda, S., Birch, P. R. J., & Moleleki, L. N. (2016). Genome-wide identification of potato long intergenic noncoding RNAs responsive to *Pectobacterium carotovorum* subspecies *brasiliense* infection. *BMC Genomics*, *17*. https://doi.org/10.1186/s12864-016-2967-9

Lagercrantz, U., & Lydiate, D. J. (1996). Comparative genome mapping in *Brassica*. *Genetics*, *144*, 1903–1910. https://doi.org/10.1093/genetics/144.4.1903

Laila, R., Park, J. I., Robin, A. H. K., Natarajan, S., Vijayakumar, H., Shirasawa, K., Isobe, S., Kim, H. T., & Nou, I. S. (2019). Mapping of a novel clubroot resistance QTL using ddRAD-seq in Chinese cabbage (*Brassica rapa* L.). *BMC Plant Biology*, *19*, 1–9. https://doi.org/10.1186/s12870-018-1615-8

Lamesch, P., Berardini, T. Z., Li, D., Swarbreck, D., Wilks, C., Sasidharan, R., Muller, R., Dreher, K., Alexander, D. L., Garcia-Hernandez, M., Karthikeyan, A. S., Lee, C. H., Nelson, W. D., Ploetz, L., Singh, S., Wensel, A., & Huala, E. (2012). The *Arabidopsis* Information Resource (TAIR): Improved gene annotation and new tools. *Nucleic Acids Research*, *40*, 1202–1210. https://doi.org/10.1093/nar/gkr1090

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., Fitzhugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., Levine, R., McEwan, P., … Morgan, M. J. (2001). Initial sequencing and analysis of the human genome. *Nature*, *409*, 860–921. https://doi.org/10.1038/35057062

Landry, B. S., Hubert, N., Crete, R., Chang, M. S., Lincoln, S. E., & 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.

Lee, J. T. (2012). Epigenetic regulation by long noncoding RNAs. *Science*, *338*, 1435–1439. https://doi.org/10.1126/science.1231776

Lefever, S., Rihani, A., Van der Meulen, J., Pattyn, F., Van Maerken, T., Van Dorpe, J., ... & Vandesompele, J. (2019). Cost-effective and robust genotyping using double-mismatch allele-specific quantitative PCR. *Scientific reports*, *9*, 1-11. https://doi.org/10.1038/s41598-019-38581-z

Leister, D. (2004). Tandem and segmental gene duplication and recombination in the evolution of plant disease resistance genes. *Trends in Genetics*, *20*, 113–116. https://doi.org/10.1016/j.tig.2004.01.001

Li, G., Gao, M., Yang, B., & Quiros, C. F. (2003). Gene for gene alignment between the *Brassica* and *Arabidopsis* genomes by direct transcriptome mapping. *Theoretical and Applied Genetics*, *107*, 168–180. https://doi.org/10.1007/s00122-003-1236-x

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

Li, H., Li, X., Xuan, Y., Jiang, J., Wei, Y., & Piao, Z. (2018). Genome wide identification and expression profiling of *SWEET* genes family reveals its role during *Plasmodiophora brassicae*-induced formation of clubroot in *Brassica rapa*. *Frontiers in Plant Science*, *9*, 1–16. https://doi.org/10.3389/fpls.2018.00207

Li, H., Ribaut, J. M., Li, Z., & Wang, J. (2008). Inclusive composite interval mapping (ICIM) for digenic epistasis of quantitative traits in biparental populations. *Theoretical and Applied Genetics*, *116*, 243-260. https://doi.org/10.1007/s00122-007-0663-5

Li, J., & Liu, C. (2019). Coding or noncoding, the converging concepts of RNAs. *Frontiers in Genetics*, *10*, 1–10. https://doi.org/10.3389/fgene.2019.00496

Li, L., Petsch, K., Shimizu, R., Liu, S., Xu, W.W., Ying, K., Yu, J., Scanlon, M.J., Schnable, P.S., Timmermans, M.C.P., Springer, N.M., & Muehlbauer, G.J. (2013) Mendelian and non-mendelian regulation of gene expression in maize. *PLoS Genetics, 9*, e1003202. doi:10.1371/journal.pgen.1003202

Liu, J., Huang, S., Sun, M., Liu, S., Liu, Y., Wang, W., Zhang, X., Wang, H., & Hua, W. (2012). An improved allele-specific PCR primer design method for SNP marker analysis and its application. *Plant Methods*, *8*, 1–9. https://doi.org/10.1186/1746-4811-8-34

Liu, S., Liu, Y., Yang, X., Tong, C., Edwards, D., Parkin, I. A. P., Zhao, M., Ma, J., Yu, J., Huang, S., Wang, X., Wang, J., Lu, K., Fang, Z., Bancroft, I., Yang, T. J., Hu, Q., Wang, X.,

Yue, Z., … Paterson, A. H. (2014). The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes. *Nature Communications*, *5*. https://doi.org/10.1038/ncomms4930

Liu, Y., Xu, A., Liang, F., Yao, X., Wang, Y., Liu, X., Zhang, Y., Dalelhan, J., Zhang, B., Qin, M., Huang, Z., & Shaolin, L. (2018). Screening of clubroot-resistant varieties and transfer of clubroot resistance genes to *Brassica napus* using distant hybridization. *Breeding Science*, *68*, 258–267. https://doi.org/10.1270/jsbbs.17125

Livak, K. J., & 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

Lu, S., van Eck, J., Zhou, X., Lopez, A. B., O'Halloran, D. M., Cosman, K. M., Conlin, B. J., Paolillo, D. J., Garvin, D. F., Vrebalov, J., Kochian, L. v., Küpper, H., Earle, E. D., Cao, J., & Li, L. (2006). The cauliflower *Or* gene encodes a *DnaJ* cysteine-rich domain-containing protein that mediates high levels of β-carotene accumulation. *Plant Cell*, *18*, 3594–3605. https://doi.org/10.1105/tpc.106.046417

Ludwig-Müller, J., Auer, S., Jülke, S., & Marschollek, S. (2017). Auxins and cytokinins in plant biology. *Auxins and Cytokinins in Plant Biology: Methods and Protocols* (Vol. 1569). https://doi.org/10.1007/978-1-4939-6831-2

Ludwig-Müller, J., Bennett, R. N., Kiddle, G., Ihmig, S., Ruppel, M., & Hilgenberg, W. (1999). The host range of *Plasmodiophora brassicae* and its relationship to endogenous glucosinolate content. *New Phytologist*, *141*, 443–458. https://doi.org/10.1046/j.1469-8137.1999.00368.x

Lv, H., Fang, Z., Yang, L., Zhang, Y., & Wang, Y. (2020). An update on the arsenal: mining resistance genes for disease management of *Brassica* crops in the genomic era. *Horticulture Research*, *7*. https://doi.org/10.1038/s41438-020-0257-9

Magwene, P. M., Willis, J. H., & Kelly, J. K. (2011). The statistics of bulk segregant analysis using next generation sequencing. *PLoS Computational Biology*, *7*, 1–9. https://doi.org/10.1371/journal.pcbi.1002255

Malinowski, R., Truman, W., & Blicharz, S. (2019). Genius architect or clever thief — How *Plasmodiophora brassicae* reprograms host development to establish a pathogen-oriented

physiological sink. *Molecular Plant-Microbe Interactions*, *32*, 1259–1266. https://doi.org/10.1094/MPMI-03-19-0069-CR

Manitoba Agriculture. (2013). Clubroot testing underscores importance of disease prevention. *Manitoba Agriculture, Food, and Rural Initiatives*.

Mansfeld, B. N., & Grumet, R. (2018). QTLseqr: An R package for bulk segregant analysis with next-generation sequencing. *The Plant Genome*, *11*, 180006. https://doi.org/10.3835/plantgenome2018.01.0006

Mason, A. S., Huteau, V., Eber, F., Coriton, O., Yan, G., Nelson, M. N., Cowling, W. A., & Chèvre, A. M. (2010). Genome structure affects the rate of autosyndesis and allosyndesis in AABC, BBAC and CCAB *Brassica* interspecific hybrids. *Chromosome Research*, *18*, 655–666. https://doi.org/10.1007/s10577-010-9140-0

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

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

Mertens, P. P. C., & Sangar, D. V. (1985). Analysis of the terminal sequences of the genome segments of four orbiviruses. *Virology*, *140*, 55–67.

Mithen, R., & Magrath, R. (1992). A contribution to the life history of *Plasmodiophora brassicae*: secondary plasmodia development in root galls of *Arabidopsis thaliana*. *Mycological Research*, *96*, 877–885. https://doi.org/10.1016/S0953-7562(09)81035-6

Mizuno, T., & Chou, M. Y. (1984). A unique mechanism regulating gene expression: Translational inhibition by a complementary RNA transcript (micRNA). *Biochemistry*, *81*, 1966–1970. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC345417/pdf/pnas00608-0043.pdf

Morillon, A. (2018). Non-coding RNA, Its history and discovery timeline. *Long Non-coding RNA The Dark Side of the Genome*. Elsevier. https://doi.org/https://doi.org/10.1016/B978-1-78548-265-6.50001-0

Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., & Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods*, *5*, 621–628. https://doi.org/10.1038/nmeth.1226

Mun, J. H., Kwon, S. J., Yang, T. J., Seol, Y. J., Jin, M., Kim, J. A., Lim, M. H., Kim, J. S., Baek, S., Choi, B. S., Yu, H. J., Kim, D. S., Kim, N., Lim, K. B., Lee, S. I., Hahn, J. H., Lim, Y. P., Bancroft, I., & Park, B. S. (2009). Genome-wide comparative analysis of the *Brassica rapa* gene space reveals genome shrinkage and differential loss of duplicated genes after whole genome triplication. *Genome Biology*, *10*. https://doi.org/10.1186/gb-2009-10-10-r111

Nagaharu U (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

Nagaoka, T., Doullah, M. A. U., Matsumoto, S., Kawasaki, S., Ishikawa, T., Hori, H., & Okazaki, K. (2010). Identification of QTLs that control clubroot resistance in *Brassica oleracea* and comparative analysis of clubroot resistance genes between *B. rapa* and *B. oleracea*. *Theoretical and Applied Genetics*, *120*, 1335–1346. https://doi.org/10.1007/s00122-010-1259-z

Naiki, T., Tanahashi, K., & Kageyama, K. (1984). The relationship club formation between *Brassicae* among root cruciferous infection and Species with *Plasmodiophora* subsequent club formation among Cruciferous. *Annals of the Phytopathological Society of Japan*, *50*, 211–215.

Navabi, Z. K., Huebert, T., Sharpe, A. G., O'Neill, C. M., Bancroft, I., & Parkin, I. A. P. (2013). Conserved microstructure of the *Brassica* B Genome of *Brassica nigra* in relation to homologous regions of *Arabidopsis thaliana*, *B. rapa* and *B. oleracea*. *BMC Genomics*, *14*. https://doi.org/10.1186/1471-2164-14-250

Nei, M., & Kumar, S. (2000). *Molecular evolution and phylogenetics*. Oxford university press.

Nguyen, K. le, Grondin, A., Courtois, B., & Gantet, P. (2019). Next-Generation Sequencing Accelerates Crop Gene Discovery. *Trends in Plant Science*, *24*, 263–274. https://doi.org/10.1016/j.tplants.2018.11.008

Niedringhaus, T. P., Milanova, D., Kerby, M. B., Snyder, M. P., & Barron, A. E. (2011). Landscape of next-generation sequencing technologies. *Analytical Chemistry*, *83*, 4327–4341. https://doi.org/10.1021/ac2010857

Nijhuis, K. (1995). Greek doctors and Roman patients: a medical anthropological approach. In *Ancient Medicine in Its Socio-Cultural Context, Volume 1*. Brill Rodopi.

Olsen, O., & Sorensen, H. (1980). Sinalbin and other glucosinolates in seeds of Double low Rape Species and *Brassica napus* cv. Bronowski. *Journal of Agricultural and Food Chemistry*, *28*, 43–48.

Onodera, Y., Haag, J. R., Ream, T., Nunes, P. C., Pontes, O., & Pikaard, C. S. (2005). Plant nuclear *RNA polymerase IV* mediates siRNA and DNA methylation- dependent heterochromatin formation. *Cell*, *120*, 613–622. https://doi.org/10.1016/j.cell.2005.02.007

Pandey, M. K., Khan, A. W., Singh, V. K., Vishwakarma, M. K., Shasidhar, Y., Kumar, V., Garg, V., Bhat, R. S., Chitikineni, A., Janila, P., Guo, B., & Varshney, R. K. (2017). QTL-seq approach identified genomic regions and diagnostic markers for rust and late leaf spot resistance in groundnut (*Arachis hypogaea* L.). *Plant Biotechnology Journal*, *15*, 927–941. https://doi.org/10.1111/pbi.12686

Pang, W., Fu, P., Li, X., Zhan, Z., Yu, S., & Piao, Z. (2018). Identification and mapping of the clubroot resistance gene *CRd* in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*). *Frontiers in Plant Science*, *9*, 1–9. https://doi.org/10.3389/fpls.2018.00653

Pang, W., Liang, Y., Zhan, Z., Li, X., & Piao, Z. (2020). Development of a sinitic clubroot differential set for the pathotype classification of *Plasmodiophora brassicae*. *Frontiers in Plant Science*, *11*, 1–9. https://doi.org/10.3389/fpls.2020.568771

Park, Y. S., Kim, S., Park, D. G., Kim, D. H., Yoon, K. W., Shin, W., & Han, K. (2019). Comparison of library construction kits for mRNA sequencing in the Illumina platform. *Genes and Genomics*, *41*, 1233–1240. https://doi.org/10.1007/s13258-019-00853-3

Peer, W. A., Brown, D. E., Tague, B. W., Muday, G. K., Taiz, L., & Murphy, A. S. (2001). Flavonoid accumulation patterns of transparent test a mutants of *Arabidopsis*. *Plant Physiology*, *126*, 536–548. https://doi.org/10.1104/pp.126.2.536

Peng, G., Falk, K. C., Gugel, R. K., Franke, C., Yu, F., James, B., Strelkov, S. E., Hwang, S. F., & McGregor, L. (2014a). Sources of resistance to *Plasmodiophora brassicae* (clubroot) pathotypes virulent on canola. *Canadian Journal of Plant Pathology*, *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., & Strelkov, S. E. (2014b). Special Issue: Crop rotation, cultivar resistance, and fungicides/

biofungicides for managing clubroot (*Plasmodiophora brassicae*) on canola. *Canadian Journal of Plant Pathology*, *36*, 99–112. https://doi.org/10.1080/07060661.2013.860398

Perumal, S., Koh, C. S., Jin, L., Buchwaldt, M., Higgins, E. E., Zheng, C., Sankoff, D., Robinson, S. J., Kagale, S., Navabi, Z. K., Tang, L., Horner, K. N., He, Z., Bancroft, I., Chalhoub, B., Sharpe, A. G., & Parkin, I. A. P. (2020). A high-contiguity *Brassica nigra* genome localizes active centromeres and defines the ancestral *Brassica* genome. *Nature Plants*, *6*, 929–941. https://doi.org/10.1038/s41477-020-0735-y

Piao, Z. Y., Deng, Y. Q., Choi, S. R., Park, Y. J., & Lim, Y. P. (2004). SCAR and CAPS mapping of *CRb*, a gene conferring resistance to *Plasmodiophora brassicae* in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*). *Theoretical and Applied Genetics*, *108*, 1458–1465. https://doi.org/10.1007/s00122-003-1577-5

Piao, Z., Ramchiary, N., & Lim, Y. P. (2009). Genetics of clubroot resistance in *Brassica* species. *Journal of Plant Growth Regulation*, *28*, 252–264. https://doi.org/10.1007/s00344-009-9093-8

Pop, M., & Salzberg, S. L. (2008). Bioinformatics challenges of new sequencing technology. *Trends in Genetics*, *24*, 142–149. https://doi.org/10.1016/j.tig.2007.12.006

Przybylski, R., & 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., & McDonald, B. E. (2005). Canola Oil. *Bailey'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

Rahman, H., Peng, G., Yu, F., Falk, K. C., Kulkarni, M., & Selvaraj, G. (2014). Special Issue: Genetics and breeding for clubroot resistance in Canadian spring canola (*Brassica napus* L.). *Canadian Journal of Plant Pathology*, *36*, 122–134. https://doi.org/10.1080/07060661.2013.862571

Rahman, H. (2001). Production of yellow-seeded *Brassica napus* through interspecific crosses. *Plant Breeding*, *120*, 463–472. https://doi.org/10.1046/j.1439-0523.2001.00640.x

Rahman, H. (2005). Resynthesis of *Brassica napus* L. for self-incompatibility: self-incompatibility reaction, inheritance and breeding potential. *Plant Breeding*, *124*, 13–19.

Ramzi, N., Kaczmarek, J., & Jedryczka, M. (2018). Identification of clubroot resistance sources from world gene bank accessions. *IOBC-WPRS Bulletin*, *136*, 144–147. https://www.researchgate.net/publication/336281408

Rempel, C. B., Hutton, S. N., & Jurke, C. J. (2014). Clubroot and the importance of canola in Canada. *Canadian Journal of Plant Pathology*, *36*, 19–26. https://doi.org/10.1080/07060661.2013.864336

Ren, L., Jia, J., Li, M., Liu, F., Cheng, Y., Zhang, M., Chen, K., & Fang, X. (2012). Distribution of rapeseed clubroot disease in Hubei Province and evaluation of yield loss. *Agricultural Science & Technology-Hunan*, *13*, 773–775.

Rocherieux, J., Glory, P., Giboulot, A., Boury, S., Barbeyron, G., Thomas, G., & Manzanares-Dauleux, M. J. (2004). Isolate-specific and broad-spectrum QTLs are involved in the control of clubroot in *Brassica oleracea*. *Theoretical and Applied Genetics*, *108*, 1555–1563. https://doi.org/10.1007/s00122-003-1580-x

Rousseau-Gueutin, M., Belser, C., Silva, C. da, Richard, G., Istace, B., Cruaud, C., Falentin, C., Boideau, F., Boutte, J., Delourme, R., Deniot, G., Engelen, S., de Carvalho, J. F., Lemainque, A., Maillet, L., Morice, J., Wincker, P., Denoeud, F., Chèvre, A. M., & Aury, J. M. (2021). Long-read assembly of the *Brassica napus* reference genome Darmor-bzh. *GigaScience*, *9*, 1–16. https://doi.org/10.1093/gigascience/giaa137

Saito, M., Kubo, N., Matsumoto, S., Suwabe, K., Tsukada, M., & Hirai, M. (2006). Fine mapping of the clubroot resistance gene, *Crr3*, in *Brassica rapa*. *Theoretical and Applied Genetics*, *114*, 81–91. https://doi.org/10.1007/s00122-006-0412-1

Sakamoto, K., Saito, A., Hayashida, N., Taguchi, G., & Matsumoto, E. (2008). Mapping of isolate-specific QTLs for clubroot resistance in Chinese cabbage (*Brassica rapa* L. ssp. *pekinensis*). *Theoretical and Applied Genetics*, *117*, 759–767. https://doi.org/10.1007/s00122-008-0817-0

Sarris, P. F., Cevik, V., Dagdas, G., Jones, J. D. G., & Krasileva, K. V. (2016). Comparative analysis of plant immune receptor architectures uncovers host proteins likely targeted by pathogens. *BMC Biology*, *14*. https://doi.org/10.1186/s12915-016-0228-7

Scheben, A., Batley, J., & Edwards, D. (2017). Genotyping-by-sequencing approaches to characterize crop genomes: choosing the right tool for the right application. *Plant Biotechnology Journal*, *15*, 149–161. https://doi.org/10.1111/pbi.12645

Schulze-Lefert, P., & Panstruga, R. (2011). A molecular evolutionary concept connecting nonhost resistance, pathogen host range, and pathogen speciation. *Trends in Plant Science*, *16*, 117–125. https://doi.org/10.1016/j.tplants.2011.01.001

Seol, Y. J., Kim, K., Kang, S. H., Perumal, S., Lee, J., & Kim, C. K. (2017). The complete chloroplast genome of two *Brassica* species, *Brassica nigra* and *B. oleracea*. *Mitochondrial DNA Part A: DNA Mapping, Sequencing, and Analysis*, *28*, 167–168. https://doi.org/10.3109/19401736.2015.1115493

Shafiq, S., Li, J., & Sun, Q. (2016). Functions of plants long non-coding RNAs. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, *1859*, 155–162. https://doi.org/10.1016/j.bbagrm.2015.06.009

Slatko, B. E., Gardner, A. F., & Ausubel, F. M. (2018). Overview of next-generation sequencing technologies. *Current Protocols in Molecular Biology*, *122*, 1–15. https://doi.org/10.1002/cpmb.59

Somé, A., Manzanares, M. J., Laurens, F., Baron, F., Thomas, G., & Rouxel, F. (1996). Variation for virulence on *Brassica napus* L. amongst *Plasmodiophora brassicae* collections from France and derived single-spore isolates. *Plant Pathology*, *45*, 432–439.

Spaner, D. (2002). Agronomic and horticultural characters of rutabaga in eastern Canada. *Canadian Journal of Plant Science*, *82*, 221–224. https://doi.org/10.4141/P01-086

Stam, R., & McDonald, B. A. (2018). When resistance gene pyramids are not durable—the role of pathogen diversity. *Molecular Plant Pathology*, *19*, 521–524. https://doi.org/10.1111/mpp.12636

Strelkov, S. E., Hwang, S. F., Manolii, V. P., Cao, T., & Feindel, D. (2016). Emergence of new virulence phenotypes of *Plasmodiophora brassicae* on canola (*Brassica napus*) in Alberta, Canada. *European Journal of Plant Pathology*, *145*, 517–529. https://doi.org/10.1007/s10658-016-0888-8

Strelkov, S. E., Hwang, S. F., Manolii, V. P., Cao, T., Fredua-Agyeman, R., Harding, M. W., Peng, G., Gossen, B. D., Mcdonald, M. R., & Feindel, D. (2018). Virulence and pathotype classification of *Plasmodiophora brassicae* populations collected from clubroot resistant canola (*Brassica napus*) in Canada. *Canadian Journal of Plant Pathology*, *40*, 284–298. https://doi.org/10.1080/07060661.2018.1459851

Strelkov, S. E., Hwang, S. F., Manolii, V. P., Turnbull, G., Fredua-Agyeman, R., Hollman, K., & Kaus, S. (2020). Characterization of clubroot (*Plasmodiophora brassicae*) from canola (*Brassica napus*) in the Peace Country of Alberta, Canada. *Canadian Journal of Plant Pathology*, *00*, 1–7. https://doi.org/10.1080/07060661.2020.1776931

Strelkov, S. E., Manolii, V. P., Cao, T., Xue, S., & Hwang, S. F. (2007). Pathotype classification of *Plasmodiophora brassicae* and its occurrence in *Brassica napus* in Alberta, Canada. *Journal of Phytopathology*, *155*, 706–712. https://doi.org/10.1111/j.1439-0434.2007.01303.x

Strelkov, S. E., Manolii, V. P., Harding, M. W., Daniels, G. C., Nuffer, P., Aigu, Y., & Hwang, S. F. (2020). The occurrence and spread of clubroot on canola in Alberta in 2019. *Canadian Plant Disease Survey 2020 Volume 100: Disease Highlights 2019*, 117–120.

Subramanian, S., & Kumar, S. (2003). Neutral substitutions occur at a faster rate in exons than in noncoding DNA in primate genomes. *Genome Research*, *13*, 838–844. https://doi.org/10.1101/gr.1152803

Summanwar, A., Basu, U., Rahman, H., & 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. *Molecular Plant-Microbe Interactions*, *32*, 1360–1377. https://doi.org/10.1094/MPMI-12-18-0341-R

Summanwar, A., Farid, M., Basu, U., Kav, N., & Rahman, H. (2021). Comparative transcriptome analysis of canola carrying clubroot resistance from 'Mendel' or rutabaga and the development of molecular markers. *Physiological and Molecular Plant Pathology*, *114*, 101640. https://doi.org/10.1016/j.pmpp.2021.101640

Sun, X., Zheng, H., & Sui, N. (2018). Regulation mechanism of long non-coding RNA in plant response to stress. *Biochemical and Biophysical Research Communications*, *503*, 402–407. https://doi.org/10.1016/j.bbrc.2018.07.072

Suwabe, K., Tsukazaki, H., Iketani, H., Hatakeyama, K., Kondo, M., Fujimura, M., ... & Matsumoto, S. (2006). Simple sequence repeat-based comparative genomics between *Brassica rapa* and *Arabidopsis thaliana*: the genetic origin of clubroot resistance. *Genetics*, *173*, 309-319. https://doi.org/10.1534/genetics.104.038968

Szaker, H. M., Darkó, É., Medzihradszky, A., Janda, T., Liu, H. C., Charng, Y. Y., & Csorba, T. (2019). *mir824*/*agamous-like16* Module integrates recurring environmental heat stress

changes to fine-tune poststress development. *Frontiers in Plant Science*, *10*, 1–19. https://doi.org/10.3389/fpls.2019.01454

Takagi, H., Abe, A., Yoshida, K., Kosugi, S., Natsume, S., Mitsuoka, C., Uemura, A., Utsushi, H., Tamiru, M., Takuno, S., Innan, H., Cano, L. M., Kamoun, S., & Terauchi, R. (2013). QTL-seq: Rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. *Plant Journal*, *74*, 174–183. https://doi.org/10.1111/tpj.12105

Takahashi, K. (1994). Influences of some environmental factors on the viability of resting spores of *Plasmodiophora brassicae Wor*. Incubated in Sterile Soil. *Japanese Journal of Phytopathology*, *60*, 658–666. https://doi.org/10.3186/jjphytopath.60.658

Tewari, J. P., Strelkov, S. E., Orchard, D., Hartman, M., Lange, R. M., & Turkington, T. K. (2005). Identification of clubroot of crucifers on canola (*Brassica napus*) in Alberta. *Canadian Journal of Plant Pathology*, *27*, 143–144. https://doi.org/10.1080/07060660509507206

Thordal-Christensen, H. (2020). A holistic view on plant effector-triggered immunity presented as an iceberg model. *Cellular and Molecular Life Sciences*, *77*, 3963–3976. https://doi.org/10.1007/s00018-020-03515-w

Ton, J., Flors, V., & Mauch-Mani, B. (2009). The multifaceted role of ABA in disease resistance. *Trends in Plant Science*, *14*, 310–317. https://doi.org/10.1016/j.tplants.2009.03.006

Tonosaki, K., Michiba, K., Bang, S. W., Kitashiba, H., Kaneko, Y., & Nishio, T. (2013). Genetic analysis of hybrid seed formation ability of *Brassica rapa* in intergeneric crossings with *Raphanus sativus*. *Theoretical and applied genetics*, *126*, 837-846.

Treutter, D. (2005). Significance of flavonoids in plant resistance and enhancement of their biosynthesis. *Plant Biology*, *7*, 581–591.

Tso, H. H., Galindo-González, L., & Strelkov, S. E. (2021). Review current and future pathotyping platforms for *Plasmodiophora brassicae* in Canada. *Plants*, *10*. https://doi.org/10.3390/plants10071446

Ueno, H., Matsumoto, E., Aruga, D., Kitagawa, S., Matsumura, H., & Hayashida, N. (2012). Molecular characterization of the *CRa* gene conferring clubroot resistance in *Brassica rapa*. *Plant Molecular Biology*, *80*, 621–629. https://doi.org/10.1007/s11103-012-9971-5

Urquiaga, M. C. de O., Thiebaut, F., Hemerly, A. S., & Ferreira, P. C. G. (2021). From trash to luxury: the potential role of plant LncRNA in DNA methylation during abiotic stress. *Frontiers in Plant Science*, *11*, 1–10. https://doi.org/10.3389/fpls.2020.603246

Varshney, R. K., Pandey, M. K., & Chitikineni, A. (2018). Plant genetics and molecular biology: An introduction. *Advances in Biochemical Engineering/Biotechnology*. https://doi.org/10.1007/10_2017_45

Vig, A. P., Rampal, G., Thind, T. S., & Arora, S. (2009). Bio-protective effects of glucosinolates–A review. *LWT-Food Science and Technology*, *42*, 1561–1572.

Walker, J. C. (1939). Resistance to clubroot in varieties of turnip and rutabaga. *Journal of Agriculture Research. 59*, 815–828.

Walerowski, P., Gündel, A., Yahaya, N., Truman, W., Sobczak, M., Olszak, M., Rolfe, S., Borisjuk, L., & Malinowski, R. (2018). Clubroot disease stimulates early steps of phloem differentiation and recruits *SWEET* sucrose transporters within developing galls. *Plant Cell*, *30*, 3058–3073. https://doi.org/10.1105/tpc.18.00283

Wallenhammar, A. C. (1996). Prevalence of *Plasmodiophora brassicae* in a spring oilseed rape growing area in central Sweden and factors influencing soil infestation levels. *Plant Pathology*, *45*, 710–719. https://doi.org/10.1046/j.1365-3059.1996.d01-173.x

Waminal, N. E., Perumal, S., Lee, J., Kim, H. H., & Yang, T.-J. (2016). Repeat Evolution in *Brassica rapa* (AA), *B. oleracea* (CC), and *B. napus* (AACC) Genomes . *Plant Breeding and Biotechnology*, *4*, 107–122. https://doi.org/10.9787/pbb.2016.4.2.107

Wang, J., Chapman, S. C., Bonnett, D. G., Rebetzke, G. J., & Crouch, J. (2007). Application of population genetic theory and simulation models to efficiently pyramid multiple genes via marker-assisted selection. *Crop Science*, *47*, 582–590. https://doi.org/10.2135/cropsci2006.05.0341

Wang, J., Huang, Y., Hu, X. L., Niu, Y. Z., Li, X. L., & Liang, Y. (2008). Study on symptom, yield loss of clubroot and modality of *Plasmodiophora brassicae* in rape. *Chinese Journal of Oil Crop Sciences*, *30*, 112.

Wang, S., Yu, F., Zhang, W., Tang, J., Li, J., Yu, L., Wang, H., & Jiang, J. (2019). Comparative transcriptomic analysis reveals gene expression changes during early stages of *Plasmodiophora brassicae* infection in cabbage (*Brassica oleracea* var. *capitata* L.).

Canadian Journal of Plant Pathology, *41*, 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*. *Nature Genetics*, *43*, 1035–1040. https://doi.org/10.1038/ng.919

Wang, Z., Gerstein, M., & Snyder, M. (2010). *RNA-Seq: a revolutionary tool for transcriptomics*. *10*, 57–63. https://doi.org/10.1038/nrg2484.RNA-Seq

Werner, S., Diederichsen, E., Frauen, M., Schondelmaier, J., & Jung, C. (2008). Genetic mapping of clubroot resistance genes in oilseed rape. *Theoretical and Applied Genetics*, *116*, 363–372. https://doi.org/10.1007/s00122-007-0674-2

Weyen, J. (2021). Applications of doubled haploids in plant breeding and applied research. *Methods in Molecular Biology*, *2287*, 23–39. https://doi.org/10.1007/978-1-0716-1315-3_2

Wierzbick, A. T., Haag, J. R., & Pikaard, C. S. (2008). Noncoding transcription by RNA polymerase *Pol IVb/Pol V* mediates transcriptional silencing of overlapping and adjacent genes. *Cell*, *135*, 635–648. https://doi.org/10.1016/j.cell.2008.09.035

Williams, P. H. (1966). A system for the determination of races of *Plasmodiophora brassicae* that infect cabbage and rutabaga. *Phytopathology*, *56*, 624–626.

Xin, M., Wang, Y., Yao, Y., Song, N., Hu, Z., Qin, D., Xie, C., Peng, H., Ni, Z., & Sun, Q. (2011). Identification and characterization of wheat long non-protein coding RNAs responsive to powdery mildew infection and heat stress by using microarray analysis and SBS sequencing. *BMC Plant Biology*, *11*. https://doi.org/10.1186/1471-2229-11-61

Xu, M. Y., Dong, Y., Zhang, Q. X., Zhang, L., Luo, Y. Z., Sun, J., Fan, Y. L., & Wang, L. (2012). Identification of miRNAs and their targets from *Brassica napus* by high-throughput sequencing and degradome analysis. *BMC Genomics*, *13*, 1. https://doi.org/10.1186/1471-2164-13-421

Xue, S., Cao, T., Howard, R. J., Hwang, S. F., & Strelkov, S. E. (2008). Isolation and variation in virulence of single-spore isolates of *Plasmodiophora brassicae* from Canada. *Plant Disease*, *92*, 456–462. https://doi.org/10.1094/PDIS-92-3-0456

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. *Nature Genetics*, *48*, 1225–1232. https://doi.org/10.1038/ng.3657

Yang, J., Liu, G., Zhao, N., Chen, S., Liu, D., Ma, W., Hu, Z., & Zhang, M. (2016b). Comparative mitochondrial genome analysis reveals the evolutionary rearrangement mechanism in *Brassica*. *Plant Biology*, *18*, 527–536. https://doi.org/10.1111/plb.12414

Yang, Y. W., Lai, K. N., Tai, P. Y., & Li, W. H. (1999). Rates of nucleotide substitution in angiosperm mitochondrial DNA sequences and dates of divergence between *Brassica* and other angiosperm lineages. *Journal of Molecular Evolution*, *48*, 597–604. https://doi.org/10.1007/PL00006502

Yao, Y., Li, K., Liu, H., Duncan, R. W., Guo, S., Xiao, L., & Du, D. (2017). Whole-genome re-sequencing and fine mapping of an orange petal color gene (*Bnpc1*) in spring *Brassica napus* L. to a 151-kb region. *Euphytica*, *213*, 1–12. https://doi.org/10.1007/s10681-017-1959-4

Yi, X., Zhang, Z., Ling, Y., Xu, W., & Su, Z. (2015). PNRD: A plant non-coding RNA database. *Nucleic Acids Research*, *43*, D982–D989. https://doi.org/10.1093/nar/gku1162

You, F. M., Huo, N., Gu, Y. Q., 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*, 1–13. https://doi.org/10.1186/1471-2105-9-253

Yu, F., Zhang, X., Huang, Z., Chu, M., Song, T., Falk, K. C., Deora, A., Chen, Q., Zhang, Y., McGregor, L., Gossen, B. D., McDonald, M. R., & Peng, G. (2016). Identification of genome-wide variants and discovery of variants associated with *Brassica rapa* clubroot resistance gene *Rcr1* through bulked segregant RNA sequencing. *PLoS ONE*, *11*, 1–23. https://doi.org/10.1371/journal.pone.0153218

Yu, F., Zhang, X., Peng, G., Falk, K. C., Strelkov, S. E., & Gossen, B. D. (2017). Genotyping-by-sequencing reveals three QTL for clubroot resistance to six pathotypes of *Plasmodiophora brassicae* in *Brassica rapa*. *Scientific Reports*, *7*, 1–11. https://doi.org/10.1038/s41598-017-04903-2

Zampetaki, A., Albrecht, A., & Steinhofel, K. (2018). Long non-coding RNA structure and function: Is there a link? *Frontiers in Physiology*, *9*, 1–8. https://doi.org/10.3389/fphys.2018.01201

Zhan, Z., Jiang, Z., Zhu, Y., Zhang, C., Yang, Q., Li, Q., Hou, Z., Gong, J., Cheng, Y., Wu, J., & Fu, T. (2015). Development of close linked marker to *PbBa8. 1* conferring canola resistance to *Plasmodiophora brassicae*. *Chinese Journal of Oil Crop Sciences*, *37*, 766.

Zhang, H., Guo, H., Hu, W., & Ji, W. (2020). The emerging role of long non-coding RNAs in plant defense against fungal stress. *International Journal of Molecular Sciences*, *21*. https://doi.org/10.3390/ijms21082659

Zhang, L., Wang, M., Li, N., Wang, H., Qiu, P., Pei, L., Xu, Z., Wang, T., Gao, E., Liu, J., Liu, S., Hu, Q., Miao, Y., Lindsey, K., Tu, L., Zhu, L., & Zhang, X. (2018). Long noncoding RNAs involve in resistance to *Verticillium dahliae*, a fungal disease in cotton. *Plant Biotechnology Journal*, *16*, 1172–1185. https://doi.org/10.1111/pbi.12861

Zhang, W., Wang, S., Yu, F., Tang, J., Shan, X., Bao, K., Yu, L., Wang, H., Fei, Z., & Li, J. (2019a). Genome-wide characterization and expression profiling of *SWEET* genes in cabbage (*Brassica oleracea* var. *capitata* L.) reveal their roles in chilling and clubroot disease responses. *BMC Genomics*, *20*, 1–16. https://doi.org/10.1186/s12864-019-5454-2

Zhang, W., Wang, S., Yu, F., Tang, J., Yu, L., Wang, H., & Li, J. (2019b). Genome-wide identification and expression profiling of sugar transporter protein (*STP*) family genes in cabbage (*Brassica oleracea* var. *capitata* L.) reveals their involvement in clubroot disease responses. *Genes*, *10*. https://doi.org/10.3390/genes10010071

Zhang, X., Liu, Y., Fang, Z., Li, Z., Yang, L., Zhuang, M., Zhang, Y., & 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. *Frontiers in Plant Science*, *7*, 1–15. https://doi.org/10.3389/fpls.2016.01929

Zhang, X., Li, R., Chen, L., Niu, S., Chen, L., Gao, J., Wen, J., Yi, B., Ma, C., Tu, J., Fu, T., & Shen, J. (2018). Fine-mapping and candidate gene analysis of the *Brassica juncea* white-flowered mutant *Bjpc2* using the whole-genome resequencing. *Molecular Genetics and Genomics*, *293*, 359–370. https://doi.org/10.1007/s00438-017-1390-5

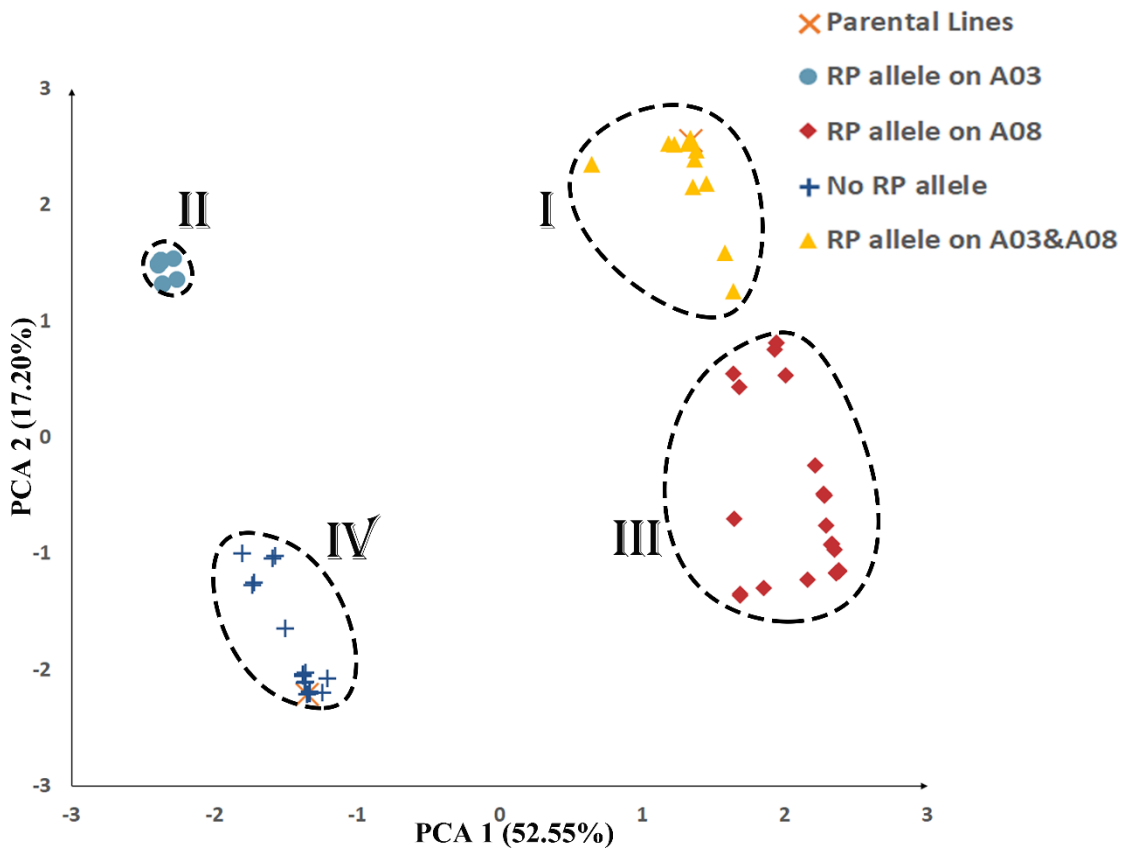
Zheng, J., Zhao, H., Ma, Y., Jiang, M., Zhan, Z., Li, X., & Piao, Z. (2022). Marker-assisted pyramiding of genes for multilocular ovaries, self-compatibility, and clubroot resistance in chinese cabbage (*Brassica rapa* L. ssp. *pekinensis*). *Horticulturae*, *8*, 139. https://doi.org/10.3390/horticulturae8020139

Zhou, Q., Galindo-González, L., Hwang, S. F., & Strelkov, S. E. (2021). Application of genomics and transcriptomics to accelerate development of clubroot resistant canola. *Canadian Journal of Plant Pathology*, *43*, 189–208. https://doi.org/10.1080/07060661.2020.1794541

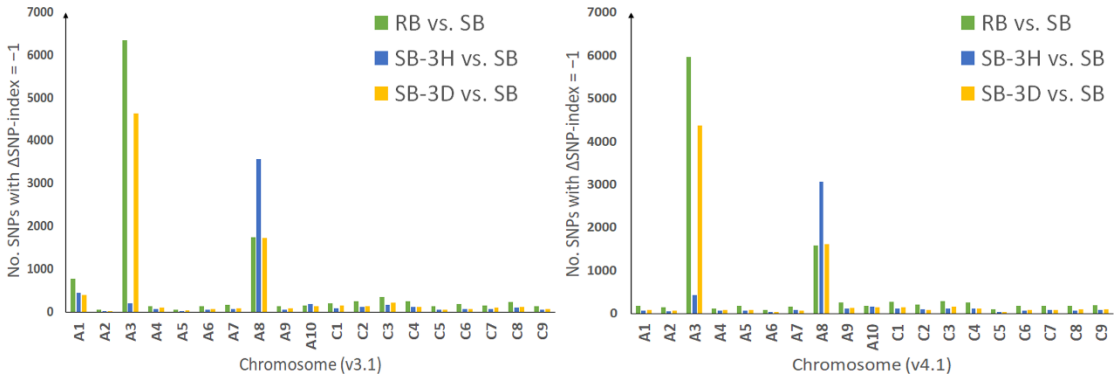
Zhou, Q., Hwang, S. F., Strelkov, S. E., Fredua-Agyeman, R., & Manolii, V. P. (2018). A molecular marker for the specific detection of new pathotype 5-like strains of *Plasmodiophora brassicae* in canola. *Plant Pathology*, *67*, 1582–1588. https://doi.org/10.1111/ppa.12868

Zhou, X. G., Ren, L. F., Li, Y. T., Zhang, M., Yu, Y. de, & Yu, J. (2010). The next-generation sequencing technology: A technology review and future perspective. *Science China Life Sciences*, *53*, 44–57. https://doi.org/10.1007/s11427-010-0023-6

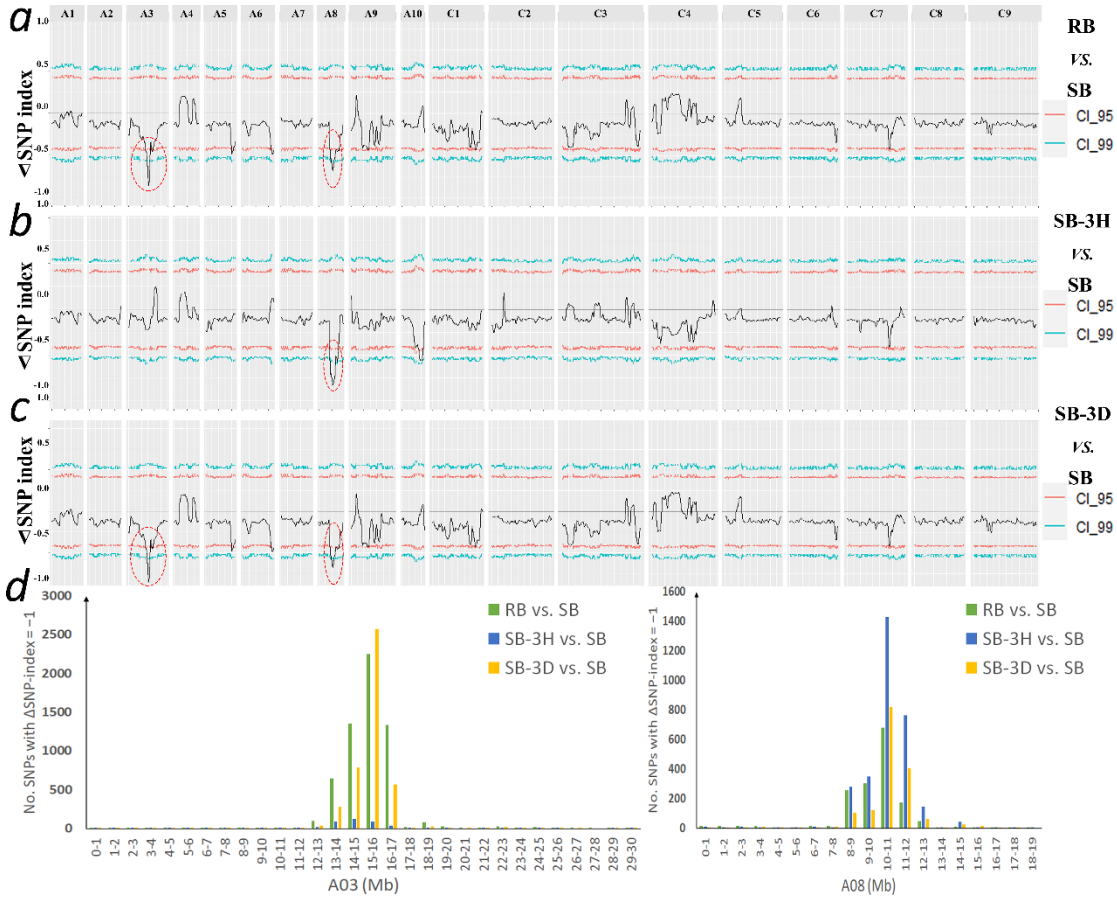
Zhu, H., Zhai, W., Li, X., & Zhu, Y. (2019). Two QTLs controlling Clubroot resistance identified from Bulked Segregant Sequencing in Pakchoi (*Brassica campestris* ssp. *chinensis* Makino). *Scientific Reports*, *9*, 1–9. https://doi.org/10.1038/s41598-019-44724-z



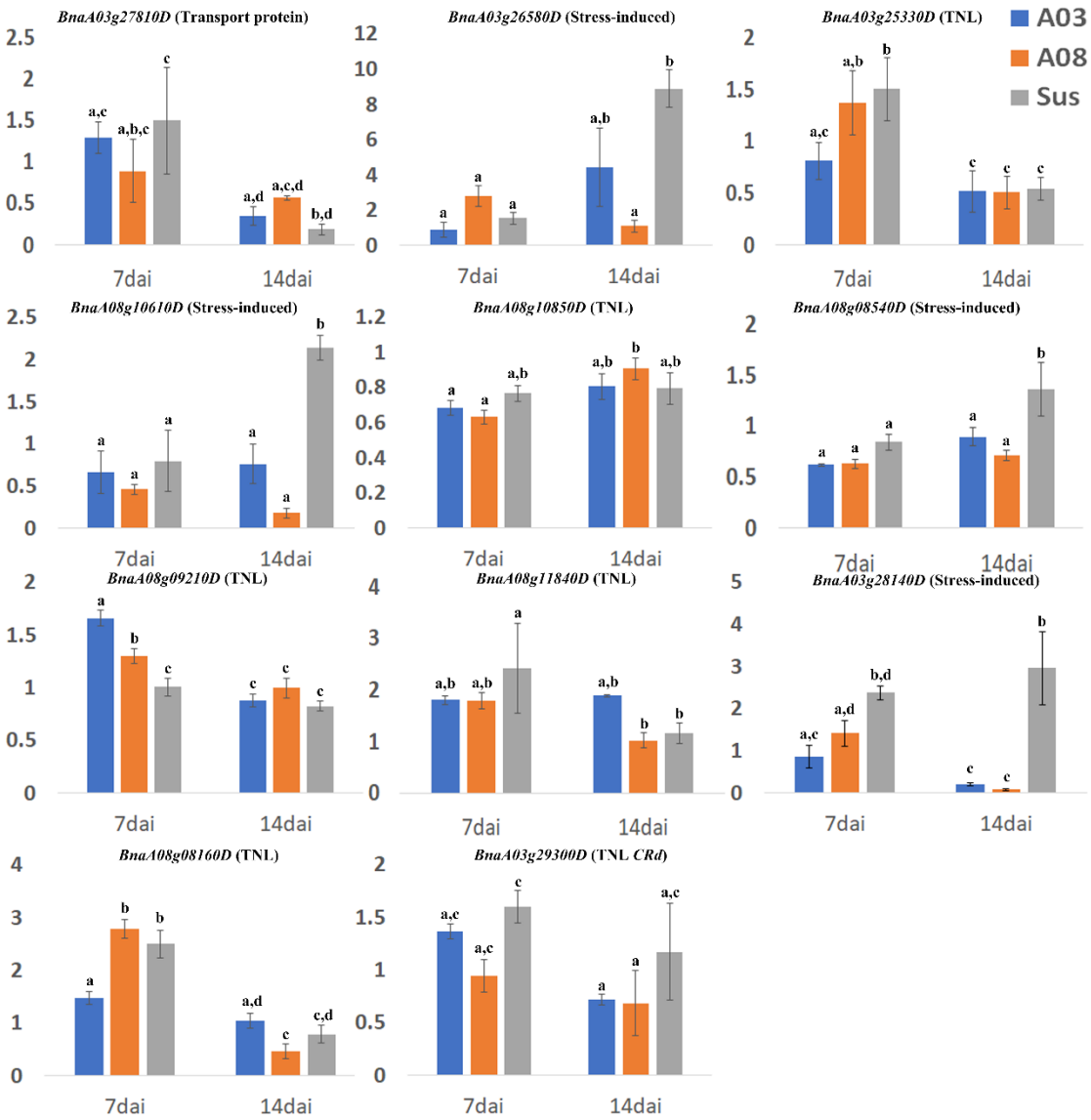
Supplementary Figure S2.1 Principal component analysis (PCA) of the *Brassica napus* doubled haploid (DH) population (*n* = 90) based on genotypic data of 31 markers from the QTL regions associated with clubroot resistance. For PCA analysis, '1' was used for resistant parent (RP) allele while '0' was used for susceptible parent (SP) allele. The parents are showed with orange-cross. The PCA placed the population into four groups: I, II, III and IV, which are marked with yellow-triangle (RP alleles on both A03 and A08), pale blue-circle (RP allele on A03, but SP allele on A08), red-diamond (RP allele on A08, but SP allele on A03) and dark blue-plus (SP alleles on both A03 and A08), respectively.



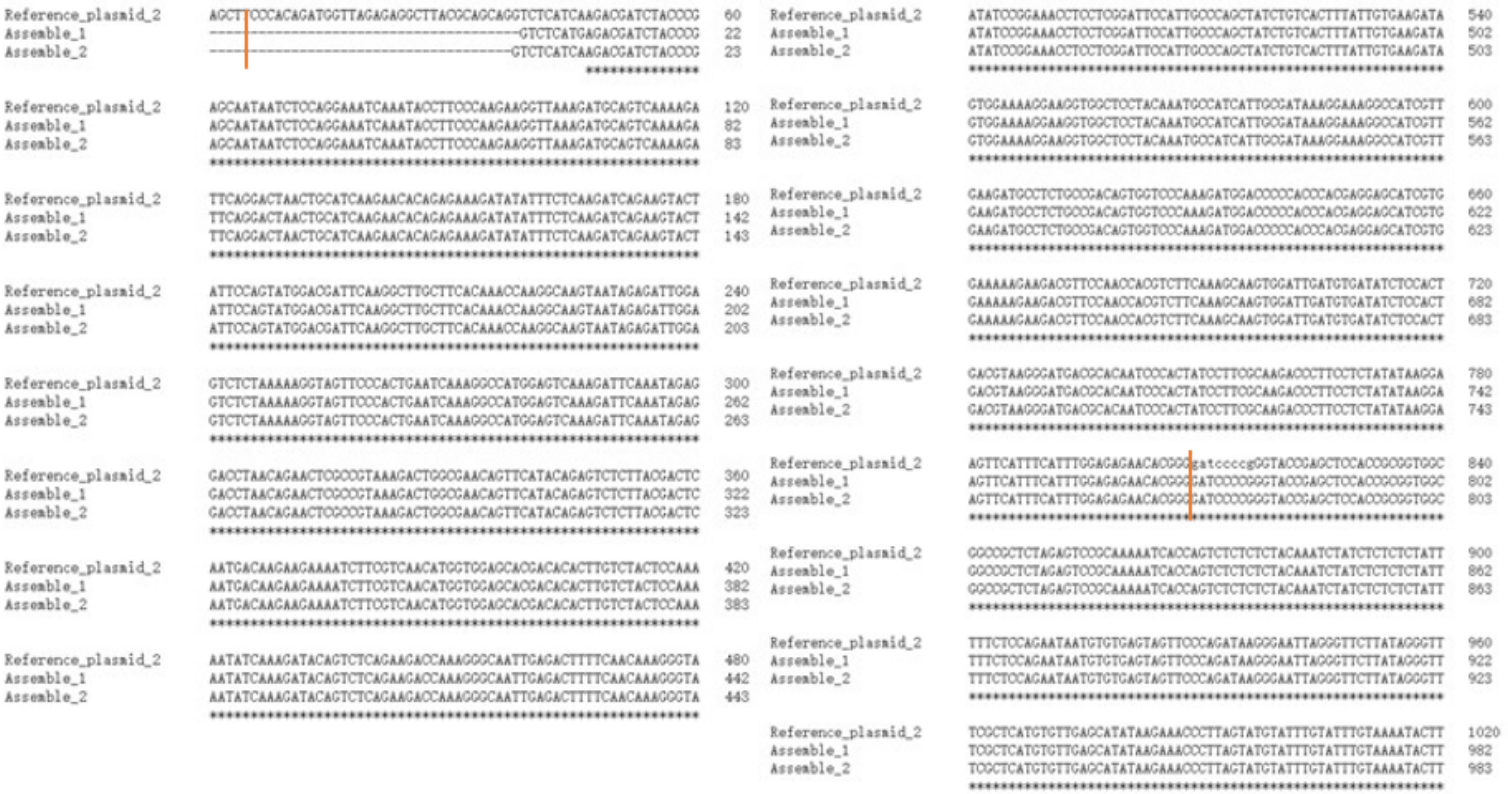
Supplementary Figure S2.2 The distribution of the SNP with ΔSNP-index = −1on *Brassica napus* chromosomes, and at Mb positions of the chromosomes A03 and A08; the SNPs were detected by using the reference genomes Brassica_napus_v3.1 (http://cruciferseq.ca/?q=node/6) (v3.1) and Brassica_napus_v4.1 (Chalhoub et al., 2014) (v4.1). The ΔSNP-indices were identified by using three bulks of *B. napus* doubled haploid (DH) lines: Bulk resistant to pathotype 3H, 3A and 3D (RB); bulk susceptible to 3H but resistant to 3A and 3D (SB-3H); and bulk susceptible to 3D but resistant to 3H and 3A (SB-3D) (Supplementary Table 2.3).



Supplementary Figure S2.3 The ΔSNP-index plots of *Brassica napus* chromosomes developed using three bulks of doubled haploid (DH) lines: Bulk resistant to pathotype 3H, 3A and 3D (RB); bulk susceptible to 3H but resistant to 3D and 3A (SB-3H); and bulk susceptible to 3D but resistant to 3H and 3A (SB-3D) (Supplementary Table 2.3) with two-sided confidence intervals, 95% (red) and 99% (blue). Analysis carried out using whole genome resequencing (WGRS) data based on the *B. napus* reference genome Brassica_napus_v4.1 (Chalhoub et al. 2014). Putative quantitative trait loci (QTL) for clubroot resistance are encircled.

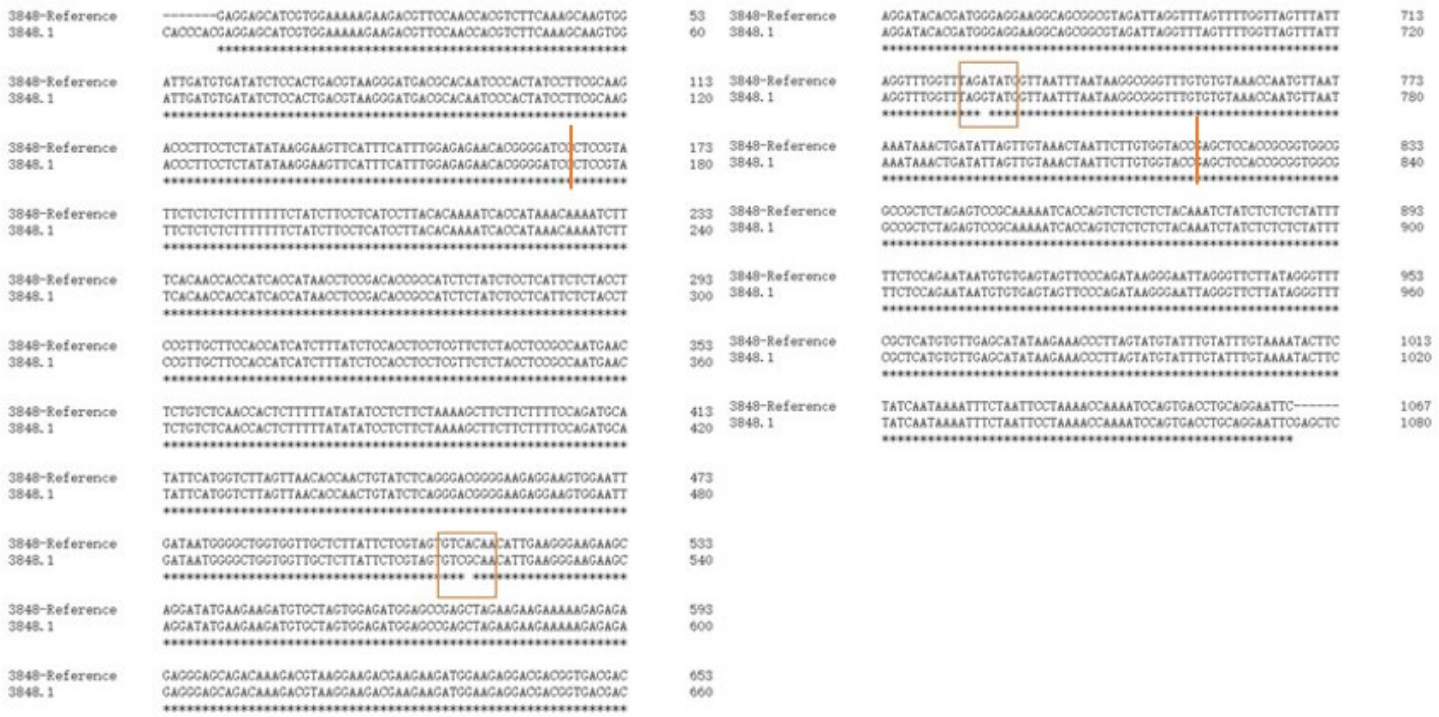


Supplementary Figure S2.4 Expression analysis of 11 genes in roots of the *Brassica napus* doubled haploid (DH) lines carrying resistance on A03 (blue) or A08 (orange), and the susceptible lines (grey) at 7 and 14 days after inoculation (dai). The name of the genes and the reasons for inclusion is shown at the top of figures. Expression of the mock-inoculated plants was set at 1 for each time point. Lines on the top of the bars show standard error.

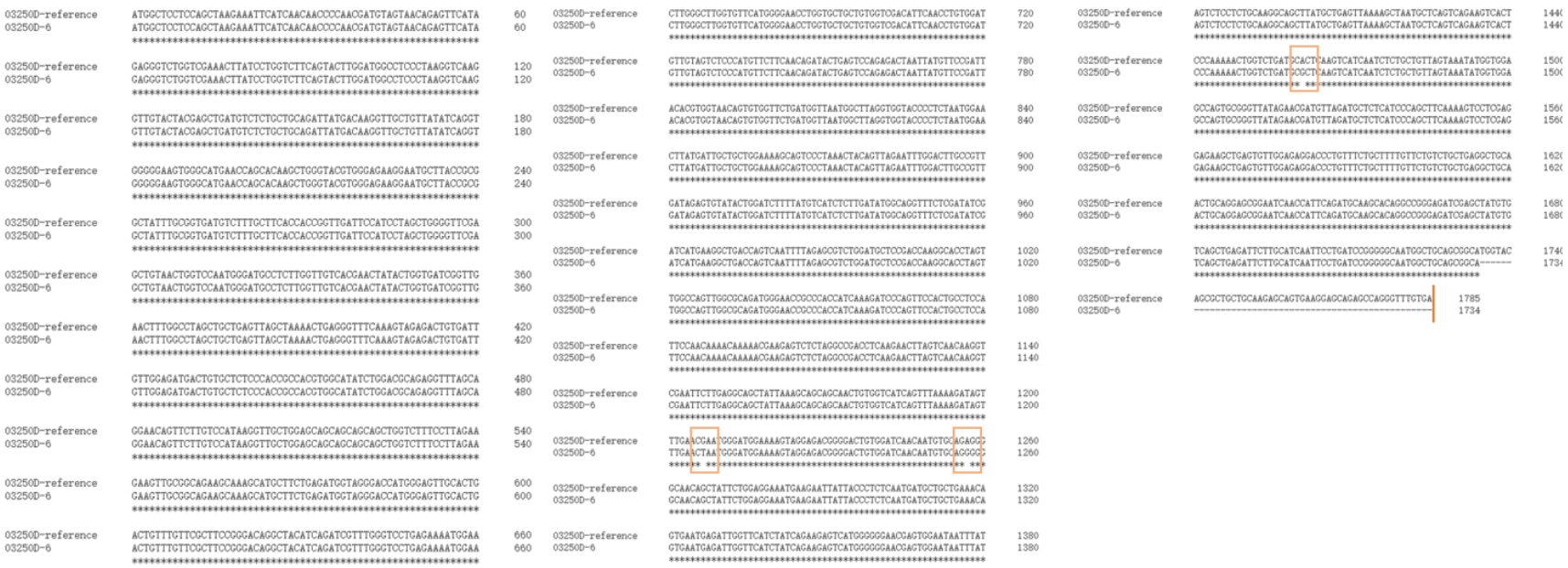


Supplementary Figure S3.1a Multiple sequence alignment of the cloned sequence with promotor CaMV35S sequence provided by NCBI. The region between two bars is the CaMV35S region, which is 100% identical (802 bp).

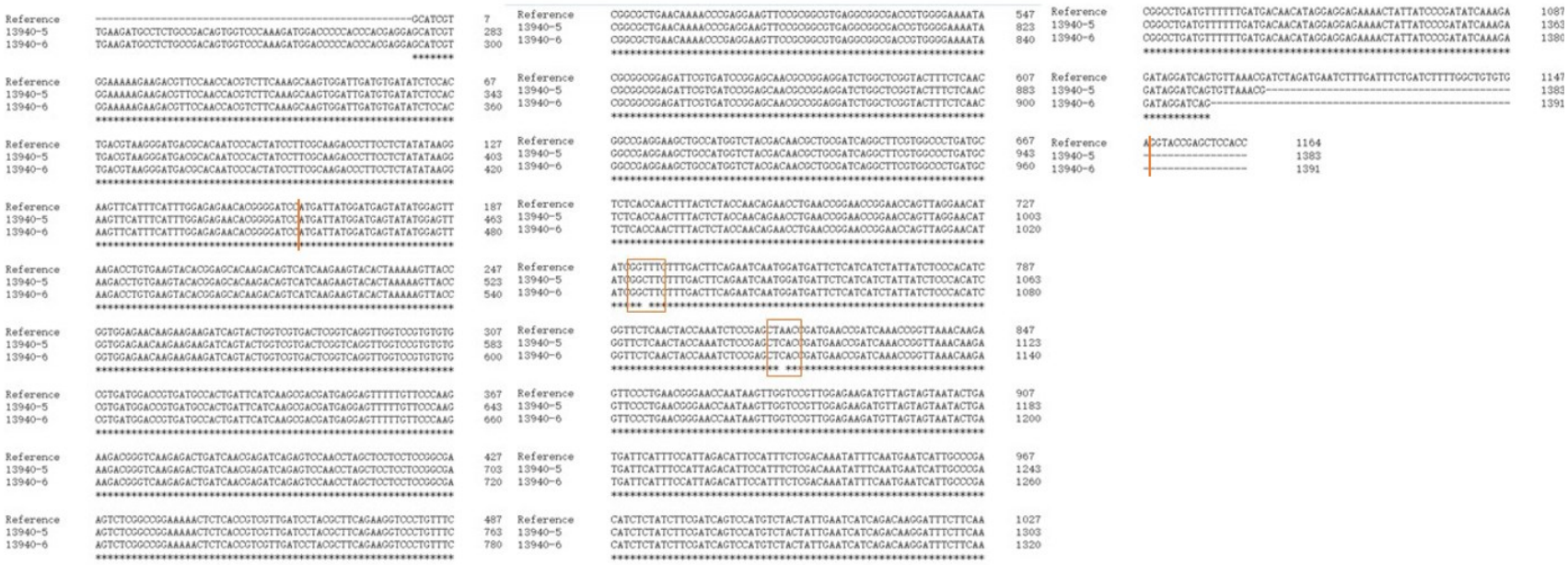
Supplementary Figure S3.1b Multiple sequence alignment of the cloned insertion of the *Lnc-001163* sequence with the *Brassica napus* reference sequence (Chalhoub et al., 2014) using BWA v0.7.12. The region between two bars is the *Lnc-001163* region, which is 96.31% identical to the reference sequence (1976bp). The variants between insertion and reference have been marked with yellow boxes.



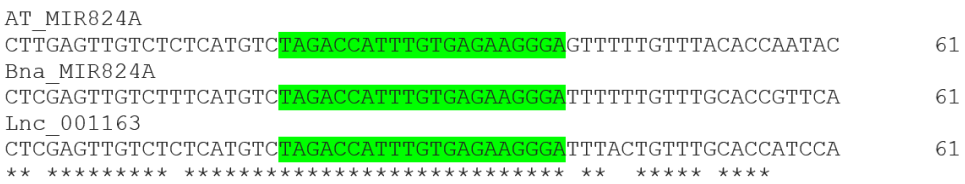
Supplementary Figure S3.1c Multiple sequence alignment of the cloned insertion *Lnc-003848* sequence with the BWA v0.7.12 reference sequence (Chalhoub et al., 2014). The region between two bars is the *Lnc-003848* region, which is 99.7% identical. (648bp). The variants between insertion and reference have been marked with yellow boxes.



Supplementary Figure S3.1d Multiple sequence alignment of the cloned insertion *BnaA08g03250D* sequence with the BWA v0.7.12 reference sequence (Chalhoub et al., 2014). The region between two bars is the *BnaA08g03250D* region, which is 99.8% identical. (1785bp). The variants between insertion and reference have been marked with yellow boxes.



Supplementary Figure S3.1e Multiple sequence alignment of the cloned insertion *BnaA08g13940D* sequence with the BWA v0.7.12 reference sequence (Chalhoub et al., 2014). The region between two bars is the *BnaA08g13940D* region, which is 99.8% identical. (987bp) The variants between insertion and reference have been marked with yellow boxes.



Supplementary Figure S3.2 Multiple sequence alignment of a partial sequence in *Arabidopsis thaliana* gene *AT_MIR824A, Brassica napus* gene *Bna_MIR824A* and the transformed gene *Lnc-001163*. The mature sequences of these genes are highlighted with green colour (Jian et al., 2018).

Please connect with the author for supplementary tables.

Supplementary Table S2.1 Evaluation of the doubled haploid (DH) lines, segregating for clubroot resistance of the rutabaga cv. Polycross, for resistance to different Plasmodiophora brassicae pathotypes.

Supplementary Table S2.2 Information of the bulk of the doubled haploid Brassica napus lines used for whole-genome resequencing (WGRS) for resistance to different Plasmodiophora brassicae pathotypes.

Supplementary Table S2.3a Allele-specific primers of the single nucleotide polymorphisms (SNPs) from QTL region of the chromosome A03 used for genotyping the doubled haploid (DH) lines.

Supplementary Table S2.3b Allele-specific primers of the single nucleotide polymorphisms (SNPs) from QTL region of the chromosome A08 used for genotyping the doubled haploid (DH) lines.

Supplementary Table S2.3c Primer information of the simple sequence repeat (SSR) markers fromQTL region of the chromosome A03 used for genotyping the doubled haploid (DH) lines.

Supplementary Table S2.3d Primer information of the simple sequence repeat (SSR) markers from QTL region of the chromosome A08 used for genotyping the doubled haploid (DH) lines.

Supplementary Table S2.4 List of the putative Brassica napus genes, to be affecting clubroot resistance, used for expression analysis. Their positions, domains, functions in regards to their translated protein, and the reason of inclusion for expression analysis also presented.

Supplementary Table S2.5 Correlation between the replications for resistance to different Plasmodiophora brassicae pathotypes.

Supplementary Table S2.6a The number of SNPs from different chromosomes with Delta_SNP index = -1 calculated based on Brassica napus reference genome Brassica_napus_v3.1 (http://cruciferseq.ca/?q=node/6) kindly provided by Dr. Isobel Parkin, AAFC, Saskatoon, SK .

Supplementary Table S2.6b The number of SNPs from different chromosomes with Delta_SNP index = -1 calculated based on Brassica napus reference genome Brassica_napus_v3.1 (Chalhoub et al. 2014).

Supplementary Table S2.7 The highly resistant (DSI <5%) and resistant (DSI <25%) double haploid (DH) lines carrying marker allele of the resistant parent (R-allele).

Supplementary Table S3.1 Information of primers used in plasmid construction and expression analysis.