Transcriptomic responses in spring canola carrying clubroot resistance introgressed from rutabaga or "Mendel"

by

Aarohi Summanwar

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

Doctor of Philosophy

in

Plant Science

Department of Agricultural, Food and Nutritional Science
University of Alberta

Abstract

Canola provides ~15% of the total global vegetable oil supply and is an important crop contributing about $26.7 billion to Canada's economy. Canola production has been threatened by clubroot disease caused by *Plasmodiophora brassicae* Woronin, which in extreme case can result complete crop failure. Development of clubroot resistant cultivars is one of the important ways of managing this disease. Clubroot resistance has been introgressed into spring canola from various sources of which the resistance introgressed from winter canola cv. 'Mendel' and rutabaga have been used most extensively in Canada; however, the molecular basis including comparative analysis of this resistance has not yet been studied. In this dissertation, I focused and compared thetranscriptomic response of these two resistances in canola in response to infection by *P. brassicae*.

Eukaryote transcriptome is more complex than previously imagined, as evidenced by the knowledge gained from recent studies. The transcriptomes primarily consist of non-coding RNAs such as long non-coding RNAs (lncRNAs) and protein coding RNAs like mRNAs. Using lncRNA-seq, we identified 530 differentially expressed (DE) lncRNAs between clubroot resistant canola lines carrying resistance of rutabaga (RtR) and susceptible lines lacking this resistance (RtS). Similarly, we also identified 464 DE lncRNAs between the canola lines carrying resistance of winter cv. 'Mendel' (MdR) and lines lacking this resistance (MdS). While comparing these two sets of DE lncRNAs, we found 12 lncRNAs were upregulated in both the resistances and could be potentially involved in regulating clubroot resistance. We also found that the predicted targets genes of these lncRNAs were similar in both RtR and MdR resistances. Quantitative real-time (qRT-PCR) analysis was carried out for the lncRNAs specifically

expressed in RtR or MdR lines, and the lncRNAs expressed in both RtR and MdR lines and their respective target genes to confirm their response to clubroot .

Pathway analysis of the target genes regulated by the 530 DE lncRNAs identified by using rutabaga-resistance and 464 DE lncRNAs identified by using Mendel-resistance showed that the primary and secondary metabolic pathways play an important role in resistance to this disease at 10 dpi; the genes involved in primary and secondary metabolic pathways were observed to be upregulated in the RtR and MdR lines. The importance of these pathways was investigated by treating, clubroot susceptible plants with trehalose. The DH lines treated with 30 mM and 60 mM trehalose showed partial resistance to this disease while the treatment with 120 mM trehalose did not change the susceptible phenotype. Additionally, we observed that the expression of the genes like pathogenesis-related protein 2, lipoxygenase and phenylalanine ammonia lyase increased in response to this treatment indicating that the modulation of primary metabolic pathways can affect clubroot resistance.

Furthermore, considering the importance of the clubroot resistance loci of BnaA03 from MdR lines and BnaA08 from RtR lines and their association with clubroot resistance, the DEGs upregulated from these two chromosomes were compared, and the gene cytokinin responsive factor (*CRF4*) was found upregulated in both resistances and, thus, could be involved in mediating clubroot resistance in both RtR and MdR lines. In addition, I also developed lncRNA- and gene-based molecular markers from BnaA08 (RtR) and BnaA03 (MdR). A simple sequence repeat (SSR) marker designed from the lncRNA LNC_000424 showed cosegregation with resistance in a mapping population carrying Mendel-resistance; this lncRNA was upregulated in the MdR lines and was predicted to regulate the target genes involved in plant defense. Additionally, the gene-based SSR markers from the DEGs viz. BnaA03g41300D and

BnaA03g44400D of BnaA03 upregulated in the MdR lines and SSR from the DEG BnaA08g03250D of BnaA08 upregulated in the RtR lines showed showed linkage association with resistance; however, these lncRNA- and gene-based markers showed 1-3% recombination between the marker and resistance phenotypes.

Preface

A version of Chapter 2 of this dissertation has been published as:

Summanwar, A, Basu, U, Rahman, H, & Kav, N. (2019) Non-coding RNAs as emerging targets for crop improvement: a review. Plant Science. Doi: 10.1016/j.plantsci.2020.110521

A version of Chapter 3 of this dissertation has been published as:

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, 1360-77. Doi: 10.1094/MPMI-12-18-0341-R

A version of Chapter 4 has been submitted as a manuscript for publication to Genome as:

Summanwar, A, Basu, U, Kav, N, & Rahman, H. (2020) Identification of lncRNAs in response to infection by *Plasmodiophora brassicae* in *Brassica napus* and development of lncRNA-based SSR markers.

A version of Chapter 5 has been submitted as a manuscript for publication to Physiological and Molecular Plant Pathology as:

Summanwar, A, Farid, M, Basu, U, Kav, N, & Rahman, H. (2020) Comparative transcriptome analysis of canola carrying clubroot resistance from 'Mendel' or Rutabaga and the development of molecular markers.

In Chapter 5, The gene-based SSRs were genotyped by Dr. Mehdi Farid.
I was responsible for conducting all experiments, analysis and interpretations of data and presentation of the results in the manuscripts of all the studies presented in this dissertation experiments after considering feedback from Dr. Kav, Dr. Urmila and Dr. Rahman. Dr. Kav, Dr. Urmila and Dr. Rahman reviewed and edited draft versions of these manuscripts.

All the raw sequencing data generated in this study are available under the NCBI Bioproject accession number PRJNA662499. The SRA accession id's for all the samples are from SRR12615531 to SRR12615542.

Dedication

This is for you, Ajoba

My grandfather, for the encouragement to pursue Science. For always being my guardian angel

Acknowledgments

I would like to express my gratitude to my supervisors Dr. Nat Kav and Dr. Habibur Rahman for giving me the opportunity to work with them, for believing in me and my abilities and for their constant encouragement and support. I would also like to thank Dr. Urmila Basu for her support, guidance, critical evaluations, and valuable suggestions for the duration of the program. I would like to thank my supervisory committee members: Dr. Nat Kav, Dr. Habibur Rahman, Dr. Urmila Basu and Dr. Stephen Strelkov for their valuable advices for this research program.

My sincerest thanks to Dr. Mehdi Farid, Dr. Rubeena Shaikh and Dr. Swati Megha for their valuable guidance, constant encouragement, and emotional support for the duration of the program. It has been a pleasure to share this experience with you all.

I would like to extend my sincerest gratitude to my colleagues Dr. Harleen Kaur, Dr. Berisso Kebede, Ann Vo and my lab mates for their support.

I would like to thank Dr. Uwe Hacke and Dr. Joselyn Ozga, for their teaching assistant opportunities, which helped me increase my work experience and gain new knowledge.

I would like to thank Robin Miles and Nikki Scott for employing me part-time.

I would like to thank Financial support from Alberta Crop Industry Development Fund (ACIDF), the Alberta Canola Producers Commission (ACPC), Alberta Innovates (AI), Alberta Agriculture and Forestry (AAF), Agriculture and Agri-Food Canada (AAFC), and the Natural Sciences and Engineering Research Council (NSERC) for financial support of this project.

I would like to thank my husband, Swaraj for his incessant love, care, and moral support. Thanks for being there when I needed all the help.

I would like to thank my Dad for his love, constant encouragement, moral and emotional support extended all the way from across the world, at every step of this long journey. Thanks, Mom, for the meals at work, staying in the university for all the late-night work and for her endless faith in me and my determination. This journey would have been difficult without you both.

Thank you, to my entire family and my Aai and Baba for their blessings and support.

Table of Contents

List of Tables

List of Figures

List of Abbreviation

Canola (DH plants) carrying resistance from rutabaga	RtR
Canola (DH plants) lacking the resistance from rutabaga	RtS
Canola (DH plants) carrying resistance from winter cv. “Mendel”	MdR
Canola (DH plants) lacking the resistance from winter cv. “Mendel”	MdS
Cycle threshold	Ct
Differentially expressed genes	DEGs
Differentially expressed lncRNAs	DE lncRNAs
Double haploid	DH
Days post inoculation	Dpi
Disease severity index	DSI
Ethylene	ET
Effector triggered immunity	ETI
Endogenous target mimics	eTM
Fragments Per Kilobase pair of exon model per Million fragments mapped	FPKM
Gene Ontology	GO
Hours post inoculation	HPI
Jasmonic acid	JA
Kyoto Encyclopedia of Genes and Genomes	KEGG
Long non-coding RNAs	lncRNAs
Marker assisted breeding	MAB
Marker assisted selection	MAS
Minimum Free Energy	MFE
MicroRNA	miRNA
Natural antisense transcripts	NATs
Next-Generation Sequencing	NGS
Non-coding RNAs	ncRNAs
PAMP-triggered immunity	PTI
Quantitative trait loci	QTL
Recombination frequency	RF

Resistant plants/lines	R plants/lines
RNA integrity number	RIN
Salicylic acid	SA
Small interfering RNA	siRNA
SSR	Simple Sequence Repeats
Susceptible plants/line	S plants/lines

Chapter 1: General Introduction

The global population is projected to exceed 9.7 billion by 2050 and could peak to nearly 11 billion around 2100 (United Nations 2019). Considering the current rate of yield increases of ~ 0.9% to 1.6% per year in important agricultural crops (Ray et al. 2013); it will be challenging to feed the roughly 40-86 million increased population in each year until 2050 (Fyles and Madramootoo 2016). Moreover, over the last century, the average global temperature has already increased by 0.74°C (Chakraborty and Newton 2011) and according to the Intergovernmental Panel on Climate Change, a rise in temperature of ~2°C over the next century is anticipated (Pachauri and Reisinger 2007). Such changes in climate can also influence the dynamics of plant-pathogen interactions, which may result in adverse effect on agricultural productivity (Chakraborty et al. 2000). To address this increasing demand for food arising due to population growth, a tailored solution will be needed in the agri-food sector. Development of high-yielding cultivars with stability of yield and resistance to various biotic and abiotic stresses are some of the important objectives in plant breeding and crop production. Thus, a next generation "Green Revolution" is needed to realise the sustainable food security.

Crops of the Brassicaceae family are one of the oldest cultivated crops (Raymer 2002). The Brassica triangle of U describes the relationships between the three diploid species *Brassica nigra*, *B. oleracea* and *Brassica rapa* and the three amphidiploid species *Brassica carinata*, *Brassica juncea* and *Brassica napus* which evolved through interspecific hybridization between the diploid species. In this, *B. napus* is derived from hybridization between *B. rapa* and *B. oleracea*. *B. napus* canola is one of the most important cultivated oilseed crops of the family Brassicaceae. The oilseed canola is defined as the Brassica oilseed crop species which seed oil contain <2% erucic fatty acid and seed meal contain < 30 μmol glucosinolates per gram oil-free meal (Daun 2011). Canola is known to produce one of the most nutritionally desirable vegetable oils for human consumption (Lin et al. 2013). Canola's contribution to Canada's economy is about $26.7 billion based on the data collected from the crop years 2012/13, 2013/14 and 2014/15 (www.canolacouncil.org; retrieved on 23 May 2020). Canola production has been threatened by abiotic stresses like drought, salinity, extreme temperatures; biotic stresses like pathogens causing major diseases such as clubroot, sclerotinia stem rot, blackleg, alternaria black spot, downy mildew, root rot, and white rust (www.canolacouncil.org; retrieved on 23 May 2020).

Canola production in Canada has been threatened by clubroot, a soil-borne disease of cruciferous crops caused by *Plasmodiophora brassicae* Woronin. It is a fungus like protist which has a complex life cycle and is unable to grow and multiply without a living host (Dixon 2014). Disease development causes formation of galls on the roots, ultimately resulting in the premature death of plants. *P. brassicae* spores can survive in soil for over 20 years in the form of resting spores and germinate in the presence of secretions from the plant roots (Rempel et al. 2014). In infested fields, clubroot disease cause yield loss of 30-100 % in canola (Tewari et al 2005; Strelkov and Hwang, 2014; Ernst et al. 2018; www.canolacouncil.org; retrieved on 23 May 2020). Considering the nature of the pathogenesis and its ability to produce long-living resting spores, makes clubroot,a difficult disease to control. (Dobson et al. 1983; Howard et al. 2010; Kowata-Dresch and May-De Mio 2012). Considering the high value of canola in Canada, excessive losses of canola due to clubroot may have a serious impact on Canadian economy (Rempel et al. 2014). Therefore, despite the existence of preventive resources like commercial clubroot resistant cultivars along with integrated management approaches (www.canolacouncil.org; retrieved on 23 May 2020), it is still important to achieve better understanding of the resistance mechanism, and this may help to develop cultivars carrying durable resistance.

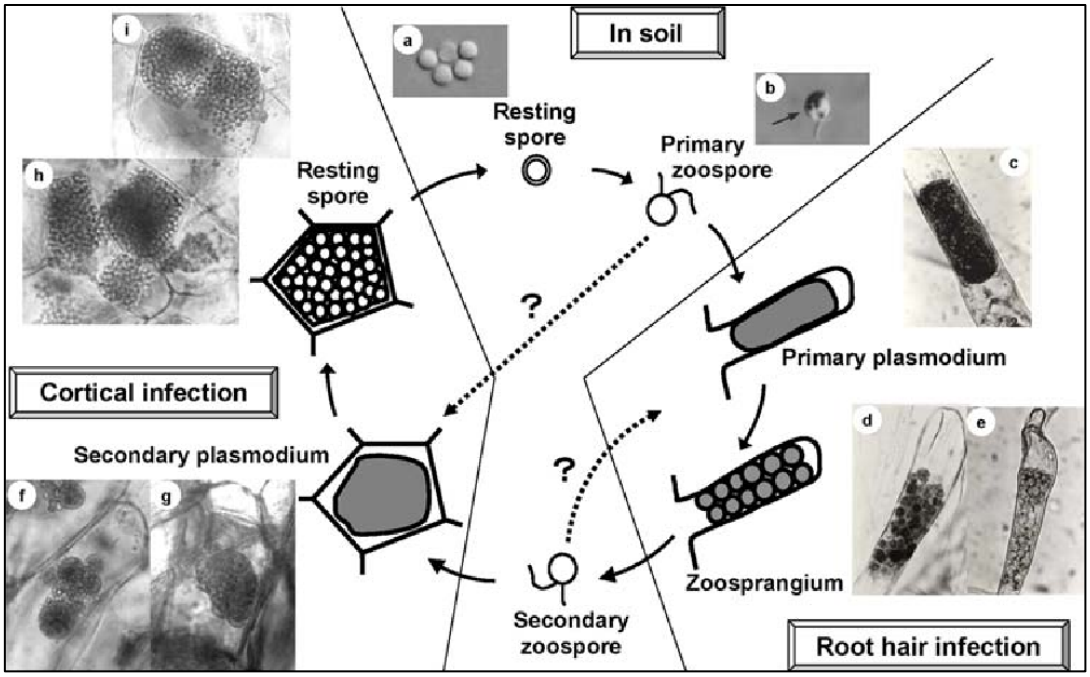


Fig. 1 Life cycle of *Plasmodiophora brassicae*. a) Resting spore. b) Primary zoospore. c) Primary plasmodium in root hair. d) Zoosporangial cluster in root hair. e) Empty zoosporangium. f, g) Secondary plasmodia in cortical cells. h, i) Resting spores in cortical cells. [Image taken from: Kageyama and Asano, 2009; Reproduced with permission from: Springer Nature.]

Clubroot resistant canola cultivars were first registered in Canada in 2009 (Hwang et al. 2017, Strelkov et al. 2018). These cultivars (first-generation hybrids) were resistant to *P. brassicae* pathotype 3 which was predominant in Alberta at that time; this enabled growing of this crop even in fields where the pathogen was well established (Peng et al. 2014, Strelkov and Hwang 2014). However, after growing these resistant cultivars only for about five years, several new pathotypes have been detected in the canola fields; some of them are more virulent than the pathotype 3 and can overcome the resistance gene present in the first-generation hybrids (Strelkov et al. 2018). Therefore, it is important to develop second-generation hybrids with additional resistance genes to make the resistance effective against a wider range of pathotypes.

Plant defence to biotic stresses like clubroot implicates a collective effort involving an array of morphological, genetic, biochemical, and molecular processes. Initial defence by the plants is in the form of physical barriers, but some pathogens can circumvent them by secretion of pathogen-associated molecular patterns (PAMPs). These PAMPs when recognized by plant pattern recognition receptors (PRRs) triggers the secondary resistance mechanism called PAMP-triggered immunity (PTI) (Jones and Dangl 2006). Successful pathogens secrete effectors to

interfere with PTI, which are recognized by the specialized NBS-LRR proteins, resulting in effector-triggered immunity (ETI) (Jones and Dangl 2006). PTI/ETI continuum eventually results in a complex signalling cascade of events resulting in the accumulation of reactive oxygen species, defense hormones (salicylic acid (SA), jasmonic acid (JA), ethylene (ET)), mitogen activated protein kinases, transcription factors and disease resistance proteins. These mechanisms either work constitutively or are induced under pathogen attack to confer tolerance or resistance (Onaga et al. 2016; Ben Rejeb et al. 2014). In addition to these genes expressed during plant defense, plants are also known to have these defense responses to be regulated by the non-coding regions of the genome for instance by long non-coding RNA (lncRNA), microRNA(miRNA) and small interfering RNA(siRNA) (Summanwar et al. 2020 and references therein).

Plants have evolved different strategies to resist or tolerate various types of biotic stresses. Recent advances in the area of NGS sequencing have made it possible to characterize the non-coding regions in the genome. Emerging evidences supports that the lncRNAs play an important role in the regulation of molecular processes like transcription, splicing of transcripts, as well as protein function and localization, in response to growth, development, and responses to abiotic and biotic stresses. Various lncRNAs have been identified to be important regulators of plant immunity and adaptation to environmental stresses (for review, (Nejat and Mantri 2018; Summanwar et al. 2020). A number of studies have identified a large number of lncRNAs in response to biotic stresses; for instance, in wheat against *Blumeria graminis,* in *B. napus* in response to *Sclerotinia sclerotiorum* (Joshi et al. 2016), in tomato against Tomato Yellow Leaf Curl Virus (Wang et al. 2015), in barley against *Fusarium graminearum* (Huang et al. 2016), in rice against *Xanthomonas oryzae* (Yu et al. 2020), in *Populus* in response to *Melampora larici-populina* (Wang et al. 2017), in cotton against *Verticillium dahlia* and *Botrytis cineria* (Zhang et al. 2018), and in *Arabidopsis* against *Pseudomonas syringae* (Seo et al. 2017).

Therefore, considering the complexity of the plant defense responses to biotic stresses, and the economic value of canola to Canada there is a need to further understand the molecular basis of clubroot resistance in this crop. This knowledge would be beneficial for the development of rational strategies aimed for attaining durable clubroot resistant cultivars in the longer term. To achieve this overarching goal, the following objectives were pursued.

Research hypotheses:

1) Transcriptome analysis of clubroot-resistant and -susceptible *B. napus* canola lines, differing for 'Mendel-resistance', in response to infection by *P. brassicae,* will result in the identification of genes located on chromosome A3 that may be potentially involved in clubroot resistance.
2) Transcriptome analysis of clubroot-resistant and -susceptible *B. napus* canola lines, differing for resistance from rutabaga, in response to infection by *P. brassicae* will result in the identification of the genes located on chromosome A8 that may be potentially involved in clubroot resistance.
3) Comparative transcriptome analysis of the canola lines carrying resistance gene located on chromosomes A3 and A8 will identify the genes that could play a fundamental role in disease resistance.
4) Molecular markers developed from the differentially expressed transcripts will show strong association with resistance and will facilitate marker-assisted selection of clubroot resistant canola germplasm.

Chapter 2: Non-coding RNAs as emerging targets for crop improvement

2.1 What is food security/insecurity?

Food security as a concept was introduced in the mid-1970s when international food-related issues were discussed at a time of global food crisis (FAO 2003). Over the years, the interpretation of the term food security has changed, reflecting global concerns regarding the volume and stable supply of food (FAO 2003). In 2018, the World Health Organization (WHO) indicated that about 821 million people in the world are undernourished and are not getting the required quantity of food; therefore, the eradication of food insecurity is of utmost importance.

Currently, the production of global staple food grains (wheat, rice and maize) is about 2 719 Mt (FAO 2020); however, this production is often challenged by different abiotic and biotic stresses, including increased temperature and drought, infestation by insects and diseases. For example, the global mean temperature has increased about 0.74°C over the last century and this increase is attributed to climate change (Chakraborty and Newton 2011). According to the Intergovernmental Panel on Climate Change (IPCC, 2000), an increase in temperature of about 2°C over the next century is inevitable (Pachauri and Reisinger 2007). In addition, changes in global sea levels observed over the past several decades have also been linked to increases in global mean temperature (Church and White 2011). A rise in sea level will continue to result in increased salinity in coastal agricultural regions with negative effects on crop productivity (Alam et al. 2017). In addition, climate change also negatively impacts biodiversity (Torquebiau et al. 2016), which is crucial for sustainable crop production.

Plants are also challenged with numerous biotic stresses, which includes fungi, bacteria, nematodes, and viruses. Effect of climate change, including an increase in temperature will result in changes to the dynamics of plant-pathogen interactions, pest behaviour, increased pathogen population and incidences of diseases and evolution of new pests leading to reduced agricultural productivity (Daugherty et al. 2017). Crop diseases have had devastating effects on agricultural productivity, altering the course of our history. For example, the infamous 19th-century Irish potato famine caused by *Phytophthora infestans* (for review, (Bourke 1993)), the Bengal famine of 1943 caused by *Helminthosporium oryzae* (for review, (Padmanabhan 1973)) or the widespread damage to the grape industry in Europe caused by *Plasmopara viticola* (Gessler et al. 2011). These experiences from the past should constantly remind us to develop cost-effective

and efficient crop-disease management strategies to prevent similar issues in the future that will be exacerbated by climate change. Therefore, the development of high yielding cultivars with increased tolerance to drought, temperature and resistance to pathogens and insect pests is absolutely critical.

2.2 Strategies for introducing resistance to biotic stresses in crop plants

2.2.1 Plant breeding

Plant breeding involves the introduction of genetic variability, primarily through crossing of genetically different parents, followed by selection and fixation of superior traits (for review, (Moose and Mumm 2008)). Conventional breeding has played an important role in the development of disease resistant cultivars. Examples include the introduction of different disease resistance genes, e.g. leaf rust 13 (*Lr13)* and *Lr34* resistance genes in wheat cultivars (German anf Kolmer 1992), and powdery mildew resistance in barley (Pickering et al. 1995). The introduction of new disease resistance genes from diverse sources into crop cultivars is a high priority for plant breeders in order to combat the rapidly evolving races/pathotypes of plant pathogens.

Different molecular tools are currently available (for review, Collard and Mackill 2008) which can enhance the scope and efficiency of classical plant breeding for disease resistance. For instance, construction of high-resolution genetic maps has been used for identification of molecular markers for disease resistance genes for their application in breeding (Fredua-Agyeman and Rahman 2016, Hasan and Rahman 2016). Alternately, bulk-segregant analysis, which does not require the construction of a genetic linkage map, has also been used for identification of markers linked to the simply inherited resistance genes (Zou et al. 2016). Recently, genome sequence information from whole genome resequencing, also called QTL-seq, has been used for mapping of resistance genes and identification of molecular markers (Takagi et al. 2013). Through the application of markers linked to resistance genes, selection for the gene can be made at the DNA level, a technique commonly referred to as marker assisted selection (MAS) (for review, (Collard and Mackill 2008)). In MAS, the accuracy of the selection becomes higher as opposed to the phenotypic selection where the influence of environment on the trait can mislead selection (Collard et al. 2005). Other breeding approaches that use molecular markers include marker-assisted backcrossing, marker-assisted recurrent selection and marker-assisted

gene pyramiding and genomic selection. Some of the examples of the aforementioned marker assisted breeding methods are provided in Table 2.1.

Although conventional breeding approaches have made significant progress over the last decades, it still remains a relatively slow process, which takes almost 8-10 years to develop a cultivar carrying the desired genes. In some cases, linkage drag, arising from the linkage of a desirable gene with an undesirable gene, can pose a challenge for incorporation of only the desired gene into a cultivar (Delourme et al. 1998, Primard-Brisset et al. 2005). Furthermore, pre-, and post-fertilization barriers pose additional challenges in conventional breeding while transferring genes from allied species through interspecific or intergeneric hybridization (Bennett et al. 2008). In order to achieve global food security in the era of population growth and challenges posed by climate change, complementary approaches to conventional plant breeding are necessary.

2.2.2 Genome modification

Some of the aforementioned challenges faced by plant breeders may be overcome though the application of genetic engineering techniques. This approach allows the introduction of any gene(s) from any donor organism into a recipient organism, and has been used to develop and commercialize genetically modified (GM) crop cultivars including soybean, canola, corn and cotton, carrying resistance to insect-pest and/or tolerance to herbicides (for review, (Ricroch and Hénard-Damave 2016)). In contrast to traditional breeding, genetic engineering results in a more precise change in the genome and reduces the time for incorporation of a new gene into an elite line for the development of an improved cultivar in crop plants (for review, (Samal and Rout 2018)). Nevertheless, genetic engineering also does not completely eliminate the need for conventional plant breeding techniques.

In genetic engineering, the desired DNA fragment or a gene is introduced into the plants following either cisgenic or transgenic approach (for review, (Holme et al. 2013)). Some of the crops that have been improved using cisgenic and transgenic approaches are listed in Table 2.2. Cisgenesis involves the introduction of a complete gene including its native promoter and terminator sequences obtained from the same organism or from a sexually compatible gene pool into the same organism, while in transgenesis, these gene elements originate from different organisms. Cisgenesis is a preferred technique for introduction of disease resistance genes into an elite susceptible cultivar from its wild relative, especially when incorporation of multiple

resistance genes is needed (for review, (Schaart et al. 2016)). The cisgenic plants, therefore, should be considered a non-GMO and similar to plants developed through conventional breeding approaches (for review, (Schouten et al. 2006)). Therefore, there is a very little chance to alter the desired combination of the genes of the elite recipient parent. In this technique, the time required for the development of new cultivars is significantly reduced.

Recently, genome editing has emerged as a promising technique for crop improvement because of its accuracy and its ability to alter gene function rapidly. Genome editing involves the application of sequence-specific nucleases like zinc-finger nucleases (Kim et al. 1996), transcription activator-like effector nucleases (TALENs) (Boch and Bonas 2010, Bogdanove et al. 2010) and clustered regularly interspersed short palindromic repeats (CRISPR/Cas9) (Lei et al. 2013). This technique has been applied for the improvement of disease resistance in different crops and have been reviewed by others (Table 2.2).

Recent technological advances in the area of massively parallel sequencing have made it possible to characterize non-coding regions in the genome. Emerging evidence supports the notion that these non-coding regions play an important role in the regulation of molecular processes like transcription, splicing and protein function and localization, in response to plant growth, development, and response to abiotic and biotic stresses. Various non-coding RNAs (ncRNAs) have been identified to be important regulators in eukaryotes and are suggested to be involved in plant immunity and adaptation to environmental stresses (for review, (Nejat and Mantri 2018)). In this review, we have focussed on the recent advances on the regulatory function of the long non-coding RNAs (lncRNAs), current understanding and knowledge gaps of the mechanism of action of the lncRNAs and their emerging roles in the regulation of immune responses in plants against various biotic stresses.

2.3 Transcriptome and the non-coding RNAome

In eukaryotes, transcription is more complex than previously envisioned, as evidenced by the knowledge gained from recent advances (Kornberg 2007). The cellular transcriptome is defined as a complete set of mRNA transcripts, the quantity of these transcripts, and their expression in a specific tissue and at a specific developmental stage or a physiological condition (for review, (Wang et al. 2009)). About 98% of the genome was traditionally referred as the "transcriptional noise" or "junk DNA" or "experimental artifacts" (for review, (Nejat and Mantri 2018, Wilusz et al. 2009)); however, an in-depth examination of the non-coding portion of the

genome using genome sequencing and bioinformatic analysis has revealed the existence of thousands of transcripts with properties like mRNAs but with no protein coding potential, known as ncRNAs (for review, (Nejat and Mantri 2018, Karlik et al. 2019)).

ncRNAs can either have housekeeping or regulatory roles. For instance, the catalytic or structural roles in a cell are generally played by the housekeeping ncRNAs which include transfer ribonucleic acids (tRNAs) and ribosomal ribonucleic acids (rRNAs). On the other hand, recently, many regulatory ncRNAs have also been reported to play important roles in a variety of biological processes (for review, (Nejat and Mantri 2018)). Based on their size and shape, regulatory ncRNAs can be classified into three subclasses: i) lncRNAs (>200 nt), ii) small ncRNAs which include microRNAs (miRNAs, 20-24 nt), small interfering RNAs (siRNAs, < 200 nt), small nuclear RNA (snRNAs) and small nucleolar RNAs (snoRNAs), and iii) circular ncRNAs (circRNAs) (for review, (Karlik et al. 2019, Ohtani 2018) and references therein). These ncRNAs and their roles in mediating plant responses to biotic stresses are discussed in subsequent sections.

2.4 Long noncoding RNAs (lncRNAs)

Although lncRNAs in plants have been studied for over 20 years (Ariel et al. 2014, Campalans et al. 2004, Ding et al. 2012, Heo and Sung 2011,Seo et al. 2017, Swiezewski et al. 2009, Wang et al. 2014), the details surrounding their roles in plant development, reproduction and in mediating responses to biotic and environmental stress by activating different signaling cascades still remain elusive (for review, (Matsui and Seki 2019)). Some of the major modes by which lncRNAs mediate their regulatory effects are through DNA methylation and acetylation, architectural modifications of DNA by nucleosome repositioning, and by RNA-RNA interactions (for review, (Wilusz et al. 2009, Matsui and Seki 2019)).

lncRNAs are known to regulate protein-coding genes either through cis- or trans-acting processes (for review, (Kornienko et al. 2013)). As the name suggests, the cis-acting lncRNAs regulate the expression of the neighbouring genes while the trans-acting lncRNAs regulate the expression of the genes that are located further away from their transcription site (for review, (Kornienko et al. 2013)). Based on their positional relationship with the adjacent protein coding genes, lncRNAs can be classified into: i) long intergenic lncRNAs (lincRNAs), ii) intronic lncRNAs (incRNAs), iii) natural antisense transcripts (NATs), iv) promoter upstream transcripts

(PROMPTs) and v) bidirectional lncRNAs (BI-lncRNAs) like enhancer RNAs (eRNAs) (for review, (Wu et al. 2017)).

2.4.1 Mechanisms of action of plant lncRNAs

In plants, various mechanisms for the expression of lncRNAs have been reported and expression of these lncRNAs has been found to occur in both spatial and temporal manner. Most well characterized lncRNAs such as *COLD INDUCED LONG ANTISENSE INTRAGENIC RNA* (*COOLAIR*)*, COLD ASSISTED INTRONIC NONCODING RNA* (*COLDAIR*) *and LONG-DAY-SPECIFIC MALE-FERTILITY-ASSOCIATED RNA* (*LDMAR*) are transcribed in a manner similar to that of mRNAs, i.e., by RNA Polymerase II (Pol II) and can undergo similar post-transcriptional modifications including 5' capping, splicing and polyadenylation (Ariel et al. 2014, Heo and Sung 2011, Swiezewski et al. 2009, Matsui and Seki 2019). Some of the lncRNAs can be transcribed in a manner similar to rRNAs and tRNAs by RNA Polymerase I (Pol I) and RNA Polymerase III (Pol III), respectively. Majority of the lncRNAs are lincRNAs which are transcribed by Pol II (for review, (Bhatia et al. 2017)). Some lncRNAs can also be transcribed by RNA Polymerase IV (Pol IV) and RNA Polymerase V (Pol V). The ncRNAs transcribed by Pol IV are responsible for generating 21-24 nt siRNAs and the lncRNAs transcribed by Pol V are responsible for guiding these siRNAs to the target loci where they are involved in the modification of chromatin structure by recruiting multi-subunit complexes for DNA methylation, via the RNA-directed DNA methylation (RdDM) pathway also known as, a plant de novo DNA methylation process ((Blevins et al. 2015), (Zheng et al. 2009), for review, (Matzke and Mosher 2014)). In spite of an increasing number of studies in plants, our knowledge on the specific mechanisms involved in lncRNA-mediated regulation of gene expression is still limited. Nevertheless, in this review we have classified the lncRNAs into four categories based on their known modes of action of transcriptional interference mediated by i) chromatin modification complexes, ii) RNA-DNA hybridization, iii) RNA-Protein interaction, and iv) RNA-RNA interaction, and summarized the recent results.

2.4.1.1 Transcriptional interference mediated by chromatin modification complexes

2.4.1.1a) Transcriptional interference by histone modifications: lncRNAs play an important role in plant epigenetic processes (Heo and Sung 2011). In *Arabidopsis*, prolonged cold has been found to epigenetically silence the *FLOWERING LOCUS C (FLC)* in a Polycomb-

mediated process called vernalization, which silences the floral repressor thereby promoting flowering in spring (Heo and Sung 2011). An intricate mechanism modulated by three lncRNAs, *COOLAIR, COLDAIR* and *COLDWRAP* establishes and maintains the *FLC* repression by implementing chromatin modifications and by stabilizing RNA-DNA interactions (Heo and Sung 2011, Swiezewski et al. 2009, Kim and Sung 2017) . *COOLAIR* is an antisense transcript, which originates from the region immediately to the 3' untranslated region of the *FLC* locus and terminates either early or extends into the *FLC* transcript and is critical for the initiation of the process for repression of the *FLC* locus (Swiezewski et al. 2009). Induction of *COOLAIR* occurs earlier than *COLDAIR* and *COLDWRAP* and transiently reduces the expression of *FLC* mRNA (Heo and Sung 2011) (Fig. 2.1A). During vernalization, *COOLAIR* is transiently induced by its cold inducible promoter and epigenetically silences the expression of the *FLC* and the maximum level of expression of *COOLAIR* occurs two weeks after exposure to cold (Swiezewski et al. 2009). *COOLAIR* functions by synchronized removal of H3 Lys36 trimethylation (H3K36me3) and accumulation of H3 Lys27 trimethylation (H3K27me3) at the *FLC* locus (Csorba et al. 2014) (Fig. 2.1A). In contrast, Helliwell et al. (2011) suggested that the expression of *COOLAIR* is not critical for vernalization. Csorba et al. (2014) showed that COOLAIR is responsible for coordinated switching of chromatin states and that its loss disrupted the synchronization process, and thus slowing down the repression of the *FLC*.

In addition to the vernalization pathway, *COOLAIR* is also involved in the autonomous pathway of flowering and can modulate the expression of *FLC* at ambient temperature (Ietswaart et al 2012, Liu et al. 2010). Two polyadenylated isoforms, *AS I* (~400 nt) and *AS II* (~750 nt), of *COOLAIR* are generally produced from the *FLC* locus through alternate splicing, and they are involved in the regulation of *FLC* expression at ambient temperature (Hawkes et al. 2016). For *COOLAIR* to participate in the autonomous pathway, proximal polyadenylation of the *COOLAIR* transcripts has to be promoted by flowering time control proteins like FCA and FPA, the 3' processing factors such as cleavage stimulating factor subunit 64 (Cstf64), Cstf77, the flowering time control protein FY, the cleavage and polyadenylation specificity factor (CPSF) and the pre-mRNA-processing-splicing factor 8 (PRP8) for the recruitment of protein flowering locus D (FLD) (H3K4me2 demethylase); all of which leads to quantitative downregulation of *FLC* expression. Similarly, distally polyadenylated forms are known to promote *FLC* upregulation (Ietswaart et al. 2012, Liu et al. 2010).

COLDAIR is a 5' capped, non-polyadenylated lncRNA transcribed by RNA Pol II in the sense direction from the first intron of the *FLC* locus (Heo and Sung 2011); Fig. 2.1B). *COLDAIR* acts in a *cis* manner during vernalization and its expression is elevated at the site of its transcription on the *FLC* locus. Prolonged exposure to cold induces the expression of VERNALIZATION INSENSITIVE 3 (VIN3) protein (a plant homeo domain (PHD) finger containing protein), which interacts with the Polycomb repressive complex 2 (PRC2) (Fig. 2.1B), and this is responsible for accumulation of the epigenetic repressive marks H3K27me3 through a histone methyltransferase, CURLY LEAF (CLF), a core component of PRC2 (Heo and Sung 2011). At the end of vernalization and in warm temperatures, the PRC2-VIN3 complex recruited by *COLDAIR* spreads the H3K27me3 across the *FLC* locus ((Angel et al. 2011) and references therein), and thus stabilizes the epigenetic repression at this locus and promotes flowering. Increase in the recruitment of PRC2-VIN3 and H3K27me3 at the *FLC* locus is a direct result of the elevated expression of *COLDAIR* and reduction of H3K36me3 by *COOLAIR* ((Heo and Sung 2011); Fig. 2.1B & C).

The interaction between *COLDAIR* and PRC2-CLF is also mediated by a stem-loop structure of *COLDAIR*, demonstrating the functional importance of the secondary structures of lncRNAs (Kim and Sung 2017). Besides *COLDAIR,* stable repression of *FLC* also requires a non-polyadenylated *COLDWRAP* lncRNA which is expressed from the repressed promoter region of the *FLC* during vernalization in *Arabidopsis* ((Kim and Sung 2017); Fig 1B). Similar to *COLDAIR*, *COLDWRAP* also associates with PRC2-VIN3 complex to maintain the repression of *FLC* ((Kim and Sung 2017); Fig. 2.1B). Both lncRNAs, *COLDAIR* and *COLDWRAP* along with PRC2 are responsible for the formation of a repressive intragenic gene loop at the *FLC* locus which is critical for the silencing of *FLC* (Heo and Sung 2011, Kim and Sung 2017) . These three lncRNAs are well-studied examples of lncRNAs associated with the transcriptional interference mediated by histone modifications in plants. *COLDWRAP* and *COLDAIR* has still not been found beyond *Arabidopsis*. Although, *COOLAIR* is highly expressed and shows conservation of secondary structure, despite low sequence similarity across the Brassicaceae, (Hawkes et al. 2016, Castaings et al. 2014, Li et al. 2016). Recently, Jiao et al (2019), showed that position of *COOLAIR* is conserved within the *FLC* homologs across monocot species such as *Brachypodium distachyon*. However, it is still unclear as to how PRC2 is recruited and guided by lncRNAs to the *FLC* locus. It is possible that additional mechanisms are involved in the

regulation of the expression/repression of *FLC* under different environmental conditions. Plant development including flowering can also be affected by biotic factors like pests and pathogens. Some studies have identified the association between disease resistance and flowering time by conducting QTL analyses (Pinson et al. 2010, Van Inghelandt et al. 2012, Mizobuchi et al. 2013). For instance, early flowering time is positively associated with increase in susceptibility to *Verticillium spp.* (Veronese et al. 2003), while late flowering is associated with resistance to *Fusarium oxysporum* (Lyons et al. 2015). It will be fascinating to investigate whether the overexpression of these cold induced lncRNAs are associated with the innate immunity in plants against diseases. To the best of our knowledge, there are no studies which have established the involvement of ncRNAs in crosstalk between biotic stress and flowering and plant development; this could be worth exploring in future.

2.4.1.1b) Transcriptional interference by DNA methylation: lncRNAs can contribute to epigenetic silencing via the RdDM pathway, which requires the lncRNAs transcribed by Pol IV and V as well as Pol II-derived lncRNA at same loci (Zheng et al. 2009). *LONG-DAY-SPECIFIC MALE-FERTILITY-ASSOCIATED RNA (LDMAR)* is a 1236 nt long lncRNA which regulates the Photoperiod-Sensitive Male Sterility (PSMS) in the rice cultivar Nongken 58S (NK 58S) which is a spontaneous mutant of the line Nongken 58N (NK 58N) (Ding et al. 2012). Both NK 58S and NK 58N require high expression of *LDMAR* for reproductive development under a long day condition. As illustrated in Fig. 2.1E, a change from C to G alters the structure of *LDMAR* in NK 58S, affecting the methylation levels at the entire *LDMAR* locus and thereby repressing the expression of *LDMAR* ((Ding et al. 2012a, 2012b); Fig. 2.1F). In the case of NK58S, Ding et al. (2012b) identified a siRNA, named *Psi-LDMAR,* derived from another transcript *AK111270* which was observed to overlap with the sense strand of the promoter of *LDMAR* (Fig. 2.1E). *Psi-LDMAR* was found to be expressed abundantly in this male sterile line as compared to NK 58N. It has also been suggested that repression of *LDMAR* transcripts in NK 58S is mediated by *Psi-LDMAR* through the RdDM pathway which causes an increase in methylation in the *LDMAR* promoter region (Ding et al. (2012b); Fig. 2.1E). Low *LDMAR* levels result in premature cell death in the developing anther which induces male sterility in NK 58S plants under a long day condition (Ding et al. 2012a). This example also indirectly links the expression of lncRNAs with RdDM-related siRNAs. It seems plausible from the involvement of

Psi-LDMAR in the RdDM pathway, that it could be a product of transcription by Pol IV. The involvement of RdDM pathway has also been implicated in gene activation (Au et al. 2017). According to Au et al. (2017) argonaute (AGO) is the key component of the RdDM pathway as this bind with both siRNA and lncRNA. This phenomenon has been exploited to identify the lncRNAs associated with AGO through RNA immunoprecipitation (Au et al. 2017). To understand if the target genes activated by these lncRNAs are also involved in biotic stress response, Au et al. (2017) used the transcriptomic data of *Arabidopsis* in response to *F. oxysporum* infection. It would be interesting to characterize the lncRNAs associated with the RdDM pathway in response to a pathogen attack, to understand their mechanism in regulating gene expression that results in disease resistance. It is well reported that DNA methylation levels or modifications of histones can facilitate the transcription of defense-related genes (for review, (Mauch-Mani et al. 2017, Wiliams and Gehring 2017)) and extensive changes in DNA methylation patterns have been identified in response to biotic stress in plants (Dowen et al. 2012, López Sánchez et al. 2016, Ramirez-Prado et al. 2018, Alonso et al. 2019). It will be particularly interesting if these responses could be primed for stable inheritance in subsequent generations, thus, giving rise to progeny better adapted to biotic stress. In studies on same-generation-primed plants and their descendants for resistance against virulent *Pseudomonas syringae in Arabidopsis,* it has been suggested that the alterations in the DNA c-methylation levels and their inheritance play a role in the transgenerational inheritance of priming (Slaughter et al. 2012). Future research should focus on investigation of the mechanism of action of the pathogen-responsive lncRNAs facilitated by RdDM pathway resulting in epigenetic changes in DNA methylation and/or chromatin modifications.

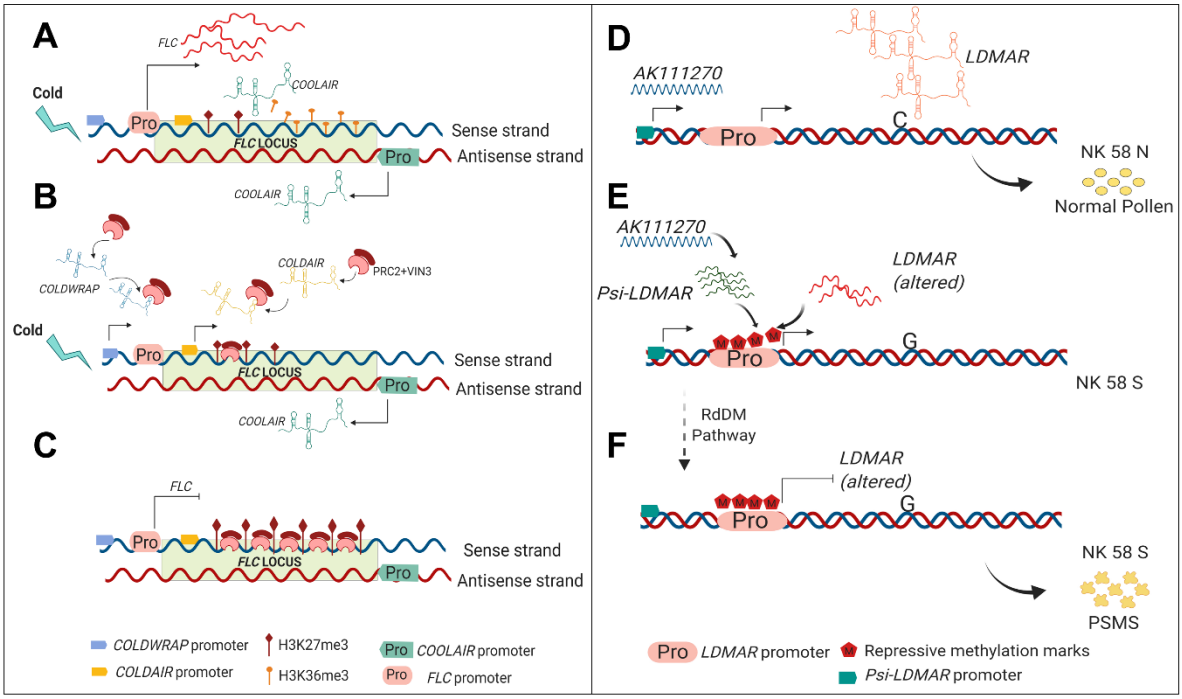


Fig. 2.1 Modification of the chromatin architecture by histone modification (A,B,C) and DNA methylation (D,E,F). A) Regulation of *FLC* locus by three lncRNAs. *FLC* is actively transcribed before exposure to cold. On exposure to cold, *COOLAIR* accumulates and synchronizes the removal of H3K36me3 and accumulation of H3K27me3. B) Next, *COLDAIR* and *COLDWRAP* are transcribed and recruit PRC2+VIN3 which are responsible for deposition of additional H3K27me3. C) Complete repression of *FLC* is achieved with H3K27me3 repression marks on *FLC* locus. D) *LDMAR* regulates reproductive development under a long day condition in rice cv. NK 58N. E) *LDMAR* regulates PSMS in mutant line NK 58S. A spontaneous mutation from C to G in the mutant line NK 58S produces altered *LDMAR* which affects the level of methylation at the promoter of *LDMAR*. An increased methylation is also the result of the RdDM pathway via *Psi-LDMAR* derived from *AK111270*, a transcript found to be overlapping the sense strand of the promoter of *LDMAR*. F) This results in premature death of pollen caused by PSMS. (*FLC*: Flowering Locus C; PRC2+VIN3: Polycomb repressive complex 2 + VERNALIZATION INSENSITIVE 3; *COOLAIR: COLD INDUCED LONG ANTISENSE INTRAGENIC RNA; COLDAIR: COLD ASSISTED INTRONIC NON-CODING RNA; LDMAR: LONG-DAY SPECIFIC MALE FERTILITY ASSOCIATED RNA*; RdDM: RNA dependent DNA Methylation; PSMS: Photoperiod Sensitive Male Sterility)

2.4.1.2 Transcriptional regulation mediated by RNA-DNA hybridization

lncRNAs can also regulate the dynamics of chromatin topology and thus affect the pattern of gene expression. A lincRNA called *AUXIN REGULATED PROMOTER LOOP RNA (APOLO)* has been reported in *Arabidopsis*. *APOLO* is transcribed by both Pol II and Pol IV/V from 5 kb upstream of the locus *PINOID* (*PID*), an important regulator of the polar auxin transport (Ariel et al. 2014). *APOLO* controls the dynamics of the chromatin loop and regulates the promoter activity of its adjacent gene PID (Fig. 2.2A). By influencing the local chromatin topology, *APOLO* adds a new layer of regulation in gene expression (Ariel et al. 2014). This is accomplished by generating a scaffold by that recruit's complexes involved in chromatin modification without involvement of siRNAs (Matsui and Seki 2018). As illustrated in Fig. 2.2A, the phytohormone auxin activates the demethylation of lncRNA locus *APOLO*, mediated by RDD (repressor of silencing 1 (ROS1), demeter-like protein 2 and 3 (DML2) and (DML3)), which then results in an opening of the chromatin loop at the promoter region of *PID*, making it accessible to Pol II (Ariel et al. 2014, Liu et al. 2015). With decreased repressive marks (H3K27me3) and increased active marks (H3K9Ac), the chromatin loop opens and transcription by Pol II results in the accumulation of *PID* and *APOLO* transcripts (Fig. 2.2A). The *APOLO* transcripts which are transcribed by Pol II subsequently recruit LIKE HETEROCHROMATIC PROTEIN 1 (LHP1) for the formation of chromatin loops and PRC2 to presumably redeposit repressive marks (H3K27me3); whereas the *APOLO* transcripts which are transcribed by Pol IV/V recruit AGO4 protein complex triggering siRNA-mediated DNA methylation of *APOLO*/*PID* (Ariel et al. 2014). Thus, *APOLO*-LHP1 and PRC2 mediate the formation of chromatin loop, which is then maintained by the DNA methylation, a process that is induced by *APOLO* transcripts produced by RNA-Pol IV/V, thereby regulating the expression of *PID* (Ariel et al. 2019). Recently, Ariel et al. (2020) showed that *APOLO* regulates plethora of auxin-responsive genes in trans by the formation of RNA-DNA hybridization. These studies clearly demonstrate that the lncRNA *APOLO* and double stranded DNA can interact to form an RNA-DNA hybrid to modulate the expression of neighbouring protein coding genes as well as distally located genes (Ariel et al. 2014, 2020, Matsui and Seki 2018). Interestingly, a recent study suggested that overexpression of *PID* gene causes perturbations in the levels of auxin and other hormones and reactive oxygen species (ROS) causing an accumulation of antioxidants, and thus, modify the growth at rosette stage and affects stress responses (Saini et al. 2017). *APOLO*

lncRNA has been known to positively regulate the *PID* gene by modulating chromatin architecture to form an inhibitory loop and by DNA methylation (Ariel et al. 2014). Considering these results, we could implicitly suggest that *APOLO* lncRNA might regulate the responses against pathogen attacks. Previous studies have already shown the importance of changes in chromatin architecture by modulations in DNA methylation events in response to pathogen attack (Dowen et al. 2012, Ramirez-Prado et al. 2018). It will be interesting to investigate if the *APOLO* lncRNA is involved in the regulation of responses against pathogens and could be exploited for crop improvement against biotic stress.

Another example of lncRNA regulating gene expression is *HIDDEN TREASURE 1 (HID1),* a 236 nt lncRNA from *Arabidopsis*. *HID1,* a trans-acting lncRNA, is assembled into an unknown nuclear protein-RNA complex which binds to the chromatin of the first intron of *PHYTOCHROME INTERACTING FACTOR 3 (PIF3)* and represses the transcription of *PIF3* (Wang et al. 2014). *PIF3* is one of the key repressors in photomorphogenesis that modulates responses to light in *Arabidopsis* (Wang et al. (2014) and references therein). Interestingly, *HID1* mediates photomorphogenesis only under red light conditions and the function of *HID1* is dependent on the conserved secondary structure of this lncRNA (Wang et al. 2014). For instance, stem-loops 2 and 4 in *HID1* are highly conserved in plants and are known to have important biological functions in *Arabidopsis* (Wang et al. 2014). It will be interesting to study the cause of interaction between the *HID1* and the nuclear protein complex and understand the role of the stem-loops of *HID1* in protein binding and target recognition in order to further delineate its mechanism in mediating plant responses to light.

Transcriptional regulatory mechanisms mediated by RNA-DNA hybridization have been uncovered in the regulation of *FLC* by Sun et al. (2013) in which the transcription of *COOLAIR* in *Arabidopsis* was found to be mediated through the formation of RNA-DNA heteroduplex (R-loop) at the 3’ end of the *COOLAIR* promoter. This research group also reported that R-loop forms naturally in the promoter of *COOLAIR* after which a nodulin homeobox protein, AtNDX, associates with a non-template single stranded DNA resulting in the stabilization of R-loop which is necessary for limiting the transcription of *COOLAIR* (Sun et al. 2013). However, it is not yet clear which factors are involved in the regulation of R-loop formation; future studies will be needed to delineate this.

2.4.1.3 Transcriptional regulation mediated by RNA-protein interaction

Another known mechanism of transcriptional regulation by lncRNAs is through their binding to specific proteins, which allows the formation of larger RNA-protein complexes and alters the localization of proteins in the cell. An example of this type of regulation is the *EARLY NODULIN 40* (*ENOD40)* which is one of the first identified lncRNAs in plants and is highly conserved in legumes (Gultyaev and Roussis 2007), rice (Kouchi et al. 1999) and maize (Compaan et al. 2003). It is involved in the regulation of symbiosis between bacteria or fungi in leguminous plants for organogenesis in root nodules (Gultyaev and Roussis 2007), as well as in mycorrhizal association in alfalfa (Rhijn et al. 1997). During such symbiotic interactions in alfalfa, the expression of *ENOD40* was observed to be induced by soil rhizobia in the root pericycle and in the cortical cells of nodule primordia (Compaan et al. 2001). *ENOD40* has been reported to have a bi-functional role in the organogenesis of nodules by producing two short peptides as well as its own transcript (Bardou et al. 2014). The two short peptides are encoded from the 5' end of *ENOD40* in soybean and are thought to influence the activity of sucrose synthase by interacting with NODULIN100, a subunit of sucrose synthase (Röhrig et al. 2002). The transcript of *ENOD40* can also function as a guide for re-localization of nuclear speckle RNA-binding proteins (NSR) (Fig. 2.2B) (Campalans et al. 2004). For instance, the NSR, *Medicago truncatula* RNA Binding Protein 1 (MtRBP1), was found to be transported by *ENOD40* into the cytoplasmic granules during nodulation (Campalans et al. 2004). This mechanism of RNA-protein interaction and transportation of proteins is required for nodule organogenesis.

It has also been demonstrated in *Arabidopsis* through the transgenic expression of *ENOD40* that this lncRNA is involved in the transport of the NSRs (AtNSRs, closest homolog of the MtRBP1) from nuclear speckles to cytoplasmic bodies (Bardou et al. 2014). Furthermore, AtNSRs have also been reported to be involved in the regulation of alternate splicing events during the initiation of lateral roots and auxin signalling responses (Bardou et al. 2014). Furthermore, AtNSR can interact with *ALTERNATIVE SPLICING COMPETITOR* (*ASCO*) lncRNA (Bardou et al. 2014). Thus, it is apparent that interaction between the *ENOD40* and MtRBP1 is involved in the re-localization of RNA-binding protein, while the interaction between *ASCO* and AtNSR is involved in the splicing activity regulated by AtNSR during auxin

signaling. It would be fascinating to discover additional naturally occurring lncRNAs that interact with other NSR proteins and assist in their re-localization in plants.

In a study by Seo et al. (2017), *ELF18-INDUCED LONG-NONCODING RNA1 (ELENA1),* a 589-nucleotide long intergenic lncRNA (NCBI Id: AT4G16355) was found to be involved in the formation of regulatory complexes through interaction with proteins. *ELENA1* is upregulated in *Arabidopsis* in response to pathogen-associated molecular patterns (PAMPs), such as elongation factor 18 and flagellin (flg22), due to infection by the bacteria *Pseudomonas syringae* (Seo et al. 2017). The knock down of *ELENA1* decreased the expression of *Pathogenesis-Related 1 (PR1)* gene and its overexpression resulted in the upregulation of plant defense genes like *PR2,* genes encoding *B-1,3-glucanase* (callose deposition), and some of the genes induced by salicylic acid (Seo et al. 2017). *ELENA1* interacts with the protein, Mediator subunit 19a (MED19a) and both of these are subsequently recruited to the *PR1* promoter (Seo et al. 2017). This research group proposed two hypotheses for the interaction between the *ELENA1* and *PR1* promoter. For the first hypothesis, a possibility of the formation of a triple helix between *ELENA1* and *PR1* promoter was explored *in silico*; however, with negative results. In another hypothesis, the possibility of interaction of *ELENA1* with other proteins, such as mediator subunits, transcription factors, cofactors or adaptors, to influence the recruitment of MED19a to the transcription machinery has been proposed (Seo et al. 2017); however, this remains to be tested. Based on these studies, *ELENA1* may be an ideal candidate for exploring its application in developing cisgenic crops or for developing markers for marker assisted breeding.

2.4.1.4 Transcriptional regulation mediated by RNA-RNA interaction : lncRNAs as miRNA target mimics

The lncRNAs which have been identified as inhibitors of miRNA expression, are termed as RNA decoys or endogenous target mimics (eTM) or RNA sponges. These lncRNAs contain a short sequence which is homologous to the binding sites of the mRNAs in miRNAs, thus acting as a competitor against mRNAs (for review, (Chekanova 2015) and references therein). The lncRNA *INDUCED BY PHOSPHATE STARVATION1 (IPS1)* of *Arabidopsis* was the first eTM to be studied in detail in plants and was found to regulate the phosphate (Pi) balance by sequestering the miR399 and increased accumulation of its target mRNA (Franco-Zorilla et al. 2007). At*IPS1* (NCBI Id: AT3G09922) is a 542 nt long transcript belonging to the *TPS1/Mt4* gene family, members of which are generally known to be induced by Pi starvation. It contains a

23 nt long motif which has been found to be conserved among *Solanum lycopersicum*, *Medicago truncatula*, *Oryza sativa*, and *Arabidopsis* (for review, (Zhu and Wang 2012)). Under Pi starvation conditions, the miR399 was upregulated and consequently targeted the *PHOSPHATE 2* (*PHO2)* gene, which encodes an E2 ubiquitin conjugase-related enzyme and causes the cleavage of mRNA mediated by miRNA (Fujii et al. 2005, Chiou et al. 2006). A reduced activity of *PHO2* gene resulted in up-regulation of two root-specific phosphate transporter genes *PHT1;8* and *PHT1;9*; these genes are involved in the uptake of Pi from the soil (Bari et al. 2006). The 23 nt motif of the lncRNA *IPS1* is partially complementary to miR399 which forms a 3 nt mismatch bulge at 11-13 nt position of miR399 and thus protects *PHO2* from the cleavage mediated by miRNA399 (Franco-Zorilla et al. 2007) and references therein; Fig. 2.2C). Through this mechanism, *IPS1* attenuates the activity of miR399 by mimicking the structure of the *PHO2* and preventing cleavage of *PHO2* by miR399, which ultimately increases the uptake of Pi by the plants (Franco-Zorilla et al. 2007). Genome-wide analyses in *Arabidopsis* have also identified eTMs which are known to sequester different miRNAs, such as miR156, miR159, miR160, miR166 and miR172 (Wu et al. 2013). Recently, Jiang et al. (2019) demonstrated that the *lncRNA23468* is able to mimic the target genes of miR482b during infection by *Phytophthera infestans* in tomato, which then allows the expression of the nucleotide-binding site leucine-rich repeat (NBS-LRR) protein-coding genes conferring resistance to *P. infestans*. Similarly, Hou et al. (2020) also demonstrated that *lncRNA39026* positively regulates the tomato's defense response against *P. infestans* by functioning as an eTM against miR168a. This attribute of eTMs can be applied to generate transgenic plants for attenuating the function of the corresponding target miRNAs and can be exploited for crop improvement.

More recent studies have indicated that the phloem lncRNAs, *IPS1* and other eTMs, can move long distance from cell-to-cell in response to Pi deficiency and are therefore involved in co-ordinating a systemic response in the early phase of Pi deficiency (Zhang et al. 2019). These findings further raise the possibility of lncRNAs exerting their function and acting as signalling molecules at a systemic level. In addition, plant extracellular vesicles such as exosomes have been shown to assist in cell-to-cell communication by carrying small RNAs and proteins, thus regulating growth and environmental inhabitation, including responses to biotic and abiotic stress stimuli (Rutter and Innes 2017, Hudzik et al. 2020). For instance, in response to *Botrytis cineria* infection in *Arabidopsis*, siRNAs are transported into fungi through exosomes resulting in

improved resistance (Cai et al. 2018). Exosomes can also be used as vehicles for transporting sRNAs that can modulate both host defense and fungal virulence through host-induced gene silencing (Koch et al. 2020). Similarly, lncRNAs are also known to have cell-to-cell movement via exosomes, which is evidenced from multiple studies associated with cancer research (Piao et al. 2020), (Wang et al. 2020). Future work on the cell-to-cell movement or host-to-pathogen movement of plant lncRNAs in response to biotic stress will further pave the path for using lncRNAs for modulating the pathogen RNAs, thus opening the possibility of improving resistance.

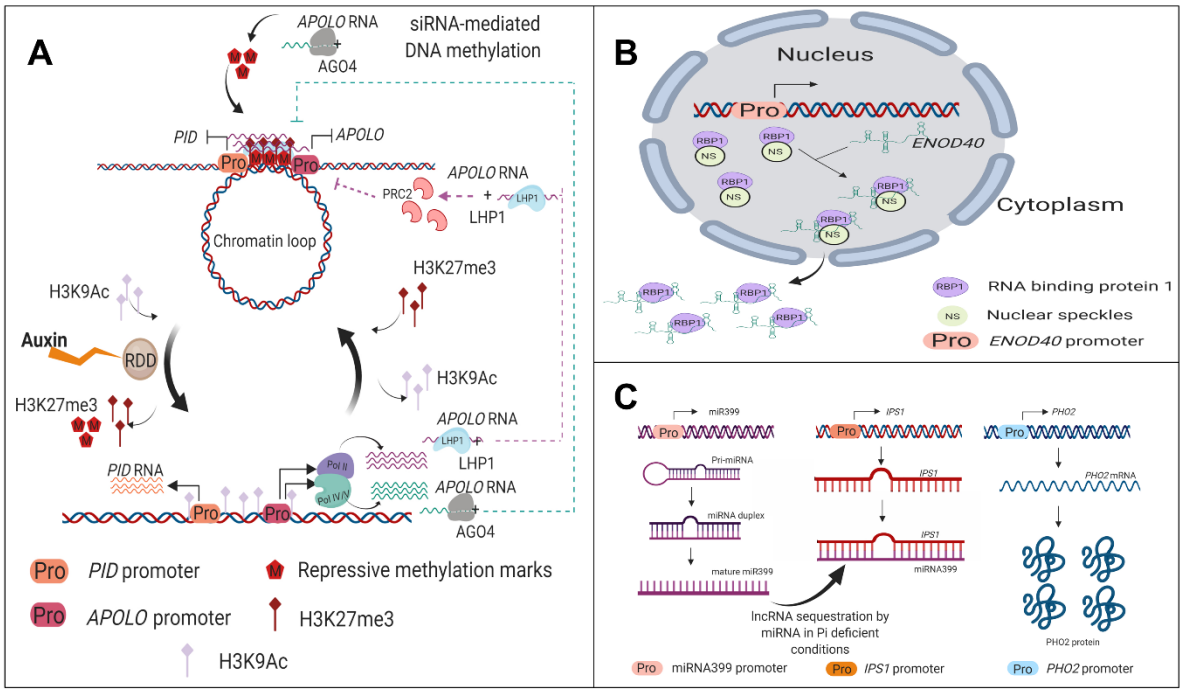


Fig. 2.2 Role of lncRNAs in transcription regulation by different RNA-dependent interaction A) RNA-DNA hybridization: Auxin activates the opening of the chromatin loop via RDD (ROS1, DML2 and DML3) mediated H3K27me3 removal and H3K9Ac deposition. Following this, the transcription of *PID* gene and the dual transcription of *APOLO* by Pol II and Pol IV/V is triggered by an increase in H3K9Ac. Subsequently, the transcripts of *APOLO* generated by Pol IV/V recruits the AGO4, triggering the siRNA-mediated DNA methylation which deposits methylation marks at *APOLO/PID* locus, and the transcripts of *APOLO* generated by Pol II and LHP1 interacts with the genomic region and act as scaffold to recruit PRC2 and deposit the H3K27me3, resulting in the formation of chromatin loop which is then maintained by methylation marks provided by siRNA-mediated DNA methylation. B) RNA-Protein interaction:

The transcript *ENOD40* is responsible for re-localization of the nuclear speckle RPB1 from nucleus to cytoplasm; this transfer is mediated by RNA-protein interaction. C) RNA-RNA interaction: In Pi-sufficient conditions, the miR399 targets the mRNA *PHO2*. In Pi-deficient conditions, the lncRNA acts as the target mimic and sequesters miR399 thus inhibiting the degradation of *PHO2* by miR399. (ROS1: Repressor of Silencing 1; DML2: Demeter-like 2; DML3: Demeter-like 3; *PID: Pinoid; APOLO: AUXIN REGULATED PROMOTER LOOP RNA*; AGO4: Argonaute 4; LHP1: LIKE HETEROCHROMATIN PROTEIN 1; PRC2: Polycomb repressive complex 2; *ENOD40: EARLY NODULIN 40*; RPB1: RNA binding protein1)

2.5 Small RNAs (sRNAs)

Both miRNAs and siRNAs, which are also ncRNAs, are about 20-24 nt long. They differ by the structure of their precursors, biogenesis pathways and mode of action, and are important for regulation of gene expression. The functions of sRNAs have been studied extensively (Borges and Martienssen 2015, Brant and Budak 2018, Chauhan et al. 2017, Hunag et al. 2016, Huang et al. 2019). The miRNAs are transcribed by Pol II and are guided by the transcriptional co-activator called Mediator (for review, (Brant and Budak 2018)); however, production of transcripts by Pol II and Pol IV is required for the biogenesis of siRNAs which are found to be more diverse than the miRNAs (for review, (Kumar et al. 2018) and references therein). The biogenesis and mode of action of the miRNAs and siRNAs are illustrated in Fig. 2.3. Most of the siRNAs are synthesized from the pre-formed dsRNAs transcribed by RNA-dependent RNA polymerases (RDRs) or from the transcription of the inverted repeat region by Pol II (for review, (Kumar et al. 2018)) (Fig. 2.3).

Both type of sRNAs are processed from long RNA precursors by Dicer-like ribonucleases and regulate the post-transcriptional process through the ribonucleoprotein silencing complex (RISC) containing the AGO protein. The miRNAs are cleaved by only RNase III enzyme Dicer-like protein 1 (DCL1) (for review, (Kumar et al. 2018)). Unlike miRNAs, the formation of different classes of siRNAs results from the cleavage of the precursors of siRNA by different types of DCL proteins (DCL1-4). In the epigenetic pathway, dsRNAs are cleaved by DCL-3 into siRNAs which are then incorporated into AGO4 (for review, (Kumar et al. 2018)). The whole machinery for DNA methylation includes the siRNA within the AGO4, the largest subunit of Pol V, and the associated proteins, which are recruited on to the homologous target DNA sequence facilitating the methylation process catalyzed by Domains rearranged methylase 2 (DRM2) (for review, (Sahu et al. 2013)). It is thus clear that siRNAs play an important role in

epigenetic regulation of plant growth and development as well as under environmental stress via transcriptional silencing through the RdDM pathway. We would like to point our readers to recent reviews (Huang et al. 2016, Huang et al. 2019) which have discussed the roles of siRNAs and miRNAs in different plant-pathogen interactions. However, each miRNA and siRNA are pathogen-dependent, and therefore can play distinct functions in plant immune responses. In order to obtain the systemic role of RNA silencing/activation in response to biotic stresses, the function of more sRNAs needs to be explored.

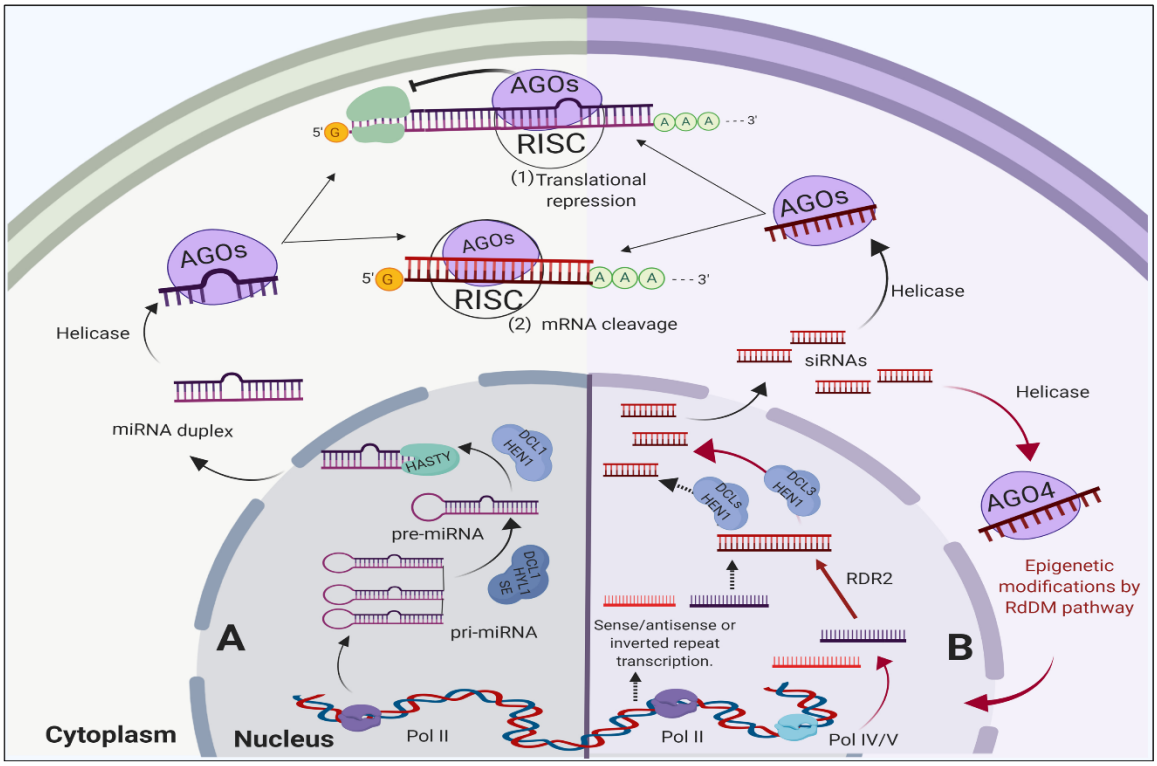


Fig. 2.3 Illustration of the biogenesis of A) miRNA and B) siRNA. A) miRNAs are transcribed by Pol II. Pri-miRNAs and pre-miRNAs are processed by DCL1, HYL1 and zinc finger protein SE to form miRNA duplex. Following methylation, HEN1, the miRNA duplex is exported to the cytoplasm by the protein HASTY. The miRNA duplex is unwound by unknown helicases, making them accessible to AGO, thus forming a RISC. MiRNA bound RISC is then guided to target mRNA, causing translation repression (1) or mRNA cleavage (2). B) siRNA precursors are transcribed by Pol II and Pol IV. Pol II transcribes inverted repeated region in genome as well as generates the sense and antisense transcripts, which results in the generation of dsRNAs. The ssRNAs transcribed by Pol IV become double stranded by RDR2. The dsRNAs are processed into siRNAs by DCLs and HEN1 and then loaded into the AGOs, thus forming RISC which is guided to target mRNA, causing mRNA cleavage or translational repression. The

siRNAs processed by DCL-3 and HEN1, unwound by helicases and then incorporated in AGO4 are important in epigenetic regulation through RdDM pathway. ***(The red arrows show the siRNA processing necessary for the epigenetic modifications by the RdDM pathway.) (DCL1: Dicer-like 1; DCL-3: Dicer-like 3; HYL1: Hyponastic Leaves 1; SE: Serrate; HEN1: HUA enhancer 1; AGO: Argonaute; RISC: RNA-induced silencing complex; AGO4: Argonaute 4; RdDM: RNA-dependent DNA methylation; RDR2: RNA-dependent RNA polymerase 2)

2.6 Circular RNAs (circRNAs)

CircRNAs are endogenous ncRNA; our current knowledge on this is still fragmentary. These have been studied in humans (Memczak et al. 2013), fruit fly (Jakub et al. 2014), various mammals (Liang et al. 2017) and plants (Lu et al. 2015, Ye et al. 2015, Liu et al. 2017, Tang et al. 2018, Wang et al. 2018, Zhu et al. 2019). CircRNAs are non-polyadenylated, single stranded covalently closed loop structures of RNA without any 5' or 3' polarity (Chen and Yang 2015, Chu et al. 2018). CircRNAs are formed during regular splicing of the pre-mRNAs (intronic circRNA) as well as by a non-canonical splicing event which is also referred as back-splicing ((Matsui and Seki 2018, Zhao et al. 2019); (Fig. 2.4A). The most common type of circRNAs is the exonic circRNA (ecRNA), which contains no introns and is formed by exon skipping and direct back-splicing mechanism (Fig. 2.4A) (for review, (Meng et al 2016) and references therein). Another interesting isoform of circRNAs is the exon-intron-circRNA (elciRNA), which forms during circularization of an intronic fragment through joining with the neighbouring exons (for review, (Meng et al 2016) and references therein). These are involved in the cis regulation of transcription of parental genes by associating with Pol II (Li et al. 2015, Zhou et al 2018). The biogenesis and possible mode of action of different types of circRNAs is illustrated in Fig. 2.4A & B. The molecular mechanism of gene expression regulated by circRNAs is still unclear; however, the potential functions of circRNAs in plants through their interaction with endogenous RNA network have been predicted using bioinformatics. These include, circRNA-miRNA interaction network in soybean (Zhao et al. 2017), circRNA-lincRNA-mRNA network in potato (Zhou et al. 2018), and circRNA-miRNA-mRNA networks at different developmental stages in *Arabidopsis* (Liu et al. 2017, Chen et al. 2017). It has also been demonstrated that in *Arabidopsis*, the circRNA originating from exon 6 of the *SEPALLATA3* (*SEP3*) binds to its cognate DNA locus to form an RNA-DNA hybrid which results in transcriptional inhibition of *SEP3*. This further promotes the abundance of the exon-skipped alternate spliced variant of

SEP3, called *SEP3*.3 (Conn et al. 2017). However, additional studies are required to delineate the role of the circRNAs in the regulation of plant gene expression.

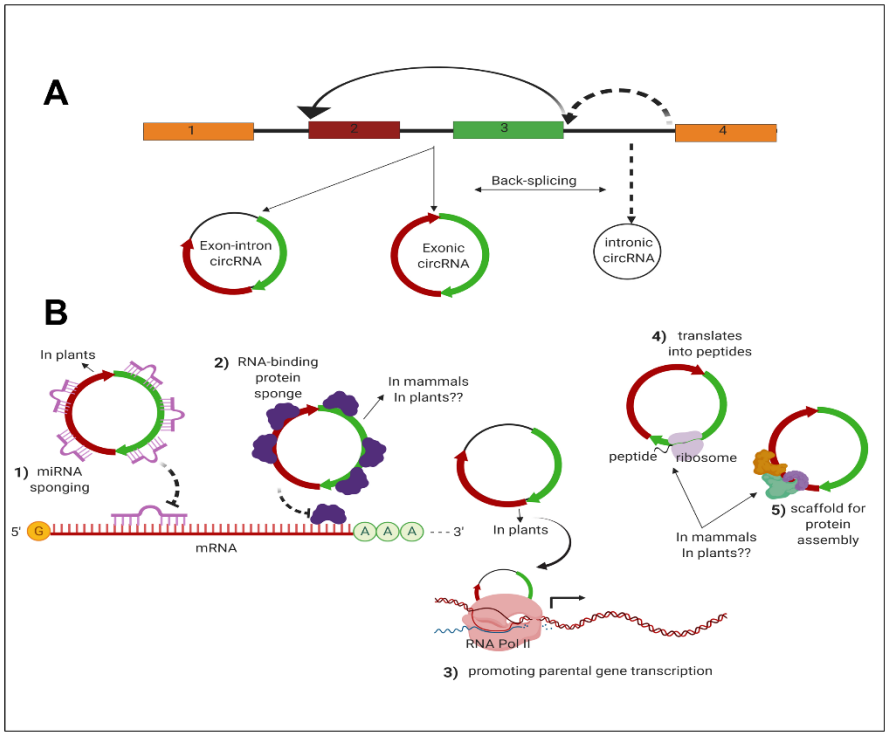


Fig. 2.4 Diagrammatic representation of the biogenesis and mode of action of circRNAs. A) The exons are shown by colored rectangles and introns by black lines. B) Modes of action: 1) The exonic circRNAs are involved in mRNA sponging, and 2) Sponging RNA-binding protein. 3) The intronic circRNAs and exon-intron circRNAs are involved in parental gene expression. 4) Some circRNAs contain ORFs encoding peptides. 5) Some circRNAs act as scaffold for protein assembly regulating the function of the associated proteins in the cell.

2.7 Transcriptomics approaches to study biotic stress response

High throughput methods like RNA-seq for transcriptome analysis and deep sequencing of small RNAs, and microarrays have made it possible to carry out an in-depth analysis of plant ncRNAs. In the following sections, we have attempted to summarize our current understanding and the knowledge gaps of the roles of ncRNA in mediating responses to biotic stresses in important crop plants.

2.7.1 lncRNAs responsive to fungal phytopathogens

Plant diseases caused by diverse fungi, viruses, oomycetes, worms and insects are one of the major problems of crop production. Fungi can cause massive loss of yield and impact quality of agriculturally important crops (for review, (Yang et al. 2017)). Deciphering host-pathogen interactions enables us to understand the infection process by the pathogens and plant responses to pathogen attack, thus, provide valuable information which can be used to develop strategies for the control of plant diseases. Pathogenic fungi can be biotrophic, hemi-biotrophic, saprotrophic and obligate biotrophic or necrotrophic (Spanu and Panstruga 2017). Despite the differences in the life cycles of different fungal pathogens, well-conserved proteins have been found to be involved in their infection processes (for review, (Yang et al. 2017)). These infections are challenged by the plant defence system through many mechanisms including the organization of chromatin structure by the genes involved in defence responses which could be regulated by lncRNAs (for review, (Ramirez-Prado et al. 2018)). The first study to document the involvement of lncRNAs in mediating response to infection by the obligate biotrophic fungus *Blumeria graminis f. sp. tritici (Bgt)* causing powdery mildew disease identified 125 putative lncRNAs in wheat using microarrays (Xin et al. 2011). In addition, this study also reported that the lncRNAs, *TalnRNA5* and *TapmlnRNA19* were the precursors of miRNA2004, *TapmlnRNA8* was the precursor of miRNA2066, and *TapmlnRNA11* was the precursor of siRNA. This indicates that these lncRNAs might be mediating their role by RNA-RNA interaction, a mechanism which has been discussed earlier. The wheat lncRNAs identified by Xin et al. (2011) were not conserved among other plant species and displayed tissue-specific expression in response to infection by *B. graminis*, suggesting their involvement in the regulation of responses to biotic stress only in wheat. This attribute of lncRNAs exhibiting tissue-specific expression was later confirmed by Deng et al (2018) through conservation analysis of the lncRNAs in plants. In a study on the expression of lncRNAs in response to the stripe rust pathogen *Puccinia striiformis* at different growth stages in wheat, Zhang et al. (2013) found that *TalncRNA18, TalncRNA73* and *TalncRNA106* were all upregulated in all stages. In contrast, *TalncRNA108* was downregulated during the early stages and it was upregulated at a later stage in the resistant plants. Similarly, two lncRNAs, *TaS1* and *TaS2*, were identified in a wheat cultivar susceptible to *Blumeria graminis*, and knockdown of these lncRNA resulted in an increased resistance by

15-22% (Li et al. 2014). These observations suggest that the aforementioned lncRNAs may be important regulators of multiple genes involved in defense response pathways.

Vascular wilt, a soil-borne root disease, caused by *Fusarium oxyporum*, affects a wide range of crops including cereals like barley, oats, rye, wheat, rice, maize, as well as other plants like melon, pepper, potato and tomato (Early 2009). Zhu et al. (2014) identified 15 NATs and 20 lincRNAs in *Arabidopsis* which were responsive to *F. oxysporum*. Knockdown of few of these lincRNAs (transcriptionally active regions (TAR)-191,-197, -212 and -224), through RNAi technique resulted in an increased development of disease in response to infection by *F. oxysporum* (Zhu et al. 2014). These results provide clear evidence for the involvement of these lincRNAs in mediating defense against *F. oxysporum*. Furthermore, the strand-specific RNA-seq analysis enabled this research group to uncover the antisense transcriptions which were responsive to infection caused by *F. oxysporum* (Zhu et al. 2014).

Analysis of differentially expressed genes and lncRNAs between tomato plants resistant or susceptible to *Phytophthora infestans* revealed that *lncRNA16397* induces the expression of *S. lycopersicum* glutaredoxin genes (*SlGRX*) involved in scavenging activity of ROS. *LncRNA16397* is positioned antisense to *SlGRX22* and is co-expressed with *SlGRX21* and S*lGRX22* upon infection by *P. infestans*. Interestingly, when *lncRNA16397* is overexpressed in a susceptible tomato plant, it resulted in a significant increase in expression of the *SlGRX21* and *SlGRX22* genes (Cui et al. 2017). This co-expression of *lncRNA* and *SlGRX* suggests reduction in the accumulation of ROS and thus caused less injury to the cell membrane, and thereby resulting an increased resistance to *P. infestans.*

To date, only a few studies on the identification of the lncRNAs in response to biotic stress have been conducted in rice and oilseed crops. However, these studies provide important information on the target genes of these lncRNAs which are involved in disease resistance. For instance, using two near-isogenic lines (NILs) for the QTL conferring resistance to fusarium head blight, Huang et al. (2016) identified 604 lncRNAs in barley responsive to infection by *Fusarium graminearum*. They also suggested that both lncNAT *XLOC_057704* and the gene *trichome birefringence like (TBR-like (TBL)*, which is required for cellulose biosynthesis, are induced in response to infection by *F. graminearum* thus, possibly playing a role in disease resistance, however, this needs to be experimentally established.

In another study, the *in-silico* identification of more than 1900 lncRNAs was performed in two rice lines resistant to *Magnaporthe oryzae* (Jain et al. 2017). Our laboratory was also the first to report the identities of the lncRNAs in oilseed *Brassica napus* involved in mediating responses to stem rot disease caused by *Sclerotinia sclerotiorum* (Joshi et al. 2016). We identified 3,181 lncRNAs including, 2,821 lincRNAs, 111 lncNATs and 41 miRNA precursors (Joshi et al. 2016); however, functional characterization of these lncRNAs and their precise roles in mediating responses to *S. sclerotiorum* has not yet been accomplished. LncRNAs have also been identified in other crops, such as cotton (Zhang et al. 2018) and banana (Muthusamy et al 2019). All of these studies indicate that lncRNAs may regulate their target genes involved in fungal pathogen-host interactions, which needs to be confirmed through functional studies. In a recent study, we have identified 530 differentially expressed lncRNAs in *B. napus* canola lines resistant to clubroot disease (Summanwar et al. 2019). Clubroot is caused by *Plasmodiophora brassicae*, a soil-borne obligate parasite belonging to the eukaryotic supergroup Rhizaria (Burki et al. 2010) and is a major threat to the production of Brassica crops including canola (for review, (Howard et al. 2010)). Most of the target genes regulated by these differentially expressed lncRNAs belong to the pathways involved in plant-pathogen interaction, plant hormone signalling, and primary and secondary metabolic pathways (Summanwar et al. 2019). Further work is needed to establish the precise role of these lncRNAs in mediating clubroot resistance. All these studies have suggested the involvement of lncRNAs and other ncRNAs during host-pathogen/pest interactions. A thorough understanding of their specific biological functions in both disease response pathways as well as in mediating resistance against plant pathogens will ultimately provide novel strategies for the improvement of disease resistance in crops.

2.7.2 lncRNAs responsive to bacterial phytopathogens

Similar to fungal pathogens, bacterial phytopathogens can cause significant crop loss. In the recent years, advances in genomics have facilitated a rapid increase of our understanding on bacterial pathogen-host plant interactions at a molecular level. This has led to the invention of additional strategies for the improvement of resistance to bacterial diseases (for review, (Sundin et al. 2016)). However, very few reports are available on the identification of lncRNAs that are responsive to bacterial pathogens in crops. Recently, Yu et al. (Yu et al. 2020) identified 567 lncRNAs in rice in response to *Xanthomonas oryzae*. They observed that lncRNAs regulating

genes involved in jasmonate pathway (JA) was significantly enriched. They also observed that the overexpression of *ALEX1* resulted in increased resistance against rice bacterial blight via the increase in activation of the JA pathway. Similarly, Li et al. (2018) reported a total of 4,709 lncRNAs in rice in response to infection by *Dickeya zeae* causing foot rot disease. Of these, four lncRNAs, *TCONS_00055598, TCONS_00043825, TCONS_00054645* and *TCONS_00054545,* and their target genes, MYB transcription factor (*OS03G0764600*), peroxidase A2 (*OS03G0235000*), plant peroxidase (*OS02G0318100*) and alternative oxidase (*OS02G0318100*), respectively, are involved in the regulation of ROS during defense responses against *D. zeae* (Li et al. 2018). Possible involvement of lncRNAs in mediating responses to bacterial diseases have also been reported in potato against *Pectobacteriam carotovorum* (Kwenda et al. 2016), kiwifruit plants against *Pseudomonas syringae pv. actinidiae* (Wang et al. 2017) and in princess tree against phytoplasmas (Cao et al. 2018, Fan et al. 2018). Future studies aimed at identifying factors controlling the synthesis and biogenesis of these lncRNAs will offer crucial insights into their function in mediating responses to bacterial pathogens in crop plants.

2.7.3 lncRNAs responsive to viral phytopathogens and pests

lncRNAs have been found to play important roles in mediating plant responses to viral infections (for review, (Nejat and Mantri 2018)). Wang et al. (2015) identified 1,565 lncRNAs in tomato associated with Tomato Yellow Leaf Curl Virus (TYLCV). Virus-Induced Gene Silencing (VIGS) of two upregulated lncRNAs, *slylnc0049* and *slylnc0761,* led to an increase in accumulated virus, suggesting that these lncRNAs may have specific roles in response to TYLCV infection (Wang et al. 2015). These two lncRNAs, therefore, can be viable targets for engineering durable of plant for resistance to TYLCV. This study also demonstrated possible roles for lncRNAs *slylnc1077* and *slylnc0195* which may function as target mimics of miRNAs. In fact, it was observed that when these lncRNAs were overexpressed, it resulted in an enhanced expression of the target genes of sly-miR399 and sly-miR166, respectively . Furthermore, the enhanced expression of these target genes, coincided with the downregulation of sly-miR399 and sly-miR166 (Wang et al. 2015). In another study, lncRNAs have been characterized in tomatoes in response to neutron irradiation followed by TYLCV infection (Zhou et al. 2019). They identified 1,563 lncRNAs and suggested they play a role in mediating plant responses to diverse stimuli, and potentially play a role in resistance to viral phytopathogens. However, no study has so far been conducted on the involvement of lncRNAs in viral diseases like barley yellow dwarf

virus, barley yellow mosaic virus, brome mosaic virus and wheat streak mosaic virus in cereals crops.

Some of the major pests of cereals crops are aphids, slugs, leatherjackets, wireworm, bulb flies and fruit flies (Finch 2002). To date, few studies have been conducted on the role of lncRNAs in response to insect pests. Recently, Zhang et al. (Zhang et al. 2019) identified and characterized the potential role of lncRNAs in cotton leaves under attack by *Aphis gossypii*. In total, 8,414 lncRNAs were identified where 1,331 showed differential expression between the plants under attack and control plants. Interestingly, many lncRNAs were found to regulate the NBS-LRR genes, which recognize the effectors secreted by the pathogen. This suggests that the lncRNAs may improve plant defence response to insects by modulating the NBS-LRR gene family; however, this hypothesis remains to be verified (Zhnag et al. 2019). Recently, Yang et al (2020) used *Pseudomonas putida* as a bio-control agent to improve resistance in tomatoes against root-knot nematodes (*Meliodogyne incognita*).They identified lncRNAs from tomato which were induced by *Pseudomonas putida* isolate Sneb821; some of these lncRNAs were also found to be involved in resistance to root-knot nematodes (*Meliodogyne incognita*). They also characterized lncRNA44664 and lncRNA48734, which function as eTMs for miR396 and miR156, respectively. These results are good examples of bio-control bacteria-induced plant resistance. It has become increasingly evident that lncRNAs play a role during host-pathogen/pest interactions. A thorough understanding of their specific biological roles in both disease responsive pathways as well as in mediating resistance against plant pathogens can ultimately lead to the development of strategies for the improvement of disease resistance in crops.

2.7.4 circRNAs responsive to biotic stress

Recently, several studies have started exploring the abundance of circRNAs in plants in response to biotic stress (for review, (Ye et al. 2015, Zhao et al. 2019)), however, our knowledge of their roles in mediating response to biotic stress still remain largely unclear. Recently, Zhou et al. (2018) identified 2,098 circRNAs in potato in response to *Pectobacterium carotovorum,* among which 44.38% were intergenic circRNAs, and a total of 429 circRNAs were differentially expressed between the resistant and susceptible cultivars. A weighted gene co-expression network analysis revealed that circRNAs are functionally related to mRNAs and lincRNAs, and generally regulate the function of their parental genes (Zhou et al. 2018). Similar to lincRNAs, circRNAs also serve as target mimics of miRNAs for sequestering the miRNAs. This research group

predicted a total of 138 circRNAs to be the target mimics of 120 miRNAs. In another study, knock down of the circRNA and its parental gene *Solyc07g043420.2.1* in tomato resulted in increased TYLCV infection; however, partial resistance was achieved after re-injecting the parental gene by agroinfiltration (Wang et al. 2018) suggesting that circRNAs can up- or down- regulate their parental genes. Similar studies have also been conducted in cotton against *Verticillium dahlia* (Xiang et al. 2018) and in kiwifruit against *P. syringae pv. actinidiae* (Psa) (Wang et al. 2017). All these studies provide evidence that circRNAs play an important role in mediating plant responses to biotic stresses. In mammals, circRNAs have also been known to sequester RNA binding proteins (Holdt et al. 2016) and regulate mRNA at post-transcriptional level, by inhibiting ribosome biogenesis and mRNA translation. CircRNAs have also been reported to function as scaffold for assembling protein complexes (Du et al. 2017) and are also known to encode functional peptides in mammals (Zhnag et al. 2018). However, it still remains to be verified if the mechanism of action of the circRNAs is similar across the plant kingdom in response to biotic stresses.

2.7.5 miRNAs derived from lncRNAs responsive to pathogens

miRNAs are involved in mediating biotic stress responses in many agronomically important crops (Chauhan et al. 2017, Khraiwesh et al. 2012). To be consistent with the main focus of this review on lncRNAs, we have highlighted only the studies those identified pathogen-responsive miRNAs which are presumably derived from a lncRNA. To accomplish this, we obtained the miRNA stem-loop sequences from www.mirbase.org and identified their precursor lncRNAs by BLAST analysis in the lncRNA plant database CANTATAdb v.2.0 (http://cantata.amu.edu.pl/) using all default parameters. Our results are presented in Table 2.3 together with some of the predicted targets of the miRNAs reported previously. For instance, a differentially expressed miRNA, the bna-miR156b, was identified in *B. napus* in response to *P. brassicae* (Verma et al. 2014). Interestingly, the above mentioned *in silico* analysis revealed that CNT20112547 to be the precursor of bna-miR156b, which was induced in *B. napus* in response to infection by *P. brassicae*. Similar results have been observed in rice (Campo et al. 2013), wheat (Kumar et al. 2014, Feng et al. 2015) and canola (Verma et al. 2014, Shen et al. 2014). The fact that the miRNAs involved in regulating biotic stress are harboured within lncRNAs broadens the current perspective of lncRNAs as ubiquitous regulators under biotic stress.

2.7.6 Potential genome editing approaches for manipulating lncRNAs for crop improvement

Aforementioned studies have elaborated on the roles of lncRNAs in response to various biotic stresses and this knowledge can be utilized for improving disease resistance in crop plants. Various versions of CRISPR/Cas9 technology have been reported for manipulating lncRNAs in mammalian systems or cell lines (Zare et al. 2018) which can be adapted to investigate the functional roles of lncRNA in plants. CRISPR-Cas9 technology is faster, cheaper, and efficient, and enables the editing of genomic elements, including promoters, enhancers as well as intergenic regions and introns. CRISPR when used with the Cas9 enzyme produces double stranded breaks (DSBs) in the target genomic region which can be repaired by non-homologous end joining (NHEJ) mechanism; thus, inducing a small-frameshift mutation and thereby producing an efficient knockout (Lei et al. 2013). This approach is known as the CRISPRn (base-editing) mutagenesis (Lei et al. 2013) has been used in some studies (Liu et al. 2018, Li et al. 2018, Chao et al. 2019), however, its applicability may be an issue for lncRNAs exerting their effects through transcription.

In another study, CRISPR excision has been used for generating lncRNA knockouts in embryonic stem cells by either excising the whole lncRNA or the promotor of the lncRNA (Yin et al. 2015). However, this method would be inappropriate if the lncRNA loci intersect or overlap with other coding/non-coding loci. Similarly, it would be inapplicable to delete the promotor of lncRNA if they are internal or bidirectional in nature. In such cases, it would be more appropriate to obtain a partial deletion of the exons that do not overlap with other functionally relevant genomic regions (Zare et al. 2018).

Another approach, CRISPRn HR which depends on homology directed repair (HDR) when repairing DSBs, can be used for knocking out of lncRNA by homology-directed insertion of transcription termination signal or premature inhibitory signal like polyadenylation-A (poly-A) or other destabilizing sequences after the transcription start site (Nie et al. 2019) or at key sites along the lncRNA loci (Zare et al. 2018). This is an efficient and less invasive strategy for silencing the expression of lncRNA; however, this strategy is not useful for studying the lncRNAs whose promoters are located within their sequences or are proximal to another neighbouring gene whose expression and function could potentially be disrupted. Therefore, before implementing CRISPR methods to investigate the lncRNAs, it is essential to have a clear understanding of their genomic loci, their impact on other genes and whether they regulate the genes in cis or in trans manner.

Another less invasive strategy for epigenetic manipulation includes CRISPR with dCas9 fused with epigenetic modifiers, such as DNA methyltransferase or histone modifiers, for targeting the desired genomic loci for modulating the histone marks and affecting lncRNA expression. For instance, Janga et al. (2018) demonstrated through silencing of the *MALAT1* lncRNA and miR-146a and miR-155 in human monocytes that the epigenetic codes could be altered for achieving desired expression of the lncRNAs; they also demonstrated that mir-155 is involved in the control of cell-cycle. Despite some of the disadvantages, CRISPR could be the preferred option for interrogating the lncRNAs as it has fewer off-target effects as compared to RNAi. Additionally, the lncRNA databases and the repositories of the plant are an asset for those interested to study their mechanism of function or manipulating them to improve disease resistance (Table 2.4).

2.8 Conclusion and future perspectives

It is clear from the published literature that lncRNAs are important players in the regulation of genetic pathways involved in mediating responses to stress in plants. LncRNAs perform these roles by regulating gene transcription through chromatin remodelling, epigenetic regulation, miRNA sponging, altering mRNA stability and localization. Although our knowledge on the function of lncRNAs in biotic stress is still in its infancy, yet it is evident that lncRNAs can be promising molecular targets in the manipulation of gene expression in disease responsive pathways. It would be advantageous to include information on the functionally validated lncRNAs in plant RNA databases. Such efforts to enhance disease resistance in crop plants are crucial as the pathogen races causing disease are continuously evolving and breaking down the available resistances in plants. This is particularly relevant when climate change is expected to increase the diversity of plant pathogens thereby posing additional threats to global food security. Further understanding of the molecular functions of the lncRNAs and their roles in mediating plant responses to pathogens will make it possible to use this knowledge in making precise change in the genome as well as to design robust biomarkers for use by plant breeders to develop disease resistant crop plants. Future research could also focus on the structural characteristics of the lncRNAs and their features critical for their biogenesis, spatial and temporal expression and localization and functional interaction and with the genes and proteins and their regulation. Such studies may advance our knowledge about lncRNAs and may lead to additional strategies to improve global food security.

Table 2.1 Examples of crops improved for disease resistance through plant breeding

PLANT	PATHOGEN / INSECT	GENE / QTL	METHOD	REFERENCE
CLASSICAL BREEDING				
Triticum aestivum	*Erysiphe graminis*	*Pm12*		Miller et al. (1988)
Triticum aestivum	*Erysiphe graminis*	*Pm16*		Reader and Miller (1991)
Oryza sativa	*Pyricularia grisea*; *Xanthor nonas oryzae*	Bacterial blight and blast resistance		Amante-Bordeos et al. (1992)
MOLECULAR BREEDING				
Arachis hypogaea	*Cercosporidium personatum*	*QTL*	MAS	Clevenger et al. (2018)
Brassica napus	*Plasmodiophora brassicae*	*RPb1*	MAS	Chu et al. (2013)
Brassica oleracea	*Plasmodiophora brassicae*	*QTL, PbBo(Anju)1*	MAGP	Tomita et al. (2013)
Cicer arietinum	*Fusarium oxyporum*	*QTL*	MAB	Mannur et al. (2018)
Hordeum vulgare	*Puccinia striiformis*	*QTL*	MAB/ MAS	Toojinda et al. (1998)
Hordeum vulgare	Barley yellow dwarf virus	*Yd2*	MAB/ MAS	Jefferies et al. (2003)
Oryza sativa	*Xanthomonas oryzae*	*xa21, xa13* and *Xa5*	MAB/ MAS	Sanchez et al. (2000)
Oryza sativa	*Xanthomonas oryzae*	*xa21, xa13* and *Xa5*	MAGP	Singh et al. (2001)
Oryza sativa	*Magnaporthe oryzae; Xanthomonas oryzae*	*Pi2* and *Pi54*; *xa13* and *xa21*	MAB	Ellur et al. (2016)
Oryza sativa	*Xanthomonas oryzae*	*xa21, xa13* and *xa5*	MAB	Baliyan et al. (2018)
Triticum aestivum	*Blumeria graminis*	*Pm43*	MAS	He et al. (2009)
Triticum aestivum	*Blumeria graminis*	*Pm55*	MAS	Zhang et al. (2016)

Triticum aestivum	*Bipolaris sorokiniana*	*Lr34*	MAB/ MAS	Vasistha et al. (2017)
Zea mays	*Sphacelotheca reiliana*	*qHSR1*	MAB/ MAS	Zhao et al. (2012)
Solanum lycopersicum	*Phytophthera infestans; Ralstonia solanacearum; Stemphyllium spp.;* Tomato yellow leaf curl virus; *Fusarium oxysporum;* Tob acco mosaic virus	*Ph-2, Ph-3, Bwr-12*; Sm; *Ty-2, Ty-3; I2;Tm2*^{2}	MAGP	Hanson et al. (2016)
Solanum lycopersicum	*Xanthomonas euvesicatoria*		GS	Liabeuf et al. (2017)
Zea mays	*Stenocarpella maydis*		GS	dos Santos et al. (2016)
Pisum sativum	*Ascochyta rabiei*		GS	Carpenter et al. (2018)

[a] MAS: Marker Assisted Selection; MAGP: Marker Assisted Gene Pyramiding; MAB: Marker Assisted Backcross

Table 2.2 Examples of crops improved for disease resistance using various genome modification techniques.

PLANT	PATHOGEN	GENE(S)	REFERENCE
CISGENESIS (Overexpression)			
Malus pumila	*Erwinia amylovora*	*FB_MR5*	Kost et al. (2015)
Solanum tuberosum	*Phytophthora infestans*	Stacking of multiple R genes	Haverkort et al. (2016)
Oryza sativa	*Magnaporthe grisea*	*Pi9*	Tamang et al. (2018)
Solanum lycopersicum	*Ralstonia solanacearum*	*SlP14a-PPC20*	Morais et al. (2019)
TRANSGENESIS (Overexpression)			
Triticum aestivum	*Puccinia recondite, Blumeria graminis, Puccinia graminis, Puccinia striiformis*	*chi26* from barley	Eissa et al. (2017)
Glycine max	*Phytophthora sojae, Fusarium virguliforme*	*PSS1* from *Arabidopsis*	Wang et al. **(2018)**
Citrus sp. 'W. Murcott' mandarin	*Xanthomonas axonopodis*	*Xa21* from rice	Omar et al. (2018)
Triticium aestivum	*Puccinia triticina,*	*MtDEF4.2 from Medicago truncatula*	Kaur et al. (2017)
TALEN (Loss of function)			
Oryza sativa	*Xanthomonas oryzae*	*Os11N3*	Li et al. (2012)
Triticum aestivum	*Blumeria graminis*	*TaMLO* ho mologs	Wang et al. (2014)
CRISPR/Cas9 (Loss of function)			

Triticum aestivum	*Blumeria graminis*	*TaMLO-A1* allele	Wang et al. (2014)
Oryza sativa	*Xanthomonas oryzae*	*OsSWEET13*	Zhou et al. (2015)
Oryza sativa	*Magnaporthe oryzae*	*OsERF922*	Wang et al. (2016)
Citrus sinesis	*Xanthomonas citri*	*CsLOB1* allele	Peng et al. 2017
Solanum lycopersicum	*Oidium neolycopersici*	*Mlo1*	Nekrasov et al. (2017)

Table 2.3 *In silico* prediction of the lncRNAs which are precursors of the previously reported miRNAs.

PLANT	PATHOGEN	MiRNA ID	PREDICTED LNCRNA PRECURSOR	PREDICTED TARGET GENES	REFERENCE
Oryza sativa	*Magnaporthe oryzae*	osa-miR7695	CNT2078696	Natural resistance-associated macrophage protein 6	Campo et al. (2013)
Oryza sativa	*Magnaporthe oryzae*	osa-miR11341	CNT20228539	------	Baldrich et al. (2015)
Oryza sativa	*Magnaporthe oryzae*	osa-miR11339	CNT2082687	Terpene synthase-Lipid metabolism	
Oryza sativa	*Magnaporthe oryzae*	osa-miR11342	CNT20228539	Transcription factor BIM2	
Oryza sativa	*Magnaporthe oryzae*	osa-miR11336	CNT2082687	Terpene synthase-Lipid metabolism	
Triticum aestivum	*Puccinia triticina*	tae-miR1122	CNT20162460	------	Kumar et al. (2014)
Triticum aestivum	*Puccinia striiformis*	Novel miRNA- PN-tae-miR5368_L-2R+3 and PN-tae-miR5368-p5	CNT2019178	------	Feng et al. (2015)
Brassica napus	*Plasmodiophora brassicae*	bna-miR156b-3p	CNT20112547	Squamosa promoter binding protein like 15	Verma et al. (2014), Wei et al. (2012)
Brassica napus	*Verticillium longisporum*	bna-miR6030	CNT20143475	Disease resistance protein	Shen et al. (2014)

[a]Retrieved on 4 December, 2019.

Table 2.4 Plant lncRNA databases

DATABASE NAME	LINKS TO DATABASE	REFERENCES
GreeNC (v1.12)	http://greenc.sciencedesigners.com/wiki/Main_Page	Gallart et al. (2016)
Cantatadb 2.0	http://cantata.amu.edu.pl.login.ezproxy.library.ualberta.ca/	Szczesniak et al. (2016)
EVLncRNAs	http://biophy.dzu.edu.cn/EVLncRNAs/	Zhou et al. (2017)
PeTMbase	http://tools.ibg.deu.edu.tr.login.ezproxy.library.ualberta.ca/petmbase/	Karakulah et al. (2016)
PLNlncRbase	http://bioinformatics.ahau.edu.cn/PLNlncRbase	Xuan et al. (2015)
lncRNAdb v2.0	http://lncrnadb.org	Quek et al. (2015)
PNRD	http://structuralbiology.cau.edu.cn/PNRD/index.php	Yi et al. (2014)
NONCODE	http://www.noncode.org/browse.php	Zhao et al. (2016); Fang et al. (2018)
PlantNATsDB	http://bis.zju.edu.cn/pnatdb/	Chen et al. (2012)
TAIR	https://www-*Arabidopsis*-org.login.ezproxy.library.ualberta.ca/	Swarbreck et al. (2007)
RNAcentral	https://rnacentral.org/	The RNAcentral Consortium (2018)

Chapter 3: Identification of lncRNAs Responsive to Infection by *Plasmodiophora brassicae* in Clubroot-Susceptible and -Resistant *Brassica napus* Lines Carrying Resistance Introgressed from Rutabaga

3.1 Introduction

Intergenic regions in the genome have been shown to be involved in transcriptional and posttranscription regulation of gene expression (Kapranov et al. 2007). These regions are ubiquitously transcribed into RNAs and are classified based on their different sizes and their regulatory functions. These types of functional RNAs are classified as noncoding RNAs (ncRNAs) (Ariel et al. 2015; Kapranov et al. 2007; Wilusz et al. 2009). Long ncRNA (lncRNAs) are defined to be greater than 200 bp (Chen 2009; Kapranov et al. 2007) and have been implicated in various gene regulatory processes, including the reprogramming of gene expression in plants in response to a variety of biotic and abiotic stresses (Nejat and Mantri 2018).

lncRNAs have been classified into different categories based on their genomic location and orientation. These include promoter upstream transcripts, enhancer RNAs, intergenic ncRNAs (lincRNAs), intronic ncRNAs, bidirectional lncRNAs, overlapping lncRNAs, and natural antisense transcripts (NATs) (Mattick and Rinn 2015; Morris and Mattick 2014; Wu et al. 2017). Similar to the mRNA transcription process, lncRNAs are also transcribed by RNA polymerase II in plants and are modified with a 5′-cap and 3′-end poly-adenylation (Guttman et al. 2009).

Advances in the next-generation sequencing have enabled genome-wide identification of lncRNAs in many eukaryotes, including several plant species. The first systematic investigation on plant lncRNAs began with the identification of 6,480 lincRNAs in *Arabidopsis thaliana* through bioinformatics analysis; however, expression profiling using microarray and RNA-seq confirmed 2,708 transcripts to be lincRNAs (Liu et al. 2012). Similarly, exploitation of expressed sequence tag databases and 30 different RNA-seq datasets resulted in the identification of 20,163 putative lncRNAs (Li et al. 2014). Among these, 1,704 were considered to be high-confidence lncRNAs in maize, which did not exhibit sequence similarity with known ncRNAs. Most (>90%) of the lncRNAs were found to be in the intergenic regions, with more than 50% of the lncRNAs observed to be expressed in a tissue-specific manner and largely affected by transgenetic factors (Li et al. 2014).

Regulatory roles for the lncRNAs have also been reported in many plant species; for example, in cellulose, lignin, and gibberellin biosynthesis as well as in wood formation in *Populus tomentosa* (Chen et al. 2015). Such regulatory roles have also been observed in flower and fruit development in diploid strawberry *Fragaria vesca* (C. Y. Kang and Liu 2015) and trifoliate orange *Poncirus trifoliata* L. Raf. (Wang et al. 2017). In *Brassica campestris*, Song et al. (2013) identified an lncRNA, *BcM11*, involved in pollen development. Additional studies have supported a crucial role of lncRNAs in plant growth and development, disease response, and genetic or epigenetic regulation of traits (Huang and Zhang 2014; Huang et al. 2018; Nejat and Mantri 2018).

lncRNAs are also known as the class of riboregulators that are expressed in response to plant biotic stress. Zhu et al. (2014) identified 35 lncRNAs, including 15 NATs and 20 lincRNAs, involved in resistance to *Fusarium oxysporum* in *A. thaliana*. Of these, antisense transcription in approximately 20% of the annotated *A. thaliana* genes was observed. The role of lncRNAs in response to diseases has also been documented in many other crop species; for example in tomato against Tomato yellow leaf curl virus (Wang et al. 2015); in *B. napus* in response to *Sclerotinia sclerotiorum*, which causes stem rot disease (Joshi et al. 2016); as well as in the forest tree genus *Populus* in response to *Melampsora larici-populina* (N. Wang et al. 2017). It has also been reported that overexpression of the lncRNA16397 in tomato induces the expression of glutaredoxin, reduces the accumulation of reactive oxygen species, and, thus, confers resistance to *Phytophthora infestans* (Cui et al. 2017). Recently, L. Zhang et al. (2018) reported that silencing of lncRNAs *GhlncNAT-ANX2* and *GhlncNAT-RLP7*, which are involved in the regulation of lipoxygenase activity, results in enhanced resistance toward *Verticillium dahliae* and *Botrytis cinerea*. ELF18-induced long noncoding RNA 1 is another functionally characterized lncRNA, which enhances resistance against *Pseudomonas syringae* in *A. thaliana* by positively regulating pathogenesis-related 1 (*PR1*) gene expression (Seo et al. 2017). Taken together, all of these reports suggest a significant role for lncRNAs in response and resistance to diseases in plants and warrant continued investigation. To our knowledge, no study has been conducted thus far to understand the role of lncRNAs in resistance to other diseases such as the clubroot (CR) disease of *Brassica* crops.

CR is a soilborne disease affecting cruciferous, crops including canola, cabbage, turnip, Brussels sprout, and many other *Brassica* crops (Howard et al. 2010). It is caused by a soilborne,

obligate parasite, *Plasmodiophora brassicae* Woronin, a member of the eukaryotic supergroup Rhizaria (Burki et al. 2010). *P. brassicae* is a fungus-like protist which is known to possess a complex lifecycle and is unable to grow and multiply without a living host (Hwang et al. 2012). This pathogen can stay in soil for over 15 years (Wallenhammar 1996) in the form of resting spores and germinates only in the presence of secretions from the plant roots (Macfarlane 1970). Disease development leads to the formation of galls on the roots, which negatively affect water and nutrient uptake by roots, leading to stunted growth (Dixon 2009). In the case of the most severely affected fields, CR disease can cause yield loss of approximately 30-100% in canola (Tewari et al. 2005). Due to the nature of the occurrence of this disease and its ability to produce long-living resting spores, it is difficult to control this disease with just implementation of improved agricultural practices and chemical treatments (Dobson et al. 1983; Howard et al. 2010; Kowata-Dresch and May-De Mio 2012).

Despite the existence of commercial, CR-resistant cultivars and integrated management approaches, it is still crucial to achieve a better understanding of plant responses to *P. brassicae* infection and to develop cultivars with durable resistance, in order to combat genetic shifts in this pathogen population and subsequent loss of effective resistance (Peng et al. 2014). Genome and proteome analyses of various plant species of the family Brassicaceae challenged with *P. brassicae* have provided some insights about molecular mechanisms triggering resistance in the host plants (Cao et al. 2008; Chen et al. 2016; Devos et al. 2006; Zhang et al. 2016). The availability of the *Brassica napus* genome sequence (Chalhoub et al. 2014) has now paved the way for investigation into the ncRNA and their regulatory roles in *Brassica* spp. The objective of this research was to identify the lncRNAs involved in the regulation of resistance to *P. brassicae* by using *B. napus* canola lines carrying resistance introgressed from rutabaga (Hasan and Rahman 2016).

3.2 Materials and Methods

3.2.1 Plant material

A doubled-haploid (DH) population derived from the F_1 generation of Rutabaga-BF × UA AlfaGold was used in this study. The parent Rutabaga-BF is a CR-R inbred line derived from the rutabaga (*B. napus* var. *napobrassica*) cultivar Brookfield through self pollination of single plants (Hasan and Rahman 2016), while UA AlfaGold is an open-pollinated CR-S spring-type *B. napus* canola cultivar (Rahman 2016). The details of the development of the DH lines

and phenotyping for resistance to CR disease have been described previously (Hasan and Rahman 2016).

3.2.2 Microscopy analysis for primary infection by *P. brassiceae*

To determine the time of primary infection by *P. brassicae* and the development of disease symptoms, 12 CR-R and 12 CR-S DH lines were grown in a greenhouse at 22 and 15°C (day and night, respectively) with a 16-h photoperiod. In total, eight plants of these lines were grown in 72-cell trays (52 by 26 by 5 cm, length by width by depth, respectively) filled with Sunshine Professional Growing Mix (Sun Gro Horticulture Canada Ltd.). Ten-day-old seedlings were inoculated with *P. brassicae* single-spore isolate SACAN-ss1 of pathotype 3 (≈ pathotype 3H) (Williams 1966) (obtained from Dr. Stephen Strelkov, University of Alberta) following a direct inoculation method. The details of inoculum preparation and inoculation procedure have been described in a previous study (Hasan and Rahman 2016). Briefly, the inoculum was prepared by crushing the galls using sterile water and the homogenized spore suspension was filtered through multilayered cheesecloth. The spore concentration was adjusted to 1×10^7 spores/ml and inoculation was done by injecting 1 ml of the spore suspension into the soil near the root of the seedlings (Voorrips and Visser 1993). Root samples of the inoculated seedlings were collected at 0 hpi and 6, 10, 14, 18, and 22 dpi. Uninoculated controls were also included in the study.

The roots were gently washed with water, cut into 10- to 15-mm segments, and fixed in a solution of formalin, acetic acid, and alcohol comprising 5 ml of commercial formalin (40% formaldehyde solution), 5 ml of glacial acetic acid, and 90 ml of 50% ethyl alcohol (Yeung and Saxena 2005) at room temperature for 24 h (Verma et al. 2014). Following this, samples were dehydrated in a series of ethanol dilutions, gradually followed by the treatment with the toluene (clearing agent) using Leica TP1020 tissue processor (Leica Biosystems). The tissues were then infiltrated with ParaplastH (Leica Biosystems) and cast into rectangular blocks. Longitudinal sections of roots, 10 mm thick, were prepared using A0 Rotary microtome (Leica Biosystems) and affixed to glass slides. The samples on the glass slides were deparaffinated with toluene, rehydrated using a graded series of ethanol, stained with Harris Hematoxylin, and counterstained with Eosin Y (Leica Biosystems). After dehydrating again with 100% ethanol and toluene, the slides were mounted with DPX mounting media (Sigma-Aldrich). The sections were viewed

with a Zeiss Axioscope, photographed with Zeiss Axiocam, and analyzed using AxioVision software (Verma et al. 2014).

3.2.3 Plant materials for lncRNA-seq analysis

In all, 12 R and 12 S DH lines were used for this study. In total, 192 seedlings of the 24 DH (12 each of R and S) lines (24 × 8 = 192) were grown in the abovementioned 72-cell trays filled with Sunshine Professional Growing Mix. Seedlings were inoculated with *P. brassicae* single-spore isolate SACAN-ss1 pathotype 3, as described above. The inoculated plants were maintained in a greenhouse at 22 and 15°C (day and night, respectively) with a 16-h photoperiod, and the soil was kept saturated for successful infection and disease development. The experiment was repeated three times as biological replicates.

3.2.4 RNA isolation, library preparation, and sequencing

For RNA isolation, roots of three seedlings from each of the 12 R and 12 S DH lines were harvested at 10 dpi as bulk of R and S samples, respectively; frozen in liquid nitrogen; and stored at −80°C until further use. The remaining plants of the R and S DH lines were grown to maturity and phenotyped for resistance in order to confirm the resistance or susceptibility of the DH lines.

Total RNA from the R and S bulks was extracted using TRIzol reagent (Invitrogen) and the Qiagen RNeasy Plant Mini Kit (Qiagen) as per the manufacturer's instructions. Root samples were homogenized into a fine powder and 100 mg of this tissue powder was treated with 1 ml of TRIzol in an Eppendorf tube. The mixture was centrifuged at 9,425 relative centrifugal force (RCF) for 15 min at 4°C and the supernatant was collected in a new tube. To achieve phase separation, 200 µl of chloroform was added to the supernatant, mixed vigorously, incubated at room temperature for 5 min, and centrifuged at 9,425 RCF. The aqueous layer was collected in a new tube without disturbing the interface layer containing DNA and 0.5 ml of chilled 95% ethanol was added and mixed. The sample was then transferred to a RNeasy minispin column (pink) of the Qiagen RNeasy Plant Mini Kit and the next steps were carried out as per the manufacturer's instructions. The quality and quantity of total RNA was determined using a NanoDrop ND-2000 spectrophotometer (Thermo Fisher Scientific). The integrity of the total RNA was determined on an Agilent Tapestation using RNA Screen Tape assay following the

manufacturer's instructions (Agilent Technologies). The RNA samples with RNA integrity number greater than nine were used for further analysis.

In total, six libraries—three R (each sample with bulks of 12 plants) and three S (each sample with bulks of 12 plants)—were prepared and 4 μg of RNA of each sample was used for library preparation. Ribosomal RNA (rRNA) was removed from the samples using an Epicentre RIbo-zero rRNA Removal Kit (Epicentre) and strand-specific libraries were prepared using a NEB Next Ultra Directional RNA Library Prep Kit (NEB) following the manufacturer's instructions. RNA library sequencing was carried out using the Illumina Hiseq platform and 2 × 150-bp paired-end reads were generated. The library preparation and sequencing were conducted by Novogene (USA).

3.2.5 Mapping to the reference genome and transcriptome assembly

The lncRNAseq raw data in FASTQ format was processed using Novogene's in-house Perl scripts and the reads containing adaptors, unrecognized bases, and low-quality reads were removed. Based on the clean reads, percentage of the total number of bases where the Phred score is greater than 20 (Q20) and 30 (Q30) was calculated. This indicated the base call accuracy and percentage of GC content in clean data. To align the reads, reference genome annotation files were downloaded from *B. napus* reference genome v5.0 (Chalhoub et al. 2014) and indexed using Bowtie v2.0.6 (Langmead and Salzberg 2012). The paired-end clean reads were aligned to this reference genome using TopHat v2.0.9 (Trapnell et al. 2012) and the mapped reads from each sample were assembled using the statistical segmentation model Scripture β2 (Guttman et al. 2010) and Cufflinks v2.2.1 (Trapnell et al. 2010). A combination of these enabled us to distinguish the expressed loci from experimental noise. The spliced reads were then used to assemble the expressed segments and a probabilistic model to simultaneously assemble and provide a maximum-likelihood explanation of the expression at a specific locus. The program Scripture was run with the default parameters, while 'min-frags-per-transfrag = 0' and '–library-type' were set for Cufflinks, with other parameters set to default. Sorting and removal of the duplicated reads and merging of the bam alignment file for each sample was accomplished by Picard- tools v1.41 and Samtools v.0.1.18.

3.2.6 lncRNA detection and target prediction

The Cufflink assemblies generated from the samples were merged using the software Cuffmerge (Trapnell et al. 2010) to remove the replicates and artifacts. The remaining transcripts were subjected to the exon number filter (exon ≥ 2) and transcript length filter (>200 nt) to remove the other known noncoding RNA (rRNA, tRNA, snRNA, snoRNA, premiRNA, and pseudogenes) using Cuffcompare (Trapnell et al. 2010). Subsequently, the transcript sequences were classified and class-coded into candidate lincRNAs, intronic lncRNAs, and antisense lncRNAs. Four different software tools, including Coding Noncoding Index v2 (Sun et al. 2013); CPC 0.9-r2 (Kong et al. 2007), Pfam scan v1.3 (Punta et al. 2012), and phylogenic codon substitution frequency (Lin et al. 2011), were used to filter out the transcripts with coding potential. Only the transcripts without coding potential were considered as the candidate set of lncRNAs. The software program Cuffdiff v2.1.1 (Trapnell et al. 2010) was used to calculate the FPKMs in each sample. Fold changes for the DE lncRNAs were calculated as log2(fold change) = log2(FPKM susceptible/FPKM resistant). A P value was assigned to each lncRNA and adjusted by the Benjamin and Hochberg (1995) approach for controlling the false-discovery rate. DE lncRNAs were considered significant when the absolute value of log2(fold change) > 1, with a q value < 0.05. The targets, where the lncRNA plays a cis role, were predicted by searching for genes at 100,000 bp upstream and downstream of the lncRNAs. To predict the trans role, correlation between expression of lncRNAs and coding genes was calculated using the Pearson correlation coefficient, and genes having absolute correlation coefficient > 0.95 were considered for further functional enrichment analysis.

3.2.7 Classification and characterization of lncRNA functions

The lncRNAs acting as precursors of the known miRNAs were identified by BLASTn analysis (e-value = 1e-5) (J. Zhang et al. 2018). For this, all of the stem loop sequences of known *B. napus* miRNA was downloaded from the miRbase database and aligned with the lncRNA sequences showing significant differential expression. The hairpin loop formation in the lncRNAs was first checked using the miRNAFold server (Tav et al. 2016), and the secondary structure was plotted using the Vienna RNAfold web server (Gruber et al. 2008) and visualized using Forna (Kerpedjiev et al. 2015)

lncRNAs mimicking the endogenous miRNA targets were predicted by psRNA-Target V2.2017 (Dai et al. 2018) by subjecting them to stringent parameters of maximum expectation

set to 3.5 and target accessibility set to 25.0. The interaction between the lncRNAs, miRNAs, and miRNA gene targets were visualized using Cytoscape v3.6.1 (Shannon et al. 2003).

NATs were subjected to RNAup (ViennaRNA Package v2.4.9) (Hofacker 2009; Lorenz et al. 2011) to calculate the thermodynamics of their interaction with their corresponding sense mRNAs and, therefore, determine the potential candidates of lncRNA-mRNA interaction. The potential dimerized secondary structures were calculated using RNAcofold (ViennaRNA Package v2.4.9) (Lorenz et al. 2011).

3.2.8 qRT-PCR analysis

The lncRNAs responsive to the *P. brassicae* infection and their respective targets from *B. napus* were validated using qRT-PCR. This included the lncRNAs from chromosome BnaA08, lncRNAs which are found to be regulating the targets on chromosome BnaA08 and elsewhere in the genome, as well as some random lncRNAs which regulate target genes predicted to play a role in plant hormone signaling. Sequences of the primers used in qRT-PCR are listed in Supplementary Table 3.1. All qRT-PCR analyses were performed on QuantStudio 6 Flex Real-Time PCR System (Thermo Fisher Scientific) using PowerUP SYBR Green Master Mix from Applied Biosystems (Thermo Fisher Scientific). RNA isolated from roots of R and S DH lines inoculated with *P. brassicae* were treated with RNase-free DNase I. The total RNA was reverse transcribed using random hexamers and the MiRNA-X miRNA First-Strand Synthesis Kit (Clontech Laboratories). Cycle threshold (Ct) values were determined for each sample based on three biological replicates and two technical replicates from each of the biological replicate. Constitutively expressed gene UBC9 (Ubiquitin-conjugating enzyme 9) was used as an endogenous control. Comparison of gene expression was performed in the two samples, R and S, each normalized using an endogenous control with S as the calibrator. Fold change was calculated using the formula $2-\Delta\Delta Ct$ = ((Ct gene of interest − Ct endogenous control)R − (Ct gene of interest − Ct endogenous control)S) (Schmittgen and Livak 2008).

3.3 Results

3.3.1 Microscopic examination of *P. brassicae* infection of *B. napus*

Microscopic examination of the roots of the resistant (R) and susceptible (S) plants inoculated with *P. brassica* pathotype 3 showed some evidence of infection at 6 days post inoculation (dpi) (Table 3.1). Infection was observed in the secondary roots at 10 dpi in both S

(Fig. 3.1C) and R (Fig. 3.1D) plants. Approximately 7.3% of cells were observed to be infected in S while 2.3% of cells were observed in R (Table 3.1); however, there was no swelling of the roots in either S or R. At 14 dpi, 8.3% cells were observed to be infected in S compared with 2.7% in R (Table 3.1). A greater number of primary plasmodia were observed in the S plants as compared with the R plants (data not shown). At 18 dpi (Fig. 3.1E and F), approximately 51.7% of cells showed the presence of secondary plasmodium in S plants whereas 2.7% of cells in R already showed the presence of resting spores. The root cells of the S plants started to show evidence of hypertrophy at 18 dpi (Fig. 3.1E). As expected, there was no indication of gall formation in the inoculated R plants and they were morphologically similar to the uninoculated R and S plants. At 22 dpi (Fig. 3.1G and H), more than 60% of cortical cells of the S plants showed the presence of mature secondary plasmodia and resting spores (Table 3.1), whereas 2.7% of infected cortical cells with resting spores could be seen in the R plants, because there was no further increase in infection observed. In conclusion, this study showed that infection by *P. brassicae* can occur in both S and R plants at 10 dpi; however, the infection did not progress further in the R plants. Based on these results, we focused on the root samples collected at 10 dpi for studies on lncRNA profiling.

Table 3.1 Percentage of infected cells observed in clubroot-susceptible (S) and -resistant (R) *Brassica napus* lines as observed microscopically at different time points[a]

Time point	S lines	R lines
0 hpi	**0**	**0**
6 dpi	**0.7**	**0**
10 dpi	**7.3**	**2.3**
14 dpi	**8.3**	**2.7**
18 dpi	**51.7**	**2.7**
22 dpi	**66.0**	**2.7**

hpi = hours post inoculation and dpi = days post inoculation; 300 cells were examined microscopically for each time point.

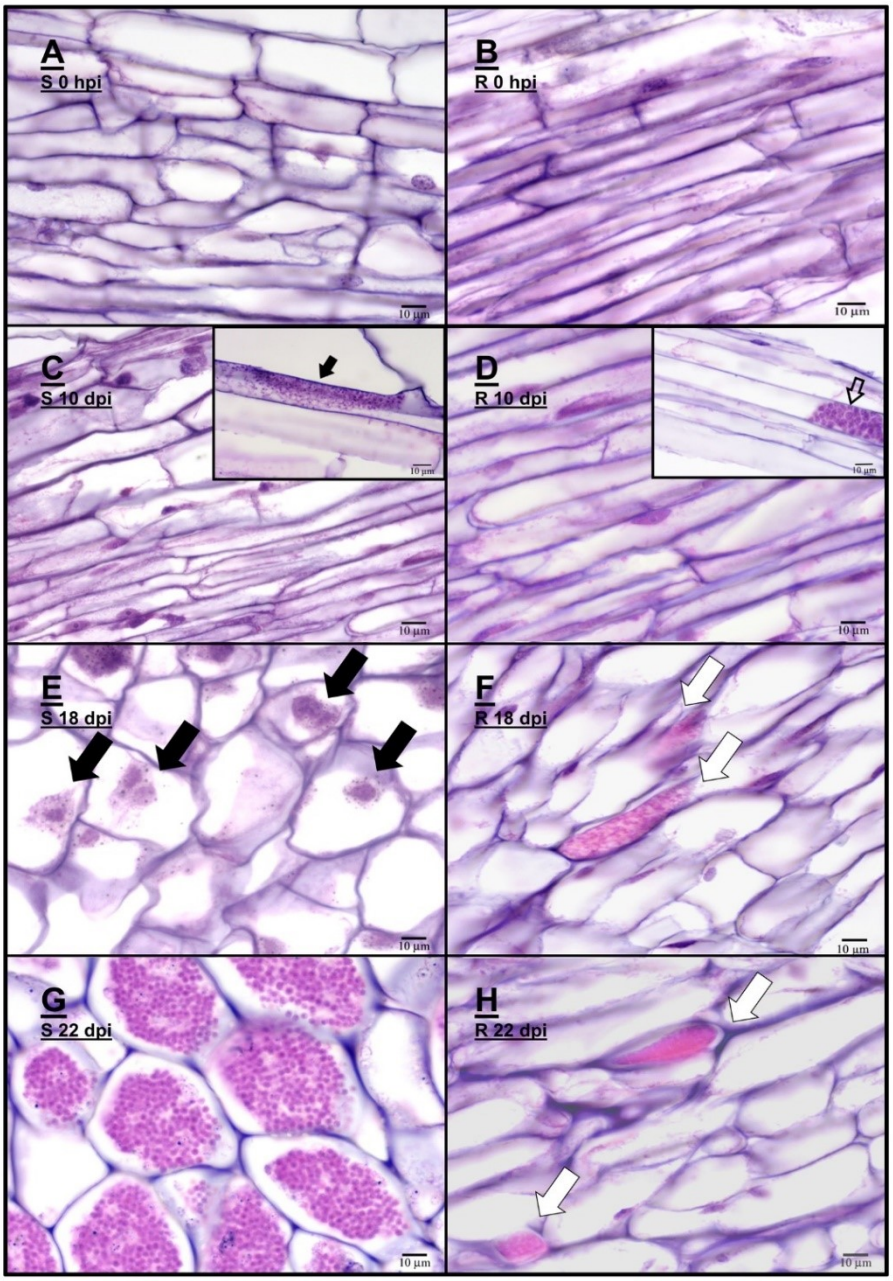


Fig. 3.1 Different stages of *Plasmodiophora brassicae* infection observed in histopathological analysis of *Brassica napus* roots (cortical tissue) at A) and B) 0 h post inoculation (hpi) and 10, 18, and 22 days post inoculation (dpi). C) and D) Presence of proliferating zoosporangia in both rutabaga-derived clubroot-susceptible (S) and clubroot-resistant (R) *Brassica napus* at 10 dpi (inset shows larger magnification); E) and F) matured secondary plasmodia in the cortical cells of S (black arrows) and resting spores in few cortical cells of R (white arrows) at 18 dpi; and G) and H) many resting spores and secondary plasmodia in S and resting spores in very few

cortical cells of R (white arrows) at 22 dpi. Bars represent the resolution of the microscopy for the particular section.

3.3.2 Identification of lncRNAs in *B. napus* following *P. brassicae* infection

Strand-specific lncRNA-Seq libraries were prepared from the roots of the R and S plants collected at 10 dpi to identify the lncRNAs responsive to *P. brassicae* infection. In total, 461 million paired-end clean reads were obtained from six strand-specific RNA-seq libraries (Table 3.2). The error rate was 0.01%, with Q30 percentage exceeding 98% for all the six libraries. After removal of the low-quality sequences and the adapter sequences and artifacts, 367 million (79.8%) clean reads were mapped on to the reference *B. napus* genome; however, by use of TopHat (v2.0.9) (Trapnell et al. 2012), 73.5% of the reads could be uniquely mapped, followed by assembly and annotation of all the transcripts in *B. napus* using Cufflinks (v2.1.1) (Trapnell et al. 2010). From the total 198,553 assembled transcripts, 78.0% of transcripts were found to be mRNAs after the application of exon number and transcript length filters, which were further filtered to remove the transcripts showing an overlap with the exon regions of the annotated functional genes (as identified using Cuffcompare, v2.1.1) (Trapnell et al. 2010). Expression of these transcripts was calculated using Cuffdiff (v2.1.1) and the transcripts with fragments per kilobase of transcript per million mapped reads (FPKM) ≥ 0.5 were selected for further analysis. The Coding Potential Calculator (CPC) (Kong et al. 2007) and Pfam Scan (v1.3) (Punta et al. 2012) were used to filter the remaining transcripts for their coding potential to distinguish the protein-coding and noncoding sequences and, in total, 4,558 transcripts were deemed to be lncRNAs. These lncRNAs included 84.7% long-intergenic RNAs and 15.3% antisense RNAs. Screening of the putative lncRNAs based on q value (<0.05) identified 530 differentially expressed (DE) lncRNAs between the R and S plants, and this included 58 (10.9%) antisense lncRNAs and 472 (89.1%) lincRNAs.

Table 3.2 Mapping statistics of the sequencing reads from six-stranded *Brassica napus* libraries post infection with *Plasmodiophora brassicae*

Replicates[a]	Total	Total mapped (%)	Multiple mapped (%)	Unique mapped (%)
R_1	**78,145,602**	**62,866,707 (80.45)**	**4,874,861 (6.24)**	**57,991,846 (74.21)**
R_2	**77,724,002**	**62,072,779 (79.86)**	**4,964,474 (6.39)**	**57,108,305 (73.48)**
R_3	**72,187,808**	**57,818,092 (80.09)**	**4,454,507 (6.17)**	**53,363,585 (73.92)**
S_1	**85,702,896**	**68,181,904 (79.56)**	**5,206,587 (6.08)**	**62,975,317 (73.48)**
S_2	**72,543,352**	**57,417,259 (79.15)**	**4,640,617 (6.4)**	**52,776,642 (72.75)**
S_3	**74,959,046**	**59,616,509 (79.53)**	**4,735,211 (6.32)**	**54,881,298 (73.22)**
Total	**461,262,706**	**367,973,250 (79.77)**	**28,876,257 (6.26)**	**339,096,993 (73.51)**

[a]R_1, R_2, and R_3 are the biological replicates of clubroot (CR)-resistant *B. napus* and S_1, S_2, and S_3 are the biological replicates of CR-susceptible *B. napus*.

3.3.3 Characterization and functional analysis of DE lncRNAs in response to *P. brassicae* infection

More than 60% of the lncRNAs were observed to be 200 to 1,000 nucleotides (nt) in size and the lengths of the 530 DE lncRNAs ranged from 206 to 8,584 nt. Approximately 18% (96 of 530) lncRNAs were longer than 1.5kb nt in size (Supplementary Fig. 3.1). Analysis of all of the DE lncRNAs for the size of their exons indicated that 78% (663 of 853 exons) had a length of 18 to 1,000 nt. Moreover, lncRNAs were observed to be significantly shorter compared with mRNAs. The number of transcripts for lncRNAs with a single exon constituted approximately 62% (329 of 530) of these lncRNAs, while 26% (138 of 530) consisted of two exons and only

2.6% (14 of 530) lncRNAs contained more than five exons. The DE lncRNAs were evenly distributed across the 19 *B. napus* chromosomes. Of the total 530 DE lncRNAs, 184 and 229 were observed to be located on the 10 "A" chromosomes and the nine "C" chromosomes, respectively, of the *B. napus* genome. However, 117 DE lncRNAs could not be assigned to any specific chromosome. A strong correlation (0.76) was observed between the distribution of the number of DE lncRNAs and chromosome size. When the distribution of DE lncRNAs on individual chromosomes was considered, the highest number was found on chromosome C03, followed by chromosomes C09, A09, and BnaA03. Interestingly, chromosome BnaA08, to which the CR resistance gene (*Crr1a, Crr1b*) has been previously mapped, was found to harbor 24 DE lncRNAs.

Of the 530 DE lncRNAs, 345 were upregulated and 185 were downregulated in the S plants as compared with R plants (Supplementary Fig. 3.2). In contrast, 170 lncRNAs were upregulated and 337 were downregulated in the R plants as compared with S plants. In all, 15 lncRNAs were detected only in the R plants while eight lncRNAs were detected only in the S plants (Table 3.3). The top 20 upregulated and downregulated lncRNAs of the R plants are presented in Supplementary Table 3.2. In all, 4 of the top 20 downregulated lncRNAs in R are located on chromosome BnaA08. Recent insights from functional analysis of the lncRNAs have demonstrated that they can be involved in regulating the expression of the overlapping or neighboring genes (cis-acting) or genes located at a long distance (transacting). Genomic regions 100 kb upstream and downstream of the lncRNAs were scanned to predict the co-located target genes of cis-acting lncRNAs (Supplementary Table 3.3). All 530 DE lncRNAs were found to pair with the annotated mRNAs from both upstream and downstream regions (Supplementary Fig. 3.3). KEGG pathway enrichment analysis revealed that the majority of the co-located target genes were from plant–pathogen interactions (15%) and the phenylpropanoid biosynthetic pathway (15%) (Fig. 3.2A). Similarly, 425 DE lncRNAs were predicted to have co-expressed targets based on correlation of expression between the lncRNAs and mRNAs. KEGG pathway enrichment analysis also revealed that the majority of the co-expressed target genes belonged to primary metabolic pathways such as biosynthesis of amino acids (19%) and carbon metabolism (18%) (Fig. 3.2B). Some of the most significantly expressed lncRNAs in R and S plants and their predicted target genes are shown in Fig. 3.3. A majority of the target genes regulated by these lncRNAs belonged to the plant defense-related pathways such as plant–pathogen interactions,

phenylpropanoid biosynthetic pathways, plant hormone-signaling pathways, and primary metabolic pathways (Fig. 3.3). Genes involved in plant–pathogen interactions include the calcium-dependent protein kinase, defensin, pathogen-related protein, and disease resistance protein (leucine-rich repeat (LRR) receptor protein) genes (Fig. 3.3); and important genes from the phenylpropanoid biosynthetic pathway upregulated in R include caffeic acid 3-O-methyltransferase (*CoMT*), shikimate O-hydroxycinnamoyltransferase (*HCT*), and cinnamoyl alcohol dehydrogenase 7 (*CADH7*) (Fig. 3.3).

Table 3.3 Long noncoding RNAs (lncRNAs) uniquely expressed in clubroot disease-resistant (R) lines carrying the resistance gene of rutabaga and clubroot-susceptible (S) lines of *Brassica napus*[a]

Transcript ID	Chromosome name and gene location	S (FPKM)	R (FPKM)
LNC_001161	A08: 12,183,521–12,183,982	0	0.954916
LNC_001063	A08: 3,202,637–3,204,852	0	1.02438
LNC_000600	A04: 14,816,008–14,816,758	0	1.3608
LNC_001126	A08: 3,447,903–3,448,931	0	1.7121
LNC_000450	A03: 27,099,999–27,101,445	0	1.98082
LNC_002916	C04: 47,030,126–47,030,803	0	1.98632
LNC_001146	A08: 9,573,084–9,573,548	0	2.28077
LNC_000108	A01: 19,635,847–19,636,338	0	2.30396
LNC_001065	A08: 3,788,461–3,789,000	0	2.92531
LNC_001154	A08: 10,153,693–10,154,043	0	3.7843
LNC_001089	A08: 12,365,659–12,366,252	0	5.01747
LNC_004045	C09: 43,229,826–43,230,101	0	6.3523
LNC_001851	Ann_random: 41,970,174–41,971,127	0	1.4904
LNC_001191	A08_random: 1,215,048–1,217,686	0	3.3501
LNC_001888	Ann_random: 47,682,468–47,684,525	0	5.2681
LNC_000089	A01: 12,314,342–12,314,908	2.07049	0
LNC_001341	A09: 11,766,717–11,767,907	2.1138	0
LNC_002432	C03: 26,497,021–26,497,291	5.19373	0

LNC_003616	C07: 28,269,085–28,273,791	1.93572	0
LNC_004790	Unn_random: 8,133,944–8,134,628	0.8604	0
LNC_001807	Ann_random: 35,009,658–35,010,033	1.4708	0
LNC_000887	A06_random: 299,855–300,207	1.5085	0
LNC_000736	A05_random: 1,032,000–1,032,588	0.8626	0

[a]FPKM = fragments per kilobase of transcript per million mapped reads. Rows in bold indicate the lncRNAs located on chromosome BnaA08.

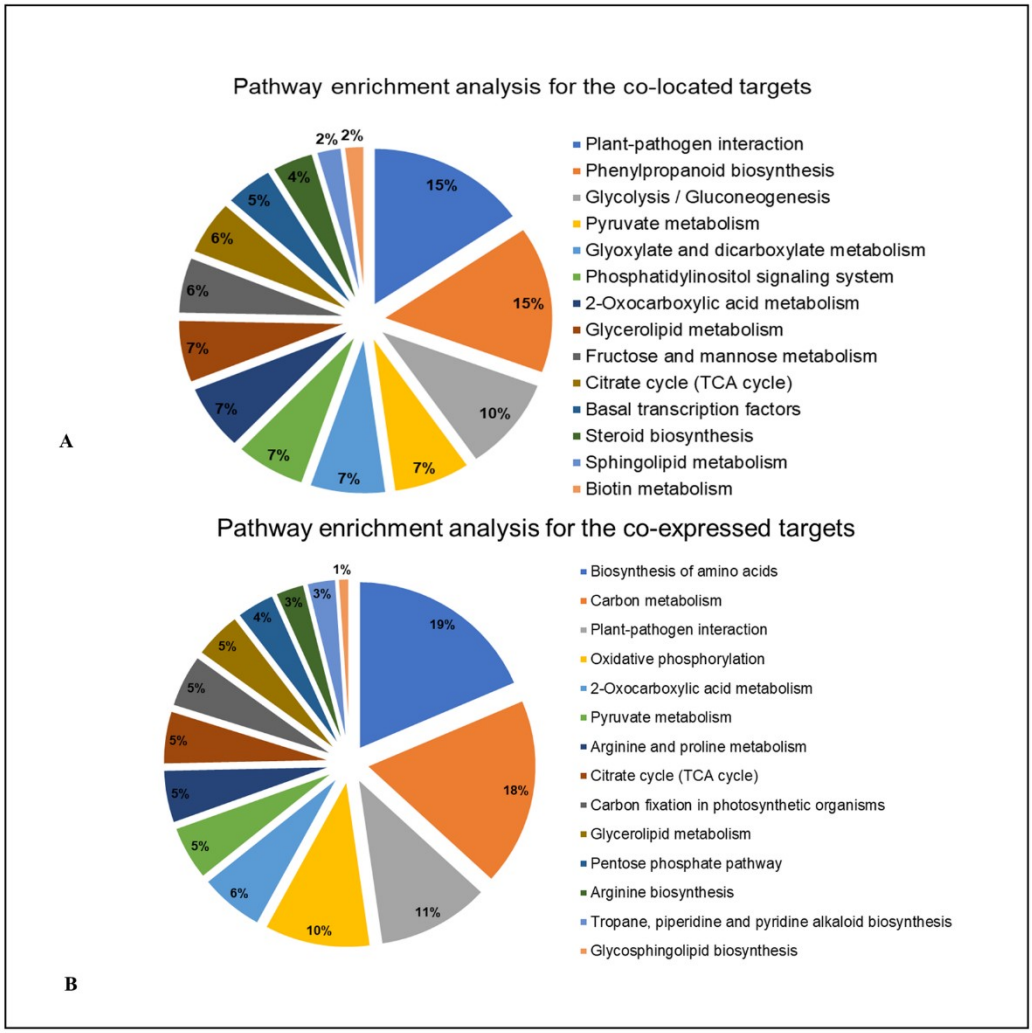
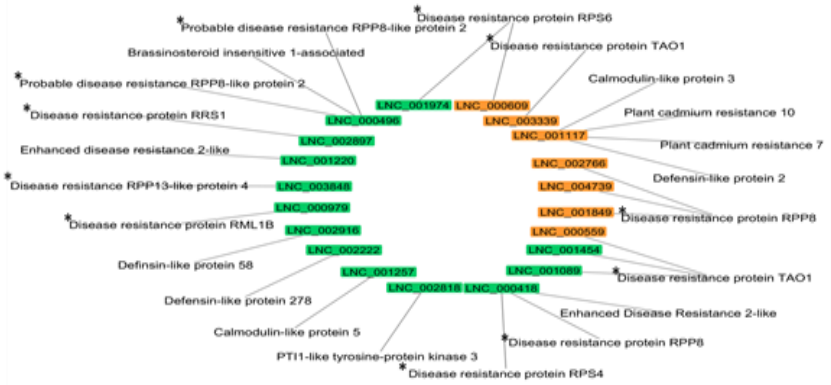


Fig. 3.2 Functional annotation of target genes of *Brassica napus* long noncoding RNAs A) Pathway enrichment analysis for the co-located targets and B) pathway enrichment analysis for the co-expressed targets [Statistical significance of the enrichment, *p*-value (<0.05) calculated using KOBAS v.20]

RESISTANT CANOLA (R)

PLANT-PATHOGEN INTERACTION



SUSCEPTIBLE CANOLA (S)

PLANT HORMONE SIGNALING

PRIMARY METABOLISM

PHENYLPROPANOID BIOSYNTHESIS

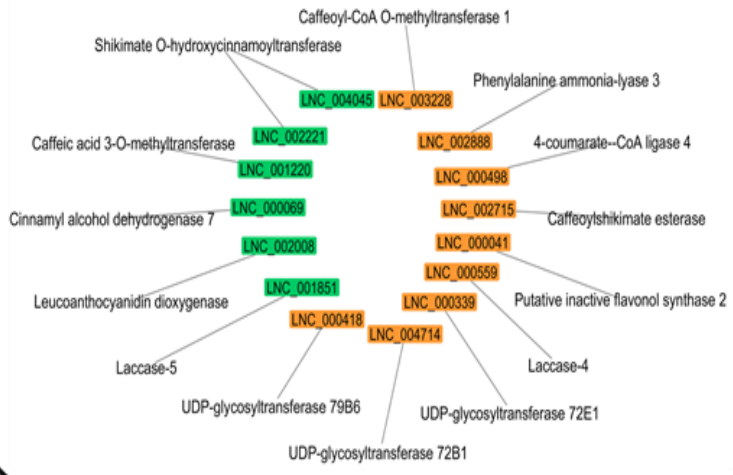


Fig. 3.3 Classification of long noncoding RNAs with the most significant differential expression and their predicted targets into the most represented functional groups. lncRNAs upregulated in R (green), lncRNAs upregulated in S (orange), and target genes with an asterisk are leucine-rich repeat protein genes.

3.3.4 Differentially expressed genes located on the BnaA08 chromosome

Of the DE lncRNAs detected at 10 dpi, 24 lncRNAs were distributed on the chromosome BnaA08 (Fig. 3.4); this chromosome has been reported to carry a genomic region conferring resistance to five *P. brassicae* pathotypes (Hasan and Rahman 2016). Among these, LNC_001065, LNC_001165, and LNC_001146 target the largest number of genes involved in plant defense (Fig. 3.4). While comparing their expression in the R and S plants, a greater number of the 24 DE lncRNAs were found to be upregulated in the S plants as compared with the R plants. In addition, 8 of the 24 lncRNAs were detected only in the R plants; however, they showed a low and inconsistent expression among the three biological replicates (Fig. 3.5; Supplementary Table 3.4). Genes targeted by these 24 lncRNAs were found to be involved in primary metabolic pathways, plant–pathogen interactions, and secondary metabolite and phenylpropanoid biosynthetic pathways (Fig. 3.4).

All 24 DE lncRNAs, which were located on chromosome BnaA08, were found to pair with the neighboring annotated target mRNAs (cis-regulation), and 21 of these were predicted to transregulate the target mRNAs. Among the 21 transregulated lncRNAs, LNC_001191, LNC_001126, LNC_001063, LNC_001117, LNC_001118, and LNC_001167 (Fig. 3.5) are known to be involved in the regulation of disease resistance protein genes distributed throughout the *B. napus* genome (Supplementary Table 3.5). Furthermore, LNC_001167 was predicted to be involved in the regulation of three different disease resistance genes encoding toll interleukin 1 receptor nucleotide-binding site LRR proteins, similar to those found in *A. thaliana*, including *RPS4* (resistance to Pseudomonas syringae 4), *RPP1* (recognition of Peronospora parasitica 1), and *EDR2* (enhanced disease resistance 2). Expression of LNC_001063, LNC_001126, and LNC_001191 and could not be detected in the R plants whereas a 0.7- to 2.7-fold increase in expression was observed for LNC_001117, LNC_001118, and LNC_001167 in the S plants.

When considering both lncRNAs and their targets located on chromosome BnaA08, 15 of the abovementioned 24 lncRNAs were predicted to transregulate their target genes located on

this chromosome (Fig. 3.5; Supplementary Table 3.6). No expression was observed for LNC_001063, LNC_001065, LNC_001089, LNC_001126, LNC_001146, LNC_001154, or LNC_001191 in the S plants, whereas a low expression was detected in the R plants (Supplementary Table 3.7). Based on KEGG pathway analysis, targets of these seven lncRNAs were found to be involved in the biosynthesis of secondary metabolites and in the primary metabolic pathways, including carbon and amino acid metabolism (Fig. 3.4). Among all of the lncRNAs located on chromosome BnaA08, the greatest level of expression was observed for LNC_001165 in the S plants (Fig. 3.5; Supplementary Table 3.8). This lncRNA was predicted to transregulate 76 genes in the *B. napus* genome; however, 29 of these target genes are located on chromosome BnaA08 and are known to be involved in carbon metabolism and biosynthesis of secondary metabolites.

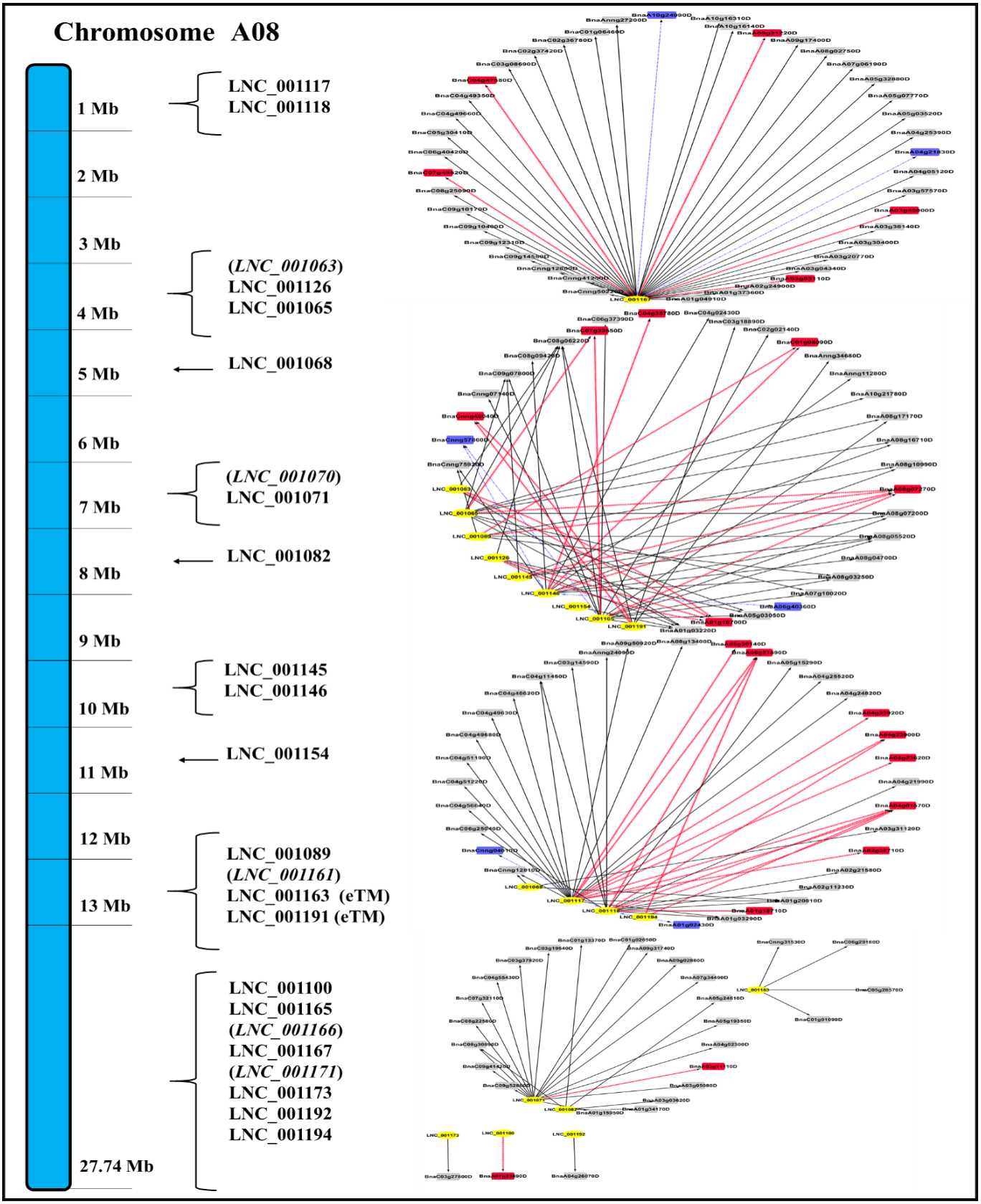

Chromosome A08
1 Mb
2 Mb
3 Mb
4 Mb
5 Mb
6 Mb
7 Mb
8 Mb
9 Mb
10 Mb
11 Mb
12 Mb
13 Mb
27.74 Mb
LNC_001117
LNC_001118
(LNC_001063)
LNC_001126
LNC_001065
LNC_001068
(LNC_001070)
LNC_001071
LNC_001082
LNC_001145
LNC_001146
LNC_001154
LNC_001089
(LNC_001161)
LNC_001163 (eTM)
LNC_001191 (eTM)
LNC_001100
LNC_001165
(LNC_001166)
LNC_001167
(LNC_001171)
LNC_001173
LNC_001192
LNC_001194

Fig. 3.4. Long noncoding RNAs (lncRNAs) on chromosome BnaA08 and their interaction with target genes in *Brassica napus* genome involved in plant defense. lncRNAs not regulating any target genes involved in plant defense are shown in italics and parentheses. Yellow ovals = lncRNAs on chromosome BnaA08, red squares = target genes involved in plant-pathogen interaction, blue squares = target genes involved in plant hormone signaling, gray squares = target genes involved in primary and secondary metabolism, and eTM = endogenous target mimic.

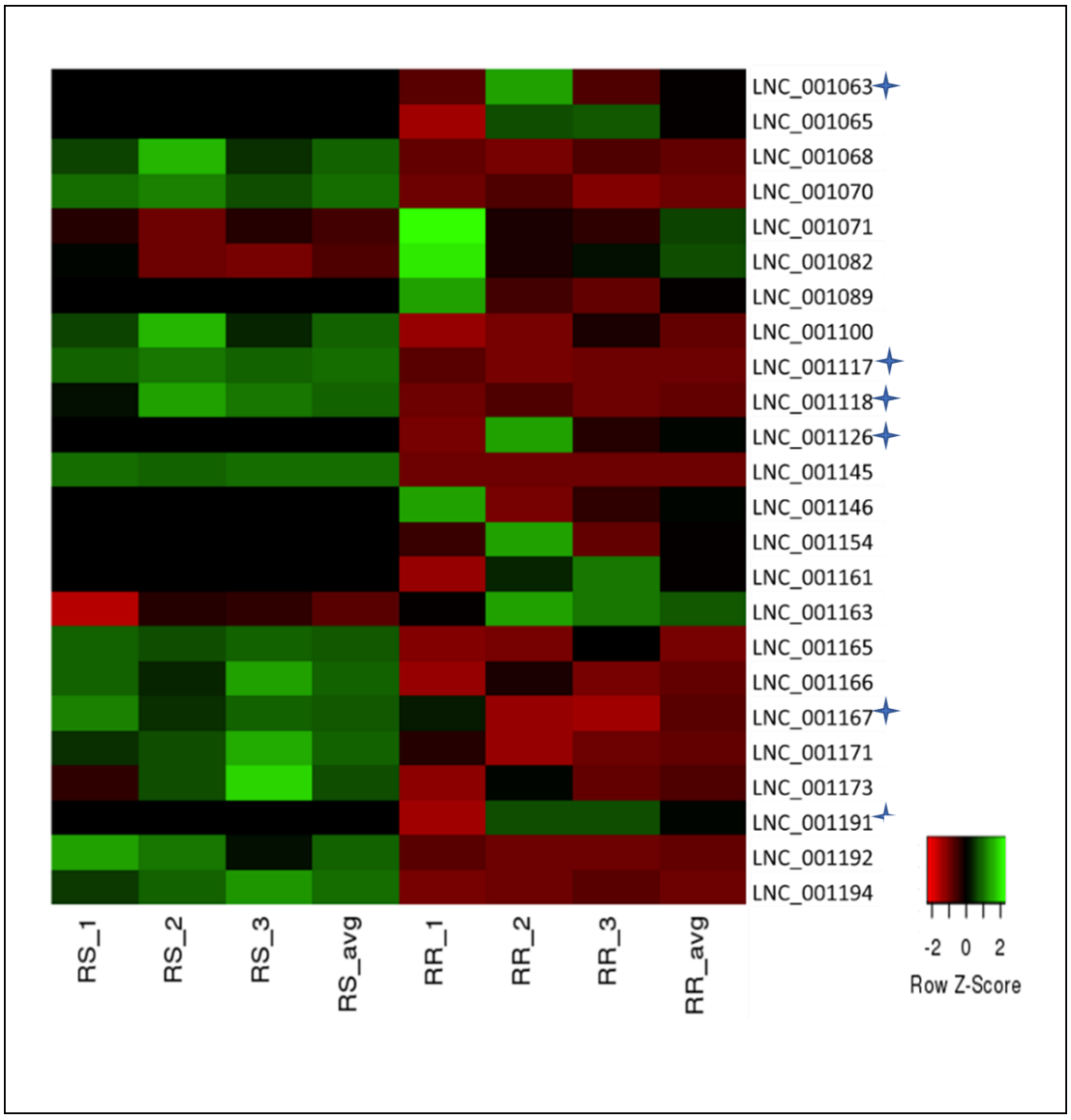


Fig. 3.5 Differentially expressed long noncoding RNAs (lncRNAs) from chromosome BnaA08 of *Brassica napus*. Expression values (FPKMs) are scaled from +2 (green) to −2 (red); black indicates no expression. S_1, S_2, and S_3 are the biological replicates of clubroot (CR)-susceptible *B. napus* and R_1, R_2, and R_3 are the biological replicates of CR-resistant *B. napus*. lncRNAs with asterisks are known to target genes encoding disease resistance proteins.

3.3.5 Quantitative real-time PCR validation for the expression of lncRNAs.

To validate our findings from lncRNA-seq analysis and to further determine whether these genes are involved in defense responsive pathways, three lncRNAs from BnaA08, three lncRNAs that are located on different chromosomes but have at least some of their target genes located on BnaA08, and three lncRNAs were randomly selected for quantitative real-time (qRT)-PCR validation. While selecting the lncRNAs, only the lncRNAs which were predicted to have a maximum of four targets were used. Expression patterns of these nine lncRNAs were found to be similar in lncRNA-seq and qRT-PCR analysis (Fig. 3.6A); however, all of the target genes of an lncRNA did not follow the same pattern (Fig. 3.6B). For example, the three targets BnaA09g52970D, BnaC03g59610D, and BnaA08g06670D of LNC_003848 showed an expression profile similar to the results of lncRNA-seq; however, the fourth target (BnaA07g30470D) followed an opposite trend. It is probable that the target BnaA07g30470D is a false positive with q value > 0.05. A similar situation was also found in the case of the targets BnaC04g09450D, BnaC03g27800D, and BnaA07g23650D of LNC_001173. Interestingly, such anomalies were not observed in the cases where lncRNAs were predicted to have only one target, such as LNC_003279 and LNC_001192. Finally, the expression of two additional genes, BnaA10g17370D and BnaC09g40660D, showed excellent correlation between RNAseq (unpublished) and qRT-PCR results.

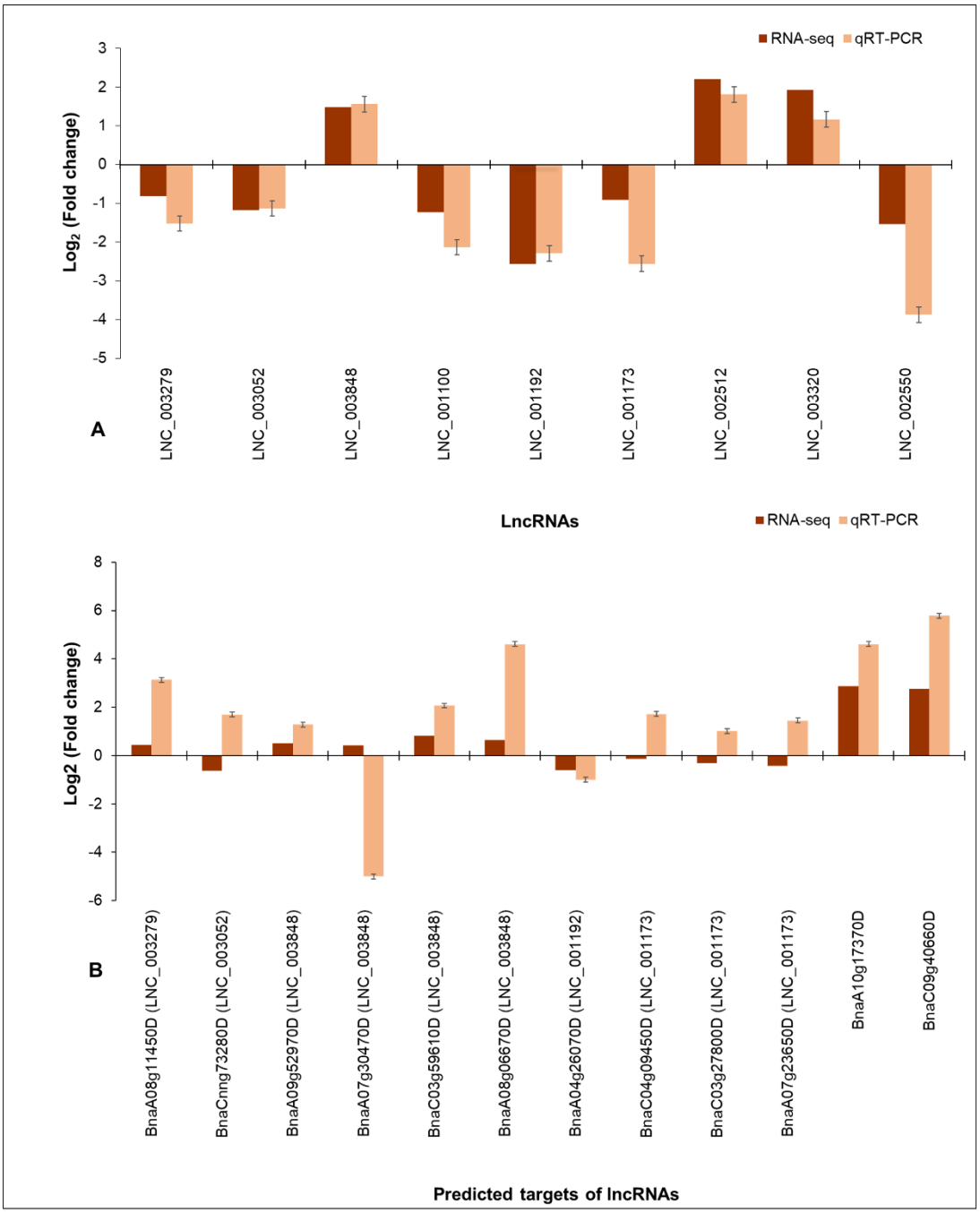


Fig. 3.6 A) Quantitative real-time (qRT)-PCR validation of pathogen-induced nine long noncoding RNAs (lncRNAs) detected by lncRNA-seq. B) qRT-PCR validation of lncRNAseq data for 10 targets predicted by correlation of expression with lncRNAs and two randomly selected genes. [The expression of lncRNAs and target genes was normalized to *UBC9* (endogenous control) and lncRNAs regulating their respective targets are indicated within parentheses. Log_2 (fold change) in R lines calculated using the formula $2^{-\Delta\Delta Ct} = ((Ct_{\text{gene of}}$

interest − Ct endogenous control)R − (Ct gene of interest − Ct endogenous control)S)]. Three biological replicates and two technical replicates were used. Error bar = standard error]

3.3.6 *B. napus* lncRNAs as endogenous target mimics for microRNAs

By use of the RNA target analysis server psRNAtarget (Dai et al. 2018), with default parameters and maximum expectation of 3.5, 15 DE lncRNAs were identified which could possibly serve as target mimics of microRNAs (miRNAs) (Supplementary Table 3.9). Among these, LNC_001191 and LNC_001163 were found to be located on chromosome BnaA08 and as targets of the miRNAs bna-miRNA6028 and bna-miRNA824, respectively. Putative target AGL16 and the target mimic LNC_001163 of bna-miRNA824 are shown in Fig. 3.7A. Among these two lncRNAs, LNC_001191 was expressed only in the R plants, and a greater level of expression was found for LNC_001163 (Fig. 3.5). qRT-PCR analysis also confirmed the upregulation of LNC_001163, bna-miRNA824, and its target AGL16 in the R plants (Fig. 3.7B). It was found that bna-miRNA6028 targets three lncRNAs (LNC_002844, LNC_001263, and LNC_001191); among these, the expression of LNC_002844 and LNC_001263 was downregulated in the R plants to a greater extent as compared with LNC_001191. From this data, six lncRNAs were predicted to target single miRNAs whereas four lncRNAs were predicted to target multiple miRNAs of the same family. The interaction analysis using Cytoscape v.3.7.0 (Fig. 3.8) depicts intricate interactions between the lncRNAs and miRNAs and mRNAs. Therefore, the abovementioned endogenous target mimics (eTMs) identified in response to *Plasmodiophora brassicae* infection were predicted to conserve the mRNA targets by disrupting the miRNA–mRNA interactions.

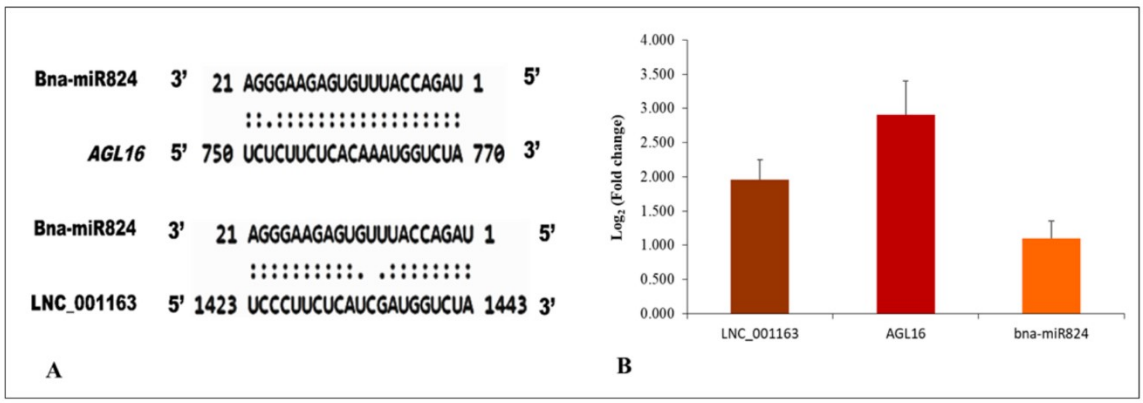


Fig. 3.7 Putative endogenous target mimic (eTM) LNC_001163 of *AGL16* inhibiting bna-miR824 in response to clubroot infection in *Brassica napus*. A) Predicted base-pairing interactions between bna-miR824 and *AGL16* and eTM long noncoding RNA LNC_001163 and bna-miR824. B) Quantitative real-time PCR results for LNC_001163, *AGL16*, and bna-miR824. [The expression of lncRNAs and target genes was normalized to *UBC9* (endogenous control). Log_2 (fold change) in R lines was calculated using the formula $2^{-\Delta\Delta Ct} = ((Ct_{\text{gene of interest}} - Ct_{\text{endogenous control}})R - (Ct_{\text{gene of interest}} - Ct_{\text{endogenous control}})S)$]. Three biological replicates and two technical replicates were used. Error bar = standard error].

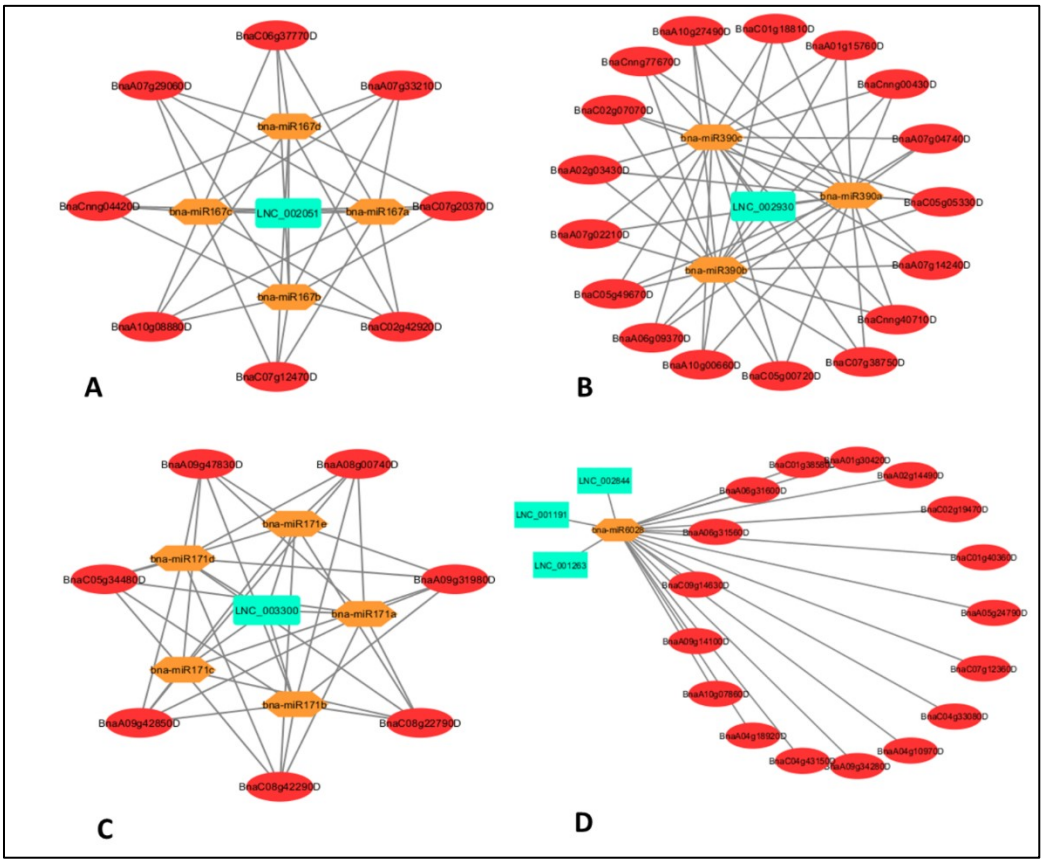


Fig. 3.8 A) B) and C) Interaction networks of a single long noncoding RNAs (lncRNAs), microRNA (miRNA) family, and the mRNAs; and D) interaction network of multiple lncRNAs with miRNA and mRNAs.

3.3.7 lncRNAs as potential miRNA precursors

lncRNAs can also function as precursors for miRNA and siRNA, playing important roles in modulating gene expression and, thus, implicating their roles in diverse biological functions in plants. In this study, 8 of the 530 DE lncRNAs were identified to be possible precursors of miRNA (Supplementary Table 3.10). For instance, LNC_001710 seems to harbor 100% of the

stem-loop sequence of miRNA-160a; and LNC_000265 and LNC_004360 were predicted to be the precursors for miRNA169n and miRNA166a, respectively (Fig. 3.9), and were expressed at higher levels in the R plants. Similarly, miRNA166a was also predicted to arise from LNC_00563; however, expression of this lncRNA was not significantly different between the R and S plants. The miRNA166 is a family of six miRNAs in *B. napus*, of which miRNA166a (LNC_004360) and miRNA166e (LNC_001636) were represented in this study.

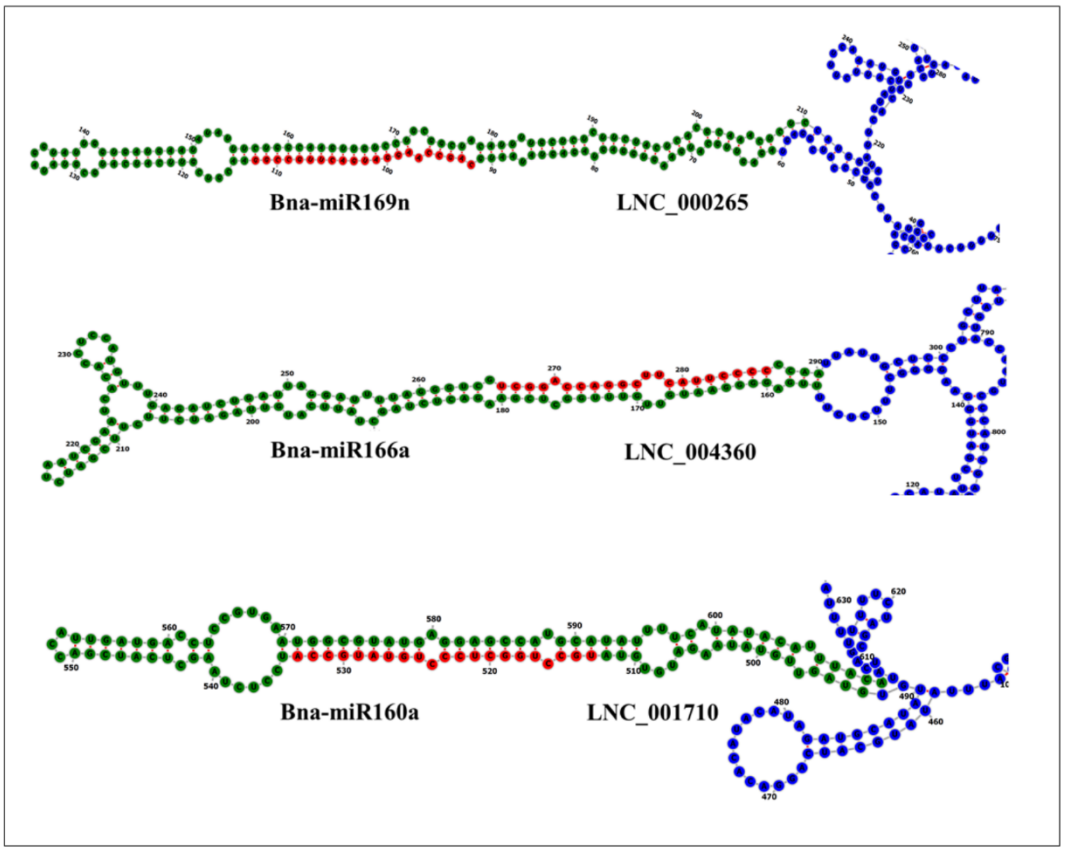


Fig. 3.9 Representative predicted long noncoding RNA functioning as microRNA (miRNA) precursor sequences.

3.3.8 Natural antisense transcripts

In total, 59 natural antisense transcripts (NATs) was identified to be regulating their corresponding sense transcripts in response to infection by *P. brassicae* at 10 dpi (Supplementary Table 3.11). All of these NATs were validated *in silico* by determining the thermodynamics of the RNA–RNA interaction (Hofacker 2009; Lorenz et al. 2011). Based on differential expression

of the lncRNAs, 38 NATs were identified. After filtering the 38 NATs based on expression of their sense transcripts in the RNA-seq data and potential pathogen defensive functions of the genes, 18 NATs were selected, of which 4 were validated using qRT-PCR (Fig. 3.10). The optimal secondary structure formation happens when thermodynamic free energy is minimum; such minimum free-energy structures of these four lncRNA-mRNA dimerized pairs, predicted using RNAcofold (Lorenz et al. 2011), are shown in Fig. 3.10.

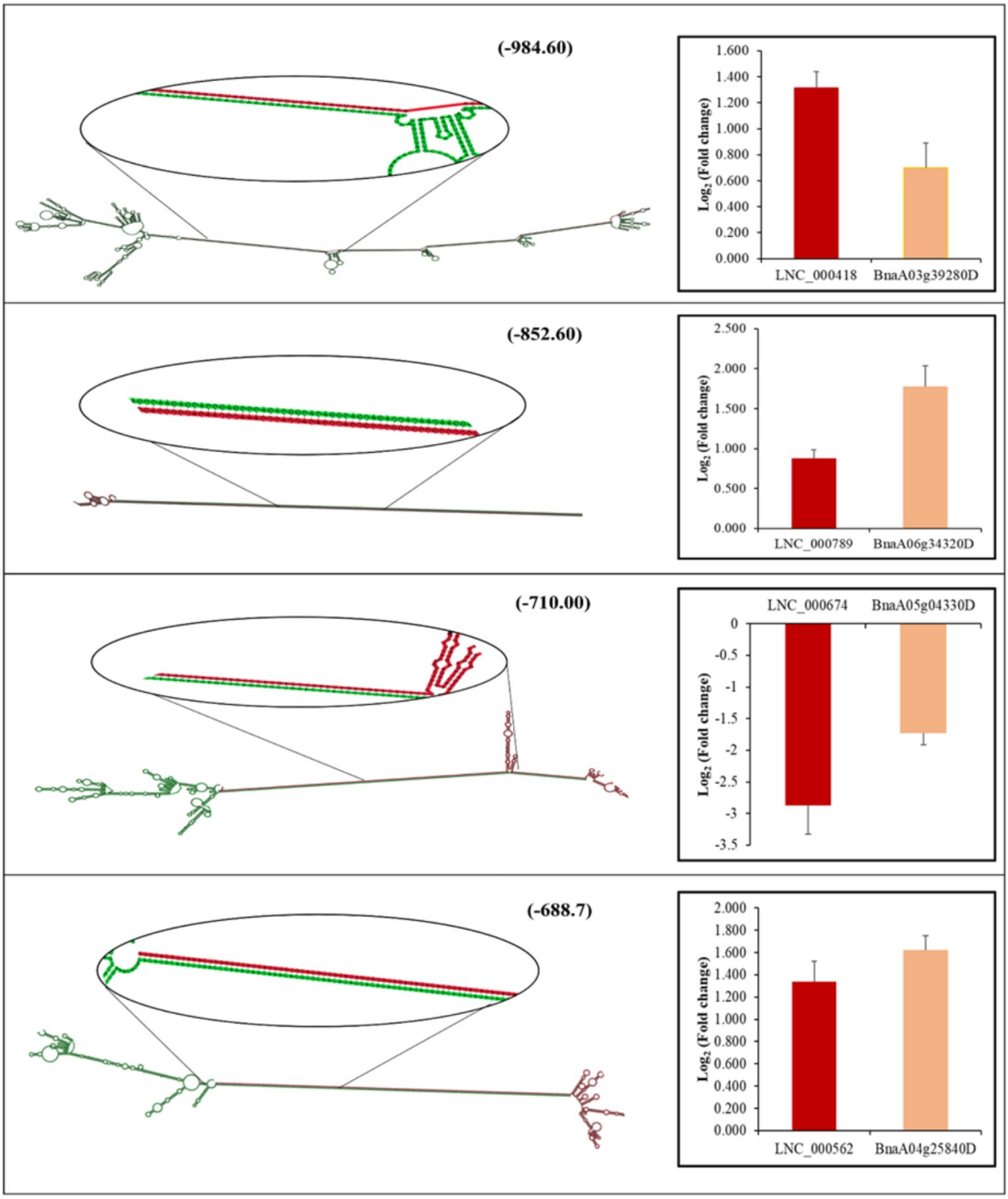


Fig. 3.10 Minimum free energy (MFE) structures for representative long noncoding RNA-mRNA pairs. MFE for each secondary structure is given in the bracket. Quantitative real-time PCR results for each pair are shown on the right side.

3.4 Discussion

CR disease is known to cause significant yield loss in *Brassica* spp., including vegetables and oilseed crops. Significant achievements have been made in combating this disease through the development of R cultivars (Piao et al. 2004; Rahman et al. 2014); however, the molecular mechanisms involved in the host plant resistance have not been well understood. Results from our study indicate the possible involvement of lncRNAs in mediating resistance in *B. napus* canola lines carrying resistance introgressed from rutabaga. We did not find evidence of infection in the roots at 6 dpi in both R and S plants; however, infection was observed in both R and S plant roots at 10 dpi. In contrast, Dobson et al. (1983) observed root hair infection in Chinese cabbage at 1 dpi. However, Asano et al. (2000) detected initial infection in turnip at 4 dpi with rapid increase up to 6 dpi; similar results have also been reported by Hwang et al. (2012) in canola. Thus, it is evident that the stage of root hair infection may be dependent on the type of plant material and the method of inoculation which, in the present study, was injection of spore suspension into the soil (direct inoculation). A noticeable reduction in the formation of primary plasmodia was found in the R plants, with no infection in the newly emerging roots at 14 dpi, whereas gall formation had already begun in the S plants at this stage. Detection of infection at an early stage in the R plants could be related to the plant development; an influence of the developmental stage of the host plant on the pathogen's ability to cause infection has been suggested by Fei et al. (2016). Thus, our study suggests that, despite microscopically visible infection that occurs in the R plants at their early growth stage, the resistance genes restrict further development of the disease.

Most of the 530 DE lncRNAs detected in this study were from the intergenic regions (lincRNAs); these lncRNAs showed considerably greater expression as compared with the expression of the antisense lncRNAs (NATs), which is consistent with the results from other plant species (Li et al. 2014; Wang et al. 2014; J. Zhang et al. 2018). The lincRNAs regulate plant responses to CR by remodeling epigenetic and chromatin architecture, stabilization and localization of proteins and RNA molecules by tethering to them and transcription regulation involving enhancer-associated activity, and cis- or trans-mediated gene activation (Kopp and Mendell 2018; Kornienko et al. 2013; Nejat and Mantri 2018; Zaynab et al. 2018). To understand the function of DE lincRNAs under CR stress, we analyzed the protein coding genes co-expressed with these lincRNAs. This putative functional prediction relied on the development of

a co-expression network. The co-expression analysis indicated that 425 lncRNAs were predicted to have transregulated targets whereas all DE 530 lncRNAs were predicted to have cis-regulated targets. KEGG pathway analysis revealed that a majority of transregulated target genes belong to carbon (18%) and amino acid biosynthesis pathways (19%); however, most of the cis-regulated target genes belong to the phenylpropanoid biosynthetic pathway (15%) and are involved in plant–pathogen interactions (15%). This suggests that the energy required to fuel the defense in R plants is probably provided by the carbon metabolism. Also, it is likely that biosynthesis of specific amino acids which are upregulated or reduced are involved in the plant defense (Rojas et al. 2014). However, more comprehensive studies are required to verify this hypothesis.

lncRNAs, like NATs, are also known to mediate the regulation of the corresponding sense transcripts in a concordant or discordant manner at a transcriptional and posttranscriptional level (Rai et al. in press; Wight and Werner 2013). In *B. napus*, 111 NATs were detected in response to *S. sclerotiorum* infection at 24 and 48 h post inoculation (hpi) (Joshi et al. 2016). In contrast, we observed very limited expression of the NATs in response to *P. brassicae* infection, which is possibly due to the fact that this type of lncRNA can be expressed in a spatial and temporal manner (Wang et al. 2014). Some of the NATs such as *COOLAIR* (Ietswaart et al. 2017), *cis* $NAT_{PHO1;2}$ (Jabnoune et al. 2013), *SHO lnc-NAT* (Zubko and Meyer 2007), and *TL-lncRNA* (Liu et al. 2018) are involved in the regulation of plant morphology and biotic or abiotic responses. In this study, we found that LNC_000418, LNC_000789, LNC_000674, and LNC_000562 are antisense to the protein-coding genes Rac-like GTP-binding protein, defensin-like protein 2, flavone 3-dioxygenase 3-like protein, and WRKY transcription factor 12, respectively; all four genes are directly or indirectly involved in plant secondary metabolism. Thus, considering the function of the predicted target genes of lncRNA, we suggest that primary and secondary metabolic pathways are also regulated by NATs and may play a role in the resistance to *P. brassicae*.

Most of the target genes identified in this study belong to the primary and secondary metabolic pathways, and include the genes involved in amino acid biosynthesis, carbon metabolism, phenylpropanoid biosynthesis, and biosynthesis of secondary metabolites, indicating their probable association with the CR disease (Fig. 3.3). The importance of phenylpropanoid biosynthesis has already been demonstrated in response to biotic stress because it is known to generate an array of secondary metabolites involved in host–pathogen interactions

(Caretto et al. 2015; Vogt 2010). Genes encoding proteins such as phenylalanine ammonia lyase (Vogt 2010), laccase 4, laccase 7 (Hu et al. 2018), and UDP-glycosyltransferases (Le Roy et al. 2016) have been implicated to play a significant role in phenylpropanoid biosynthetic pathways, and they have also been identified as targets of lncRNAs in our study. It is known that caffeoyl-CoA O-methyltransferase 1*, CoMT*, *HCT*, and *CADH7* play critical roles in the biosynthetic pathway for lignification (Harakava 2005); among these genes, *CoMT, HCT*, and *CADH7* were found to be upregulated in the R plants at 10 dpi (Fig. 3.3). This suggests that the upregulation of this pathway prevents the spread of the *P. brassicae* in the roots of R plants and, thus, contributes to the decline in disease progression. Rojas et al. (2014) suggested that induction of primary metabolic pathways may be equally important because they can modulate signal transduction cascades influencing the plant defenses. Several researchers have suggested that carbon skeletons synthesized during photosynthesis are used for growth (primary metabolism) or defense responses (secondary metabolism) (Caretto et al. 2015). Tayeh et al. (2014) reported that external application of the disaccharide sugar trehalose (consisting of two molecules of glucose; it is also known as mycose or tremalose) increases the expression of *PR protein* genes in *Arabidopsis*, and external application of glucose upregulates *PR1* and *PR5* (Xiao et al. 2000). Trehalose is considered an inducer of plant defense to biotic and abiotic stresses (Tayeh et al. 2014). In this study, we found an increased expression of LNC_002915, which is predicted to regulate an α,α-trehalose phosphate synthase 10 (*TPS10*), and a considerable decrease in the expression of LNC_00675 and its targets regulating the synthesis of trehalose phosphate phosphatase A. In addition, alkaline or neutral invertase and other similar proteins are involved in carbon metabolism and are known to cleave sucrose into glucose and fructose for energy production (Ehness et al. 1997) and we observed upregulation of these and similar genes in our study (Fig. 3.2). Functional research with these lncRNAs can provide further insight into the regulatory mechanism involved in primary metabolism and, eventually, in pathogen defense.

Hasan and Rahman (2016) mapped the CR resistance of rutabaga, which was introgressed into oilseed *B. napus*, on chromosome BnaA08 . By use of this resistance, we identified eight lncRNAs showing expression only in the R plants; however, expression these lncRNAs was low and not very consistent. Kopp and Mendell (2018) suggested that many lncRNAs can exhibit low expression in plants but still play an important role. Taking this into account, it can be assumed that these lncRNAs might play an important role in CR resistance

through regulating protein coding genes or proteins by tethering to them (Fig. 3.4). Co-expression analysis of LNC_001063, LNC_001191, and LNC_001126 had predicted 8, 46, and 47 target genes, respectively, in the *B. napus* genome, and each lncRNA was found to regulate a different resistance (*R*) gene as well as genes directly or indirectly modulating primary metabolism (Fig. 3.4). Another lncRNA (LNC_001167) was also predicted to regulate three *R* genes and 35 other genes modulating primary metabolism and phenylpropanoid biosynthesis (Fig. 3.4).

We further validated the relative expression of the selected lncRNAs, and their target genes based on their direct or indirect association with chromosome BnaA08. For instance, relative expression exhibits a positive correlation and LNC_003848 was upregulated in the R plants and exhibits a positive correlation with the expression of the target gene Transparent Testa 12 (BnaA08g06670D) located on chromosome BnaA08. Genes transparent testa 12 as well as gibberellin 2-β-dioxygenase 2 (BnaC03g59610D) are known to be involved in flavonoid and plant hormone biosynthesis and were found to be transregulated by LNC_003848. Flavonoids are known to play important roles such as protecting the plants from UV radiation, pathogen attack, and photo-oxidative stress (Falcone Ferreyra et al. 2012; Treutter 2006). The specific roles of these lncRNAs, their target genes, as well as the regulatory mechanism remain to be validated.

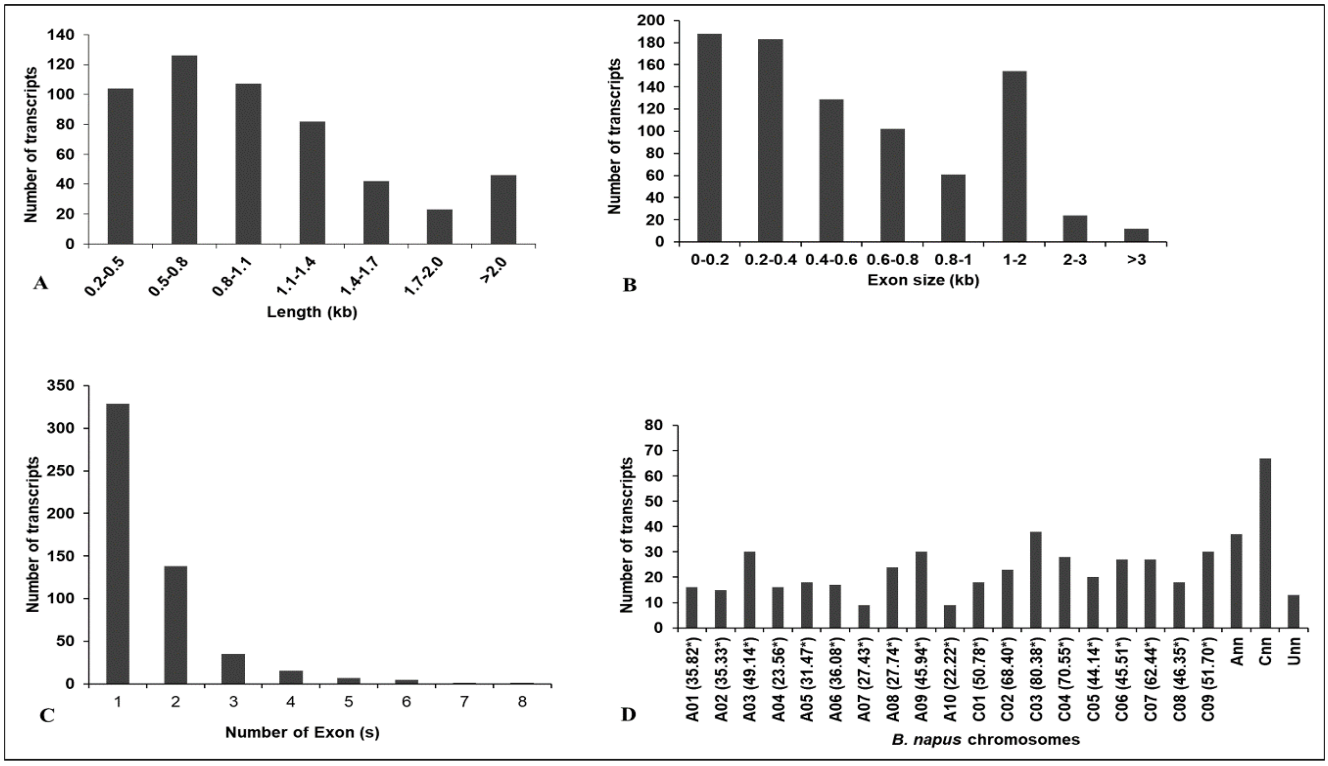
lncRNAs have also been hypothesized to form functional networks with miRNAs and mRNAs by functioning as either target mimics of miRNAs or precursors to miRNAs (Meng et al. 2012; Wu et al. 2013). Results from our study indicate that LNC_001163 and LNC_001191, located on chromosome BnaA08, could serve as an eTM of bna-miR824 and bna-miR6028, respectively. Kutter et al. (2007) reported that bna-miR824 targets the MIKC-type MADS box protein agamous-like 16 (*AGL16*), and our results indicated that the expression of *AGL16* and the predicted target mimic LNC_001163 of bna-miR824 were upregulated and positively related to each other in response to CR infection (Fig. 3.7B). Interaction analysis among the mRNAs, miRNAs, and lncRNAs suggested multiple and complex sets of interactions, including instances where a single lncRNA interacts with an miRNA family (Fig. 3.8A, B, and C) and a single miRNA interacts with multiple lncRNAs and mRNAs (Fig. 3.8D). Shumayla et al. (2017) also reported interaction of 1,047 lncRNAs with 222 miRNAs, which further interacted with 209 mRNA transcripts with various roles in *Triticum aestivum.* We also detected eight lncRNAs in

the present study to be the precursors of five miRNAs. For instance, LNC_000265 was observed to be upregulated in the R plants, and this seems to be the precursor for bna-miR169 in *B. napus* (bna-miR69n) (Fig. 3.9). bna-miR169 has been reported to be upregulated in response to drought in tomato (Zhang et al. 2011) and salt and drought stress in maize (Luan et al. 2014). Additionally, LNC_004360, a putative precursor of miRNA166a, was significantly expressed in the R plants (Fig. 3.9). The expression of miRNA166 has also been associated in response to *Botryosphaeria dothidea* in *Populus trichocarpa* (Zhao et al. 2012). Recently, J. Zhang et al. (2018) reported seven lncRNAs from *Brassica oleracea* and 15 lncRNAs from *B. rapa*, which are precursors of nine and 19 miRNAs belonging to eight and 15 miRNA families, respectively. Xin et al. (2011) and Joshi et al. (2016) demonstrated that lncRNAs can serve as precursors of miRNAs in plants and are involved in the regulation of plant responses to biotic stress. Several researchers such as Chauhan et al. (2017), Koroban et al. (2016), and Kumar (2014) also have summarized the functions of miRNAs in biotic and abiotic stresses, and results from our interaction analyses also lend credence to additional layers of gene regulation which may result from interactions of lncRNAs–miRNA and lncRNAs–target miRNAs.

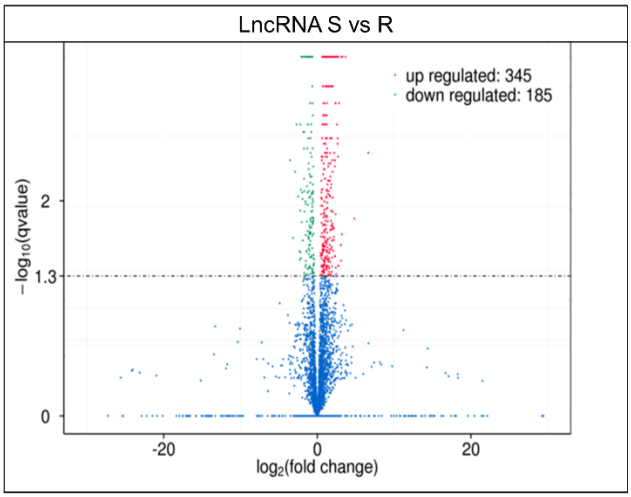
3.5 Conclusion

Using lncRNAseq analysis, we have, for the first time, identified 530 DE lncRNAs in *B. napus* in response to *Plasmodiophora brassicae*. By comparing the response of the lncRNAs in the R and S plants, we have suggested a potentially important role in mediating resistance to *P. brassicae*. qRT-PCR analyses confirmed the expression trend of the lincRNAs and NATs with that of lncRNA-seq. Of the 530 lncRNAs, 16 DE lncRNAs were predicted to function as eTMs of miRNAs and eight DE lncRNAs as the precursor of miRNAs. Of the 24 DE lncRNAs from chromosome BnaA08, eight showed low expression only in the R plants. Pathway analyses showed that the majority of the target genes regulated by these lncRNAs belong to the pathways involved in plant–pathogen interaction, plant hormone signaling, and primary-secondary metabolism. In this regard, the lncRNAs and their targets from BnaA08 provide valuable information because this chromosome is associated with resistance to multiple pathotypes. Moreover, the knowledge of the *P. brassicae*-responsive lncRNAs gained from this research can be used for knock-down or overexpression studies to understand their molecular mechanisms and regulatory functions and may be useful in the development of markers in a breeding program for CR resistance.

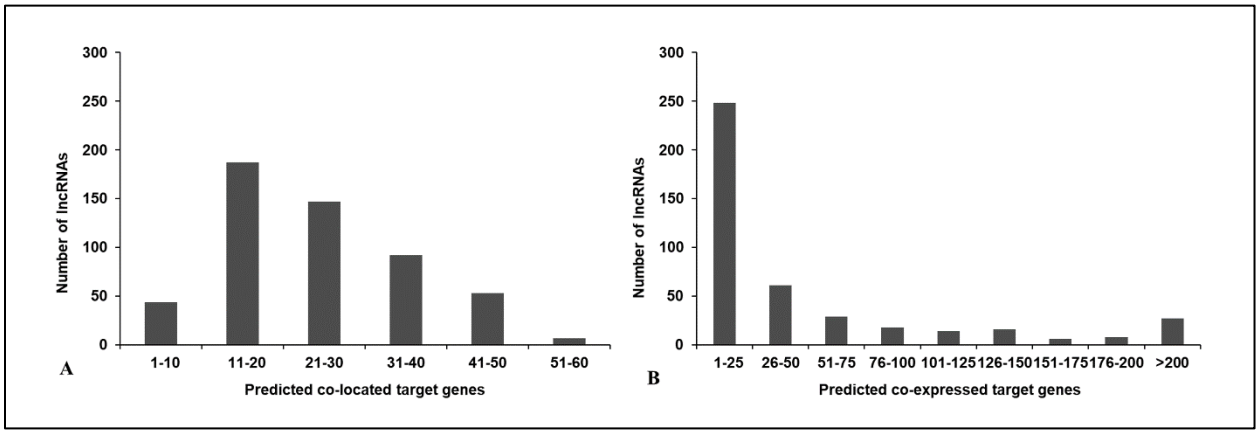
Supplementary Figures



Supplementary Fig. 3.1 Distribution and characterization of features of lncRNAs in *B. napus* in response to *P. brassicae*. A) Length distribution of DE 530 lncRNAs; B) Size distribution of exons; C) Number of exons per transcript for lncRNA expressed; D) Distribution of DE lncRNAs across all chromosomes.



Supplementary Fig. 3.2 Volcanic plot showing the differential expression of *B. napus* lncRNAs post inoculation with *P. brassicae.* Differentially expressed transcripts or genes are visualized by volcano plot where the threshold q-value <0.05 indicates significant differences in expression.



Supplementary Fig. 3.3 LncRNAs and their predicted targets. A) Number of predicted co-located targets of lncRNA; B) Number of predicted co-expressed targets of lncRNA.

Chapter 4: Identification of lncRNAs in response to infection by *Plasmodiophora brassicae* in *Brassica napus* and development of lncRNA-based SSR markers

4.1 Introduction

Clubroot disease, caused by *Plasmodiophora brassicae* Woronin, results in significant damage to Brassica oilseed and vegetable crops (Howard et al. 2010). Development of this disease starts with the infection of root hairs by *P. brassicae*, followed by the root cortex, ultimately leading to swelling of roots and formation of galls (Kageyama and Asano 2009; McDonald et al. 2014). Galls in the root restrict the uptake of water and nutrients resulting in stunted growth, yellowing of leaves and premature death of the plant (for review, (Dixon 2009)); this can result yield loss of about 25-30% (Dixon 2014; Tewari et al. 2005); however, under extreme infestation, yield loss can be such high that harvesting of the crop cannot be justified (for review, Howard et al. 2010). Different cultural and crop management practices have been investigated; however, cultivating resistant cultivars has been considered to be the most effective strategy for managing this disease (for review, see Peng et al. 2014 and Rahman et al. 2014).

Owing to the occurrence of clubroot disease in canola, breeding for the development of clubroot resistant cultivars had started in Europe few decades ago (Diederichsen and Sacristan 1996; Frauen 1999; Manzanares-Dauleux et al. 2001; Wallenhammar 1996). The German plant breeding company Norddeutsche Pflanzenzucht Hans-Georg Lembke KG (NPZ) developed and commercialized the clubroot resistant winter canola cultivar 'Mendel' in the 1990s (Frauen 1999). Clubroot resistance (CR) in 'Mendel' originated from a resynthesized *B. napus* line (AACC, 2*n* = 38), which was developed by crossing the *Brassica olarecea* (CC, 2*n* = 18) line 'ECD-15' and the *Brassica rapa* (AA, 2*n* = 20) line 'ECD-04' (Diederichsen and Sacristan 1996). Rahman et al. (2011) introgressed the clubroot resistance of 'Mendel' into Canadian spring *B. napus* canola, and this resistance has been used in breeding to develop clubroot resistant cultivars. The major gene controlling this resistance has been mapped to chromosome BnaA03 and molecular markers have been identified for use in breeding (Fredua-Agyeman and Rahman 2016). However, the molecular mechanisms and the candidate genes involved in the control of this resistance are yet to be understood; this knowledge would allow the development of gene-based markers, which are known to be more robust than the linked DNA markers.

Recently, it has been demonstrated that non-coding RNAs (ncRNAs), which are characteristically >200 nucleotides long, play a critical role in the regulation of gene expression in plants including in response to biotic and abiotic stresses (Zhu et al. 2014; Muthusamy et al. 2015; Zhang et al. 2016; Seo et al. 2017; Zhang et al. 2018; for review, see Chauhan et al. 2017). The lncRNAs can affect gene functions in different ways; this includes, deregulation of small ncRNAs like microRNAs (miRNAs), interacting with RNA-binding proteins, modulating mRNA decay, repressing target mRNA translation, by sequestering splicing factors, or through modulation of transcriptional regulators by acting as scaffolding complexes (for review, see Wang et al. 2011; Yoon et al. 2013). Several lncRNAs responsive to biotic stresses have been identified by using the next-generation sequencing technologies, for instance, the lncRNAs responsive to *Sclerotinia sclerotiorum* (Joshi et al. 2016). Recent studies into some of the lncRNAs in response to pathogen infection, such as those in *Arabidopsis* against *Fusarium oxyporum* and *Pseudomonas syringe* (Seo et al. 2017; Zhu et al. 2014), in rice against *Maganaporthe oryzae* (Jain et al. 2017), in tomato against Tomato Yellow Curl Virus (Wang et al. 2015), and recently in banana against *Mycosphaerella eumusae* and *Pratylenchus coffeae* (Muthusamy et al. 2019), have demonstrated that this type of RNAs may be important regulators of gene expression in response to pathogens. To date, only the lncRNA ELF18-INDUCED LONG-NONCODING RNA1 (ELENA1) involved in response to biotic stress has been functionally characterized (Seo et al. 2017). Overexpression of this lncRNA in *Arabidopsis* resulted in an increased expression of the genes involved in defense response pathway leading to a disease-resistant phenotype. These observations suggest that at least some of the lncRNAs play a critical role in the regulation of gene expression in response to biotic stress in plants.

The development of clubroot resistant cultivars has become an important goal for breeding programs due to the rapid spread of this disease. Several studies have demonstrated that single-gene controlled resistance can be eroded quickly (Hirai et al. 2004; Strelkov et al. 2016); therefore, strategies including the pyramiding of different CR genes in canola are needed to increase the durability of resistance in this crop. Molecular markers associated with the CR genes are essential for pyramiding of multiple CR genes through marker-assisted selection (MAS). The molecular markers of Mendel-resistance located on BnaA03 which reported by Fredua-Agyeman and Rahman (2016) showed about 2% recombination; this is an

impediment for efficient use of these markers in breeding including gene pyramiding. Recently, ncRNAs including miRNAs have been used for the development of PCR-based markers including simple sequence repeat (SSR) markers (Chen et al. 2010; Fu et al. 2013; Mondal and Ganie 2014); this type of markers can be found within the lncRNAs (Jaiswal et al. 2019; Sahu et al. 2018). SSR markers are one of the most reliable genetic markers due to their several favourable attributes, including reproducibility, multi-allelic and co-dominant nature, and abundant occurrence in the genome (Kalia et al. 2011).

The objective of this research was to identify the lncRNAs responsive to infection by *P. brassicae* in *B. napus* canola plants carrying resistance on BnaA03 derived from the cv. 'Mendel' and to develop molecular markers from the lncRNAs. These lncRNAs were also compared with the differentially expressed (DE) lncRNAs identified in clubroot resistant canola lines carrying resistance on BnaA08 (Hasan and Rahman 2016). To the best of our knowledge, this is the first study on the identification of DE lncRNAs by comparing more than one resistance source and the development of lncRNA-based molecular markers for use in marker-assisted breeding.

4.2 Material and Methods

4.2.1 Plant material and microscopic studies

Two doubled haploid (DH) spring *B. napus* canola populations, developed from F_1's of 1CA0591.323 × A07-26NR and 1CA0591.263 × A07-26NR, were used in this study; these two populations were designated as population Popl#1330 and Popl#1333, respectively. The common parent A07-26NR is susceptible to this disease, while the parents 1CA0591.263 and 1CA0591.323 are clubroot resistant spring canola lines, developed from a cross involving the winter canola cv. 'Mendel' and a Canadian spring canola line (Rahman et al. 2011); clubroot resistance in these lines is located on chromosome BnaA03. The details of the development of these DH lines and their phenotype for resistance to *P. brassicae* pathotype 3 have been reported previously (Fredua-Agyeman and Rahman 2016).

To identify the initial stage of infection of canola roots by *P. brassicae* for lncRNA-seq analysis, clubroot resistant and susceptible DH lines from popl#1333 were grown in a greenhouse at 22/15°C day/night and 16h photoperiod (Summanwar et al. 2019) and the roots were examined microscopically after inoculation with the pathogen. Seedlings, at 10 days after

germination, were inoculated with 1 ml spore suspension of a *P. brassicae* single spore isolate of pathotype 3 (SACAN-ss1; 1×10^7 spores/ml) using the pipette method (Voorips and Visser 1993; Summanwar et al. 2019). The root tissue was collected at 6-,10-,14-,18- and 22-days post inoculation (dpi) and fixed as described in the literature (Verma et al. 2014; Yeung and Saxena 2005). The tissues were sectioned, stained, examined under a microscope (Zeiss Axioscope, Germany), photographed and analyzed (Zeiss Axiocam and AxioVisionTM software, Germany) as described previously (Verma et al. 2014; Summanwar et al. 2019).

For lncRNA-seq analysis, a total of 192 seedlings were grown in a greenhouse as mentioned above; this included eight seedlings of each of 12 resistant (R) and 12 susceptible (S) lines ($24 \times 8 = 192$) of the DH population #1333. Seedlings were inoculated with *P. brassicae* pathotype 3, as mentioned above, and root tissues were collected for lncRNA-seq analysis. The experiment was repeated three times and each experiment was considered as a biological replicate.

4.2.2 RNA isolation, library preparation and sequencing

Microscopy analysis revealed that the initial stage of infection by *P. brassicae* occurred at 10 dpi; therefore, roots of three seedlings from each of the 12 R and 12 S lines were collected at this stage for each biological replicate. Roots were washed with water and the root samples from three plants of each R and S line were harvested as bulks. All samples were flash frozen with liquid nitrogen and stored at -80°C until further use. The remaining five of the eight plants of the R and S lines were grown till maturity and phenotyped for resistance to pathotype 3 for confirmation of the previously reported phenotypic data (Fredua-Agyeman and Rahman 2016).

Total RNA from the R and S bulks was extracted using TRIzol reagent (Invitrogen, USA) and Qiagen RNeasy Plant Mini Kit (Qiagen, Germany) following manufacturer's instructions, as described previously (Summanwar et al. 2019). The integrity of the total RNA was determined using RNA screen Tape assay on Agilent Tape station (Agilent Technologies, USA) following manufacturer's instruction. RNA samples with RNA integrity number (RIN) >9 were selected for further analysis. Six libraries were prepared, which included three biological replicates of R (each sample with bulk of 12 R lines) and three S (each sample with bulk of 12 S lines); library preparation and sequencing were carried out by Novogene (https://en.novogene.com/).

4.2.3 Mapping of reads to reference genome and transcriptome assembly

The raw data from lncRNA-seq analysis was processed as described in Summanwar et al. (2019). Briefly, Novogene's in-house Perl scripts were implemented to remove the reads containing adaptors, unrecognized bases, and low-quality reads. Based on the clean reads, percentage of total bases with Phred score greater than 20 (Q20) and 30 (Q30) were calculated, which indicated the base call accuracy and % GC content in the clean data. To align the clean reads to the *B. napus* genome, the genome annotation files were downloaded from *B. napus* reference genome v5.0 (Chalhoub et al. 2014), http://www.genoscope.cns.fr/brassicanapus/data/) and indexed using Bowtie v2.0.6 (Langmead and Salzberg 2012). The clean pair-end reads were aligned to this refrence genome using TopHat v2.0.9 (Trapnell et al. 2012) and the mapped reads of each sample were assembled by implementing the statistical segmentation model Scripture, beta2 (Guttman et al. 2010) and Cufflinks v2.2.1(Trapnell et al. 2010). Cufflinks was also employed to simultaneously assemble and quantify the expression levels of the lncRNAs to provide maximum likelihood explanation at a specific locus. Default parameters were used to run the program Scripture, while Cufflinks was run with 'min-frags-per-transfrag=0' and '--library-type' and other parameters were set to default. Sorting and removal of the duplicated reads and merging of the Bam alignment file for each sample was accomplished by Picard- tools v1.41 and Samtools v.0.1.18, as described in Summanwar et al. (2019).

4.2.4 LncRNA detection and target prediction

The cufflink assemblies generated from different samples were merged using the software program Cuffmerge (Trapnell et al. 2010) to eliminate the artifacts and replicates. To remove the other known ncRNAs, such as rRNA, tRNA, snRNA, snoRNA, pre-miRNA and pseudogenes from the remaining transcripts, the software program Cuffcompare (Trapnell et al. 2010) was used with exon number filter of exon>=2 and transcript length filter of length >200. The transcript sequences were subsequently classified into lincRNAs, anti-sense lncRNAs and intronic lncRNAs. Four different software tools including CNCI v2 (Coding-Non-Coding-Index) (Sun et al. 2013), CPC 0.9-r2 (Coding Potential Calculator) (Kong et al. 2007), Pfam scan v1.3 (Punta et al. 2012) and PhyloCSF (Phylogenic Codon Substitution Frequency) (Roux et al. 2011) were used to filter the coding potential transcripts. Only the transcripts without the potential for coding proteins were considered as the candidate set of lncRNAs. The FPKMs (fragments per kilobase of transcript per million mapped reads) were calculated for each sample using the

software program Cuffdiff v2.1.1 (Trapnell et al. 2010), and the fold change values for the DE lncRNAs were calculated as log2(fold change) = log2(FPKM susceptible / FPKM resistant). For each DE lncRNA, a *p*-value was assigned which was adjusted to control the false discovery rate using the approach proposed by Benjamin and Hochberg (1995) . The DE lncRNAs with absolute value of log2(fold change) > 1 with $q < 0.05$ were considered significant.

To predict if these DE lncRNAs were modulating the target genes via cis regulation, genes located at 100k bp upstream and downstream of the lncRNAs were determined. To predict if these DE lncRNAs are involved in trans regulation of the genes, Pearson correlation coefficient (r) between the expression of the lncRNAs and the coding genes was calculated; the genes with $r > 0.95$ were considered for further functional enrichment analysis.

4.2.5 Functional classification of lncRNAs

To determine if the lncRNAs were precursors of known *B. napus* miRNAs, BLASTn analysis (e-value=1e-5) was performed as described by Zhang et al. (2018). Briefly, stem loop sequences of all known *B. napus* miRNAs were downloaded from the miRbase database (http:/www.miRNAbase.org/) and were aligned with the DE lncRNAs sequences. The presence of hairpin loop structures in the lncRNAs, a characteristic feature of the precursor of miRNAs, was verified using miRNAFold (https://evryrna.ibisc.univ-evry.fr/miRNAFold; (Tav et al. 2016). Vienna RNAfold web server (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi; Gruber et al. 2008) and Forna (http://rna.tbi.univie.ac.at/forna/; Kerpedjiev et al. 2015) were used to visualize the secondary structures of the DE lncRNAs. To predict if the lncRNAs were mimicking the endogenous miRNA targets, psRNA-Target (https:/plantgrn.noble.org/psRNATarget/V2.2017; Dai et al. 2018) analysis was perfomed with maximum expectation parameters set to 3.5 and target accessibility set to 25.0 (Summanwar et al. 2019).

4.2.6 Identification and validation of SSRs in the DE lncRNAs

The web-based microsatellite finder tool MISA-web (https://webblast.ipk-gatersleben.de/misa/; Beier et al. 2017) was used with default parameters for identification of di-, tri-, tetra-, penta- and hexa-nucleotide SSRs within the DE lncRNAs. To investigate the association of the DE lncRNA-based SSR markers with resistance to *P. brassicae* pathotype 3, 18 primer pairs designed from the DE lncRNAs of the chromosome BnaA03 were tested on the

above-mentioned parents and the two DH populations #1330 and #1333 (Fredua-Agyeman and Rahman 2016). For this, genomic DNA of the resistant (1CA0591.263 and 1CA0591.323) and susceptible (A07-26NR) parents and two bulks of five resistant and five susceptible lines from each of the two DH populations (#1330 and #1333) were screened to identify the markers polymorphic between the parents and the bulks. The polymorphic markers were used to genotype the whole DH population.

4.2.7 Quantitative real time polymerase chain reaction (qRT-PCR) analysis

Expression patterns of the selected DE lncRNAs and their target genes from the *B. napus* genome were validated using qRT-PCR. The selected DE lncRNAs included those from the chromosome BnaA03, the DE lncRNAs regulating the targets on BnaA03 and elsewhere in the genome, as well as the DE lncRNAs known to play a role in plant hormone signalling. The DE lncRNAs identified from this study were also compared with the DE lncRNAs identified previously by using clubroot resistant canola lines carrying resistance of rutabaga (*B. napus* var. *napobrassica*) located on chromosome BnaA08 (Summanwar et al. 2019). We identified 12 lncRNAs which were upregulated in both R lines of Mendel-resistance and rutabaga-resistance; among these, 10 lncRNAs found to regulate plant defense-related genes. To further validate the expression of these common DE lncRNAs and some of their putative targets using qRT-PCR, we selected four of the 10 lncRNAs and their target genes known to be involved in plant defense. Sequences of the primers used in qRT-PCR are listed in Supplementary Table 4.1. All qRT-PCR analyses were performed on QuantStudio 6 Flex Real-Time PCR system (Thermo Fisher Scientific, USA) using PowerUP SYBR Green master mix from Applied Biosystem (Thermo Fisher Scientific, USA). For qRT-PCR analyses, RNAs isolated from the roots of the inoculated R and S lines were treated with DNase I prior to reverse transcription using random hexamers and MirX miRNA first strand synthesis kit (Clontech Laboratories, USA). For each sample, three biological replicates and two technical replicates from each of the biological replicate were used. A constitutively expressed UBC9 (ubiquitin conjugating enzyme 9) gene was used as an endogenous control. Fold change was calculated using the formula:
$2^{-\Delta\Delta Ct}$= ((Ct Gene of Interest – Ct endogenous control) R) – ((Ct Gene of Interest – Ct endogenous control) S) (Schmittgen and Livak 2008).

4.3 Results

4.3.1 Microscopic examination of *P. brassicae* infection

Histopathological examination of the roots of the R and S plants infected with *P. brassicae* showed no evidence of infection at 0 hpi and 6 dpi (Fig. 4.1, Supplementary Table 4.2). However, infection was visible in the secondary roots at 10 dpi in both S (Fig. 4.1C) and R (Fig. 4.1D) plants, where about 4% of the cells exhibited symptoms of infection in the S plants, as compared to only 2% cells in the R plants (Supplementary Table 4.2). However, gall formation was not evident at this stage in both S and R plants (data not shown). At 14 dpi (Fig. 4.1E), about 21% of the cells were infected in the S plants as compared to 6% in the R plants (Supplementary Table 4.2); also, a greater number of secondary plasmodia were observed in the S plants as compared to the R plants at this stage (Fig. 4.1F). From this stage, disease progression continued in the S plants, but not in the R plants, and at 22 dpi (Fig. 4.1G and 1H), more than 75% of the cortical cells of the S plants showed the presence of resting spores (Supplementary Table 4.2); while, only 8% cells of the R plants contained secondary plasmodia and resting spores. At this stage, roots of the S plants exhibited symptoms of hypertrophy and gall formation; however, these could not be detected in the roots of the R plants (data not shown). Thus, the results indicated that infection by *P. brassicae* apparently progresses similarly in both S and R plants until 10 dpi; however, after that, infection does not progress further in the R plants. Based on this observation, we focused on the root samples of the S and R plants collected at 10 dpi for lncRNA profiling.

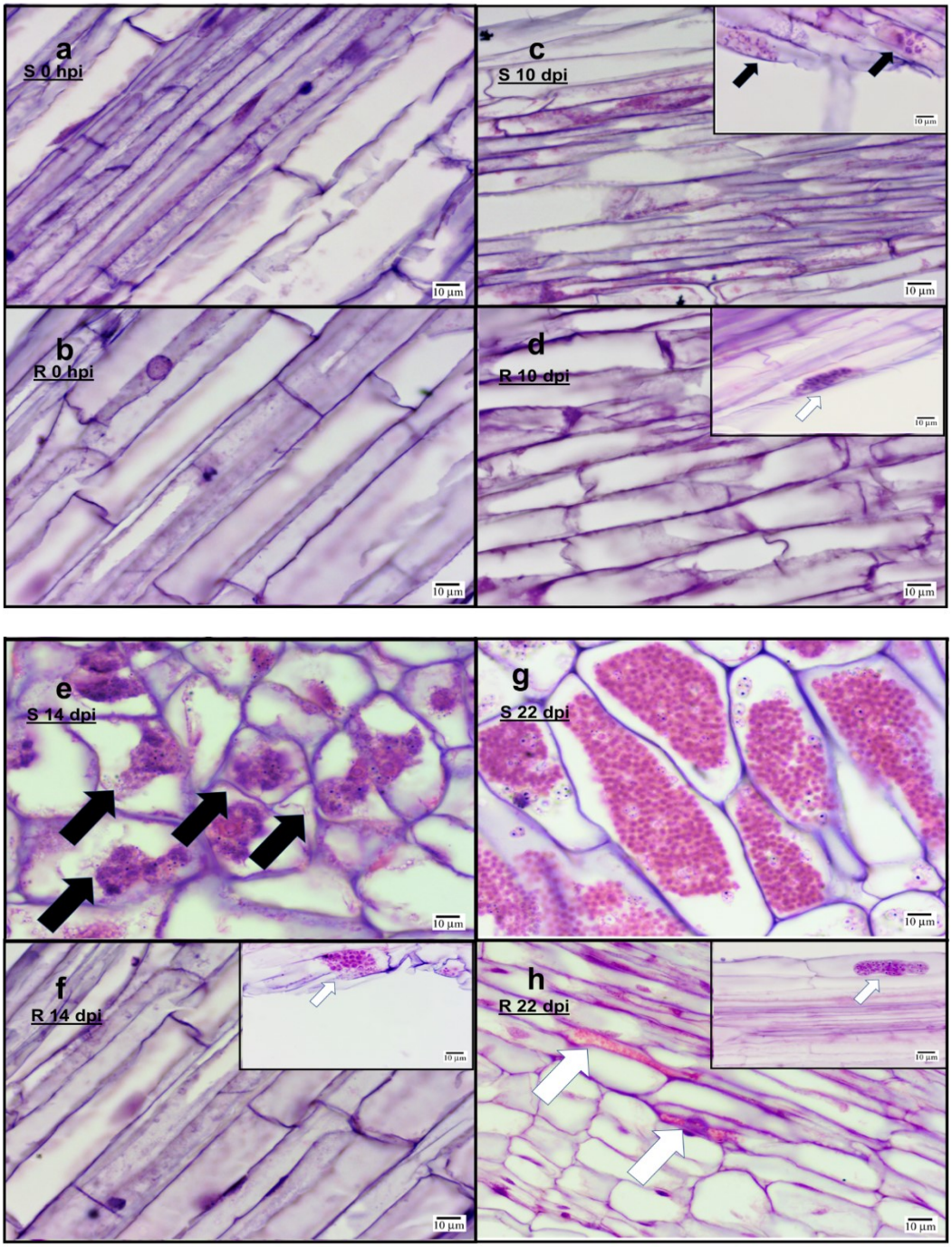


Fig. 4.1 Histopathological analysis of roots of the susceptible (S) and resistant (R) *Brassica napus* plants at 0 hour post-inoculation (hpi) and at 10, 14 and 22 days post-inoculation (dpi) by *Plasmodiophora brassicae* pathotype 3; A) and B) showing the roots of the S and R plants at 0 hpi; C) and D) (inset figure shows larger magnification) showing the presence of proliferating zoosporangia in both S and R plants at 10 dpi, E) and F) showing the secondary plasmodia in cortical cells of the S (black arrows) plants and secondary plasmodia in few cortical cells of the

R (white arrow) plants at 14 dpi, and G) and H) showing many resting spores in the S plants and resting spores and secondary plasmodia in a very few cortical cells of the R (white arrows) plants at 22 dpi. Bars represent the resolution of the microscopy for the particular section. Note: Clubroot resistance in the R plants originated from the winter canola cv. Mendel.

4.3.2 Identification and characterization of the lncRNAs following *P. brassicae* infection

A total of about 445 million paired-end clean reads were obtained from six strand-specific libraries of which 356 million (79.9%) reads could be mapped to the *B. napus* reference genome using Tophat (v2.0.9) (Supplementary Table 4.3). Through further processing of these reads, we identified a total of 4,575 putative DE lncRNA transcripts, and after screening these transcripts based on q-value (<0.05), expression of 464 DE lncRNAs were deemed to be statistically significant; this included 49 (10.6%) antisense lncRNAs and 415 (89.4%) lincRNAs.

All 464 DE lncRNAs were analyzed for the number of exons and their size. A total of 794 exons were detected of which 80.8% (642/794) were 18-1000 nt long (Supplementary Fig. 4.1A). About 59% of the lncRNAs (273/464) contained one exon, 25.9% (120/464) contained two exons, and about 1% (5/464) contained >5 exons (Supplementary Fig. 4.1B). Length of these DE lncRNA transcripts ranged from 203 to 8,238 nt, where 53% (248/464) had a length of 200 to 1000 nt (Supplementary Fig. 4.1C). Of the 464 DE lncRNAs, 165 and 205 lncRNAs were located on 10 chromosomes of the A genome and nine chromosomes of the C genome, respectively; however, 94 lncRNAs could not be assigned to any specific chromosome (Supplementary Fig. 4.1D). The greatest number of the DE lncRNAs was detected on chromosome BnaA03 (35 lncRNAs) and BnaC06 (35 lncRNAs).

Of the 464 DE lncRNAs, 187 were upregulated and 277 downregulated in the S plants (Supplementary Fig. 4.2) as compared to the R plants. Among these, 37 were expressed only in the S plants while 3 were expressed only in the R plants (Table 4.1). The identity of the 20 lncRNAs which exhibited the greatest degree of differential expression are listed in Supplementary Table 4.4; among these, three upregulated and two downregulated lncRNAs of the R plants were found located on chromosome BnaA03 where the clubroot resistance genes has been mapped (Fredua-Agyeman and Rahman 2016).

Screening of up- and down-stream genomic regions (100 kb) of the 464 DE lncRNAs identified a total of 11,779 target genes to be cis-regulated by these lncRNAs; among these, 5,946 located at the upstream and 5,833 at the downstream of the lncRNAs (Supplementary Fig. 4.3A; Supplementary Table 4.5). Using KEGG pathway enrichment analysis, a total of 4,352 co-

located target genes were annotated. Among the 1,147 genes from the top 20 biosynthetic pathways identified through KEGG analysis, about 39% (452/1,147) of the genes belonged to the pathways involved in the biosynthesis of secondary metabolites and about 7% (84/1,147) in plant-pathogen interactions (Fig 2a). Based on co-expression analysis between the lncRNAs and the mRNAs, 391 of the 464 lncRNAs were predicted to have co-expressed target genes (Supplementary Fig. 4.3B). Based on KEGG pathway enrichment analysis of the co-expressed genes from the top 20 pathways, 26% (305/1166) of the co-expressed target genes belong to the primary metabolic pathways, and this included the genes involved in biosynthesis of amino acids (13%; 157/1166) and carbon metabolism (13%; 148/1166); some of the target genes were also involved in oxidative phosphorylation (8%; 91/1166) and plant-pathogen interaction (8%; 89/1166) (Fig. 4.2B). Thus, the results from this study identified several DE lncRNAs which seem to play a role in the regulation of the genes (targets) involved in various pathogen-responsive metabolic pathways.

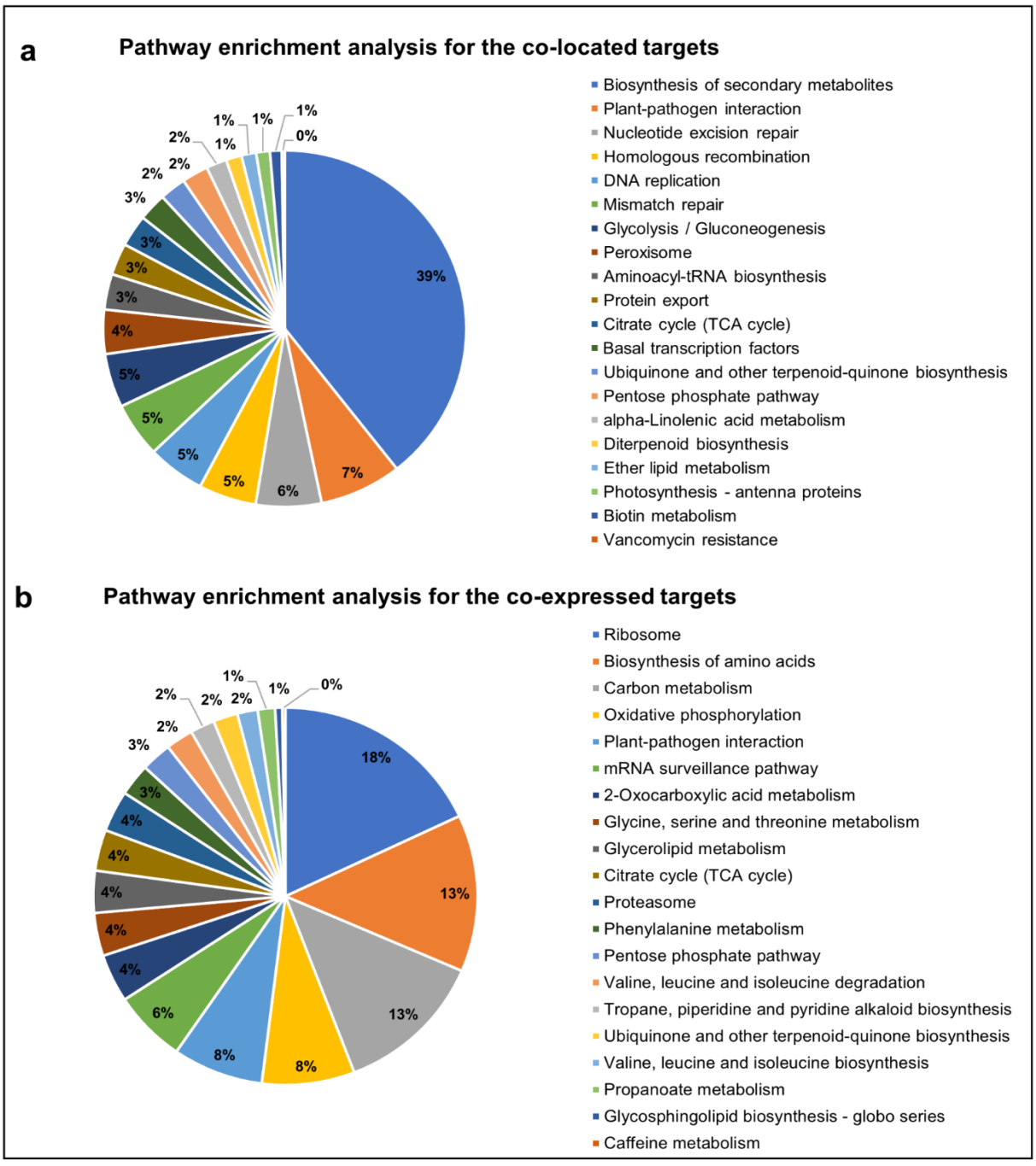


Fig. 4.2 Functional annotation of the predicted target genes of *Brassica napus* lncRNAs based on Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis. A) pathway enrichment analysis of the co-located target genes, and B) pathway enrichment analysis of the co-expressed target genes. [Statistical significance of the enrichment, *p*-value (<0.05) calculated using KOBAS v.20]

4.3.3 DE lncRNAs located on *B. napus* chromosome BnaA03

Of the 35 DE lncRNAs located on chromosome BnaA03, 19 were upregulated in the S plants and 16 in the R plants; among these, 11 expressed only in S plants with no expression in R plants (Table 4.1, Fig. 4.3). Predicted targets of these 35 lncRNAs are mostly involved in plant-pathogen interactions, primary metabolism and phenylpropanoid biosynthesis (Supplementary Table 4.6), and some of the lncRNAs had up to 145 target genes. For instance, the LNC_000396 (18 target genes), LNC_000397 (14 target genes) and LNC_000320 (34 target genes), which were highly expressed in the R plants found to be regulating caffeic acid 3-O-methyltransferase (*COMT1*), beta-glucosidase 3 (*BGL03*) and phenylalanine ammonia-lyase 4 (*PAL4*) genes involved in the phenylpropanoid biosynthetic pathway. In case of the target genes involved in carbon and amino acid metabolic pathways, 32 of the 35 lncRNAs were found to be cis-regulating and 25 of them were also trans-regulating this type of target genes (Supplementary Table 4.6). While looking at the target genes involved in disease resistance and in plant-pathogen interactions, 15 and 20 lncRNAs were predicted to trans-regulate these two types of genes, respectively (Supplementary Table 4.7 and 4.8). Thus, the results suggest that these lncRNAs may play a role in mediating resistance to clubroot disease.

Table 4.1: List of the lncRNAs uniquely expressed in clubroot disease resistant (R) or susceptible (S) lines of *Brassica napus*

Transcript_ID	Chromosome name and location (bp)	S (FPKM)	R (FPKM)
LNC_000772	BnaA06:10978686-10979071	0	1.77568
LNC_001922	BnaC01:13461970-13462387	0	0.970165
LNC_002537	BnaC03:15192320-15192823	0	0.9502
LNC_000264	BnaA02:22103587-22104364	3.03393	0
LNC_000332	**BnaA03:16695781-16698971**	**1.81157**	**0**
LNC_000348	**BnaA03:23711237-23712313**	**2.94814**	**0**
LNC_000354	**BnaA03:26012928-26013836**	**1.65033**	**0**
LNC_000421	**BnaA03:20136387-20138541**	**1.61975**	**0**
LNC_000444	**BnaA03:26259014-26260595**	**0.799901**	**0**
LNC_000445	**BnaA03:26259014-26260595**	**4.21893**	**0**
LNC_000446	**BnaA03:26259014-26260595**	**1.32141**	**0**

LNC_000456	**BnaA03:29547518-29549792**	**0.59793**	**0**
LNC_000457	**BnaA03:29547518-29549792**	**0.748802**	**0**
LNC_000458	**BnaA03:29547518-29549792**	**0.832231**	**0**
LNC_000652	BnaA05:14374128-14374551	2.22373	0
LNC_000748	BnaA06:3974486-3975726	1.99208	0
LNC_000952	BnaA07:2608742-2609009	23.5527	0
LNC_001117	BnaA08:713831-714773	3.07834	0
LNC_001293	BnaA09:29804174-29804391	180.64	0
LNC_001341	BnaA09:11766717-11767907	1.95432	0
LNC_002711	BnaC04:6728563-6728868	3.97035	0
LNC_002872	BnaC04:27296752-27297254	2.51332	0
LNC_002884	BnaC04:30849536-30851460	1.687	0
LNC_002891	BnaC04:34483346-34483939	1.34276	0
LNC_002936	BnaC05:1739372-1740288	0.989078	0
LNC_002959	BnaC05:8922090-8923406	0.814856	0
LNC_003183	BnaC06:12792054-12794831	1.18662	0
LNC_003247	BnaC06:32298723-32299069	2.76907	0
LNC_003568	BnaC07:9813085-9813401	4.77701	0
LNC_003747	BnaC08:27428746-27429208	1.09987	0
LNC_000497	**BnaA03_random:5094311-5095007**	**3.21291**	**0**
LNC_003855	BnaC08:27775749-27776276	8.16475	0
LNC_001622	BnaAnn_random:26978094-26978952	1.8183	0
LNC_001628	BnaAnn_random:28821785-28822721	4.03952	0
LNC_001731	BnaAnn_random:2089455-2092035	2.24845	0
LNC_001770	BnaAnn_random:21864290-21864748	1.71812	0
LNC_003923	BnaC08_random:4426408-4426614	238.726	0
LNC_004251	BnaCnn_random:12255877-12256505	0.962938	0
LNC_004515	BnaCnn_random:19091029-19091841	0.967721	0
LNC_004681	BnaCnn_random:69886140-69886602	1.66703	0

Note: FPKM = fragments per kilobase of transcript per million mapped reads. Rows in bold indicate the lncRNAs located on chromosome BnaA03.

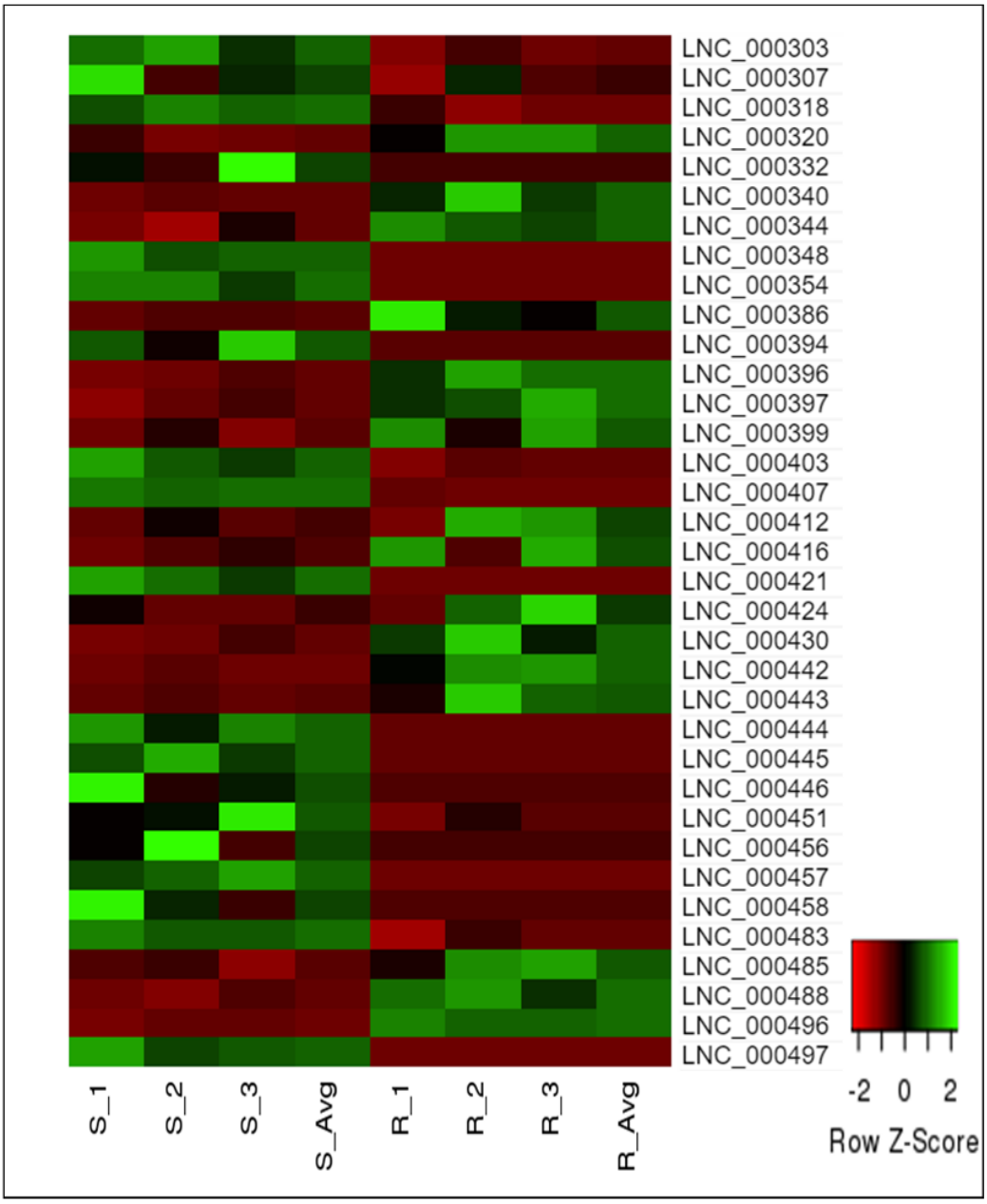


Fig. 4.3 Differentially expressed lncRNAs from chromosome BnaA03 of *Brassica napus*. Expression values (FPKMs) are scaled from minimum (-2) (red) to maximum (+2) (green); low or no expression is depicted by black color. S_1, S_2 and S_3 are the biological replicates of the clubroot susceptible *B. napus*, and R_1, R_2 and R_3 are the biological replicates of the clubroot resistant *B. napus* lines.

4.3.4 qRT-PCR validation of the expression of the lncRNAs

To validate the expression of the lncRNAs and their target genes using qRT-PCR, we selected a total of 10 lncRNAs which included three lncRNAs from chromosome BnaA03, two lncRNAs which have their putative targets on BnaA03, and five random lncRNAs whose target genes are known to be involved in plant defense; all these lncRNAs predicted to regulate a maximum of four target genes. In case of the target genes, a total of 16 genes were selected for qRT-PCR.

Among the 10 lncRNAs selected for qRT-PCR analysis, LNC_000354 and LNC_000421 from BnaA03 were expressed only in the S plants. The probability of their expression being a false positive was eliminated by validating their expression using qRT-PCR (data not shown). The target genes regulated by these two lncRNAs were also validated and their expression patterns were found to be similar in both RNAseq (unpublished data) and qRT-PCR analysis (Fig. 4.4). The remaining eight lncRNAs and their 14 target genes from RNAseq (unpublished data) were observed to follow a similar trend of expression when compared with the results obtained from qRT-PCR (Fig. 4.4). The remaining two targets, BnaCnng45310D and BnaA10g14200D, respectively, of the LNC_001630 and LNC_004419 did not show expression in either R or S plants (data not shown), which could be due to the fact that these target genes were false positives with q value >0.05.

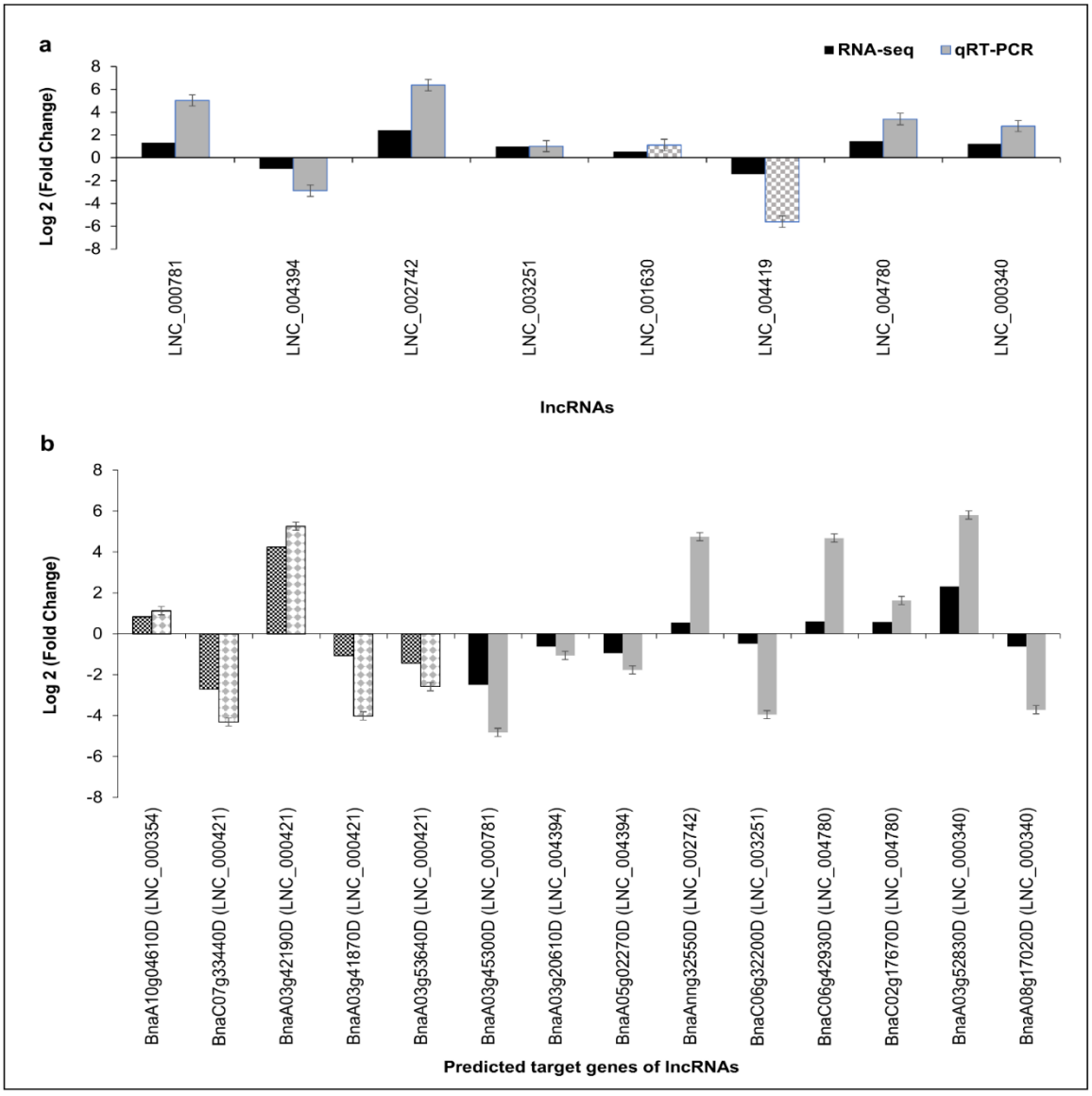


Fig. 4.4 A) qRT-PCR validation of the pathogen induced eight lncRNAs detected by lncRNAseq analysis of *Brassica napus* lines carrying clubroot resistance genes of winter canola cv. Mendel. The patterned bars represent the lncRNAs whose targets were not expressed in the resistant (R) and susceptible (S) lines. B) qRT-PCR validation of RNAseq data for 14 targets predicted by co-expression analysis; the lncRNAs regulating their targets are indicated in parentheses. The patterned bars indicate the targets of the lncRNAs (LNC_000354 and LNC_000421) from chromosome BnaA03, which are expressed only in the susceptible plants. [Expression of the lncRNAs and their target genes in both cases (A and B) was normalized to *UBC9* (endogenous control). [The expression of lncRNAs and target genes was normalized to *UBC9* (endogenous control). Log_2 (fold change) in R lines was calculated using the formula $2^{-\Delta\Delta Ct} = ((Ct_{\text{gene of interest}}$

$-$ Ct $_{\text{endogenous control}}$)R $-$ (Ct $_{\text{gene of interest}}$ $-$ Ct $_{\text{endogenous control}}$)S)]. Three biological replicates and two technical replicates were used. Error bar = standard error].

4.3.5 Comparative analysis of the DE lncRNAs of the lines carrying Mendel- or rutabaga-resistance

In Summanwar et al. (2019), we reported 530 DE lncRNAs, in response to infection by *P. brassicae*, in canola plants carrying resistance introgressed from rutabaga located on chromosome BnaA08 (Hasan and Rahman 2016). We compared these 530 lncRNAs with the 464 DE lncRNAs from this study and identified 123 (23-26%) common DE lncRNAs in these two types of resistant lines. Among the 123 DE lncRNAs, 12 (LNC_001220, LNC_000496, LNC_000801, LNC_000980, LNC_002687, LNC_002381, LNC_003929, LNC_004310, LNC_003624, LNC_003305, LNC_004780 and LNC_001968) showed a similar trend of upregulation with fold change <-1 in both lines carrying Mendel- or rutabaga-resistance; these lncRNAs might play an important role in resistance to clubroot disease (Supplementary Table 4.9). Among these, 10 lncRNAs (excluding LNC_002687 and LNC_004310) were predicted to regulate the genes involved in different defense-related pathways (Fig. 4.5A), while the two, LNC_002687 and LNC_004310, are predicted to regulate only the neighbouring genes but expression of these genes could not be detected at 10 dpi. qRT-PCR analysis further confirmed upregulation of these lncRNAs and their target genes in the lines carrying Mendel- or rutabaga-resistance in response to infection by *P. brassicae* at 10 dpi (Fig. 4.5B).

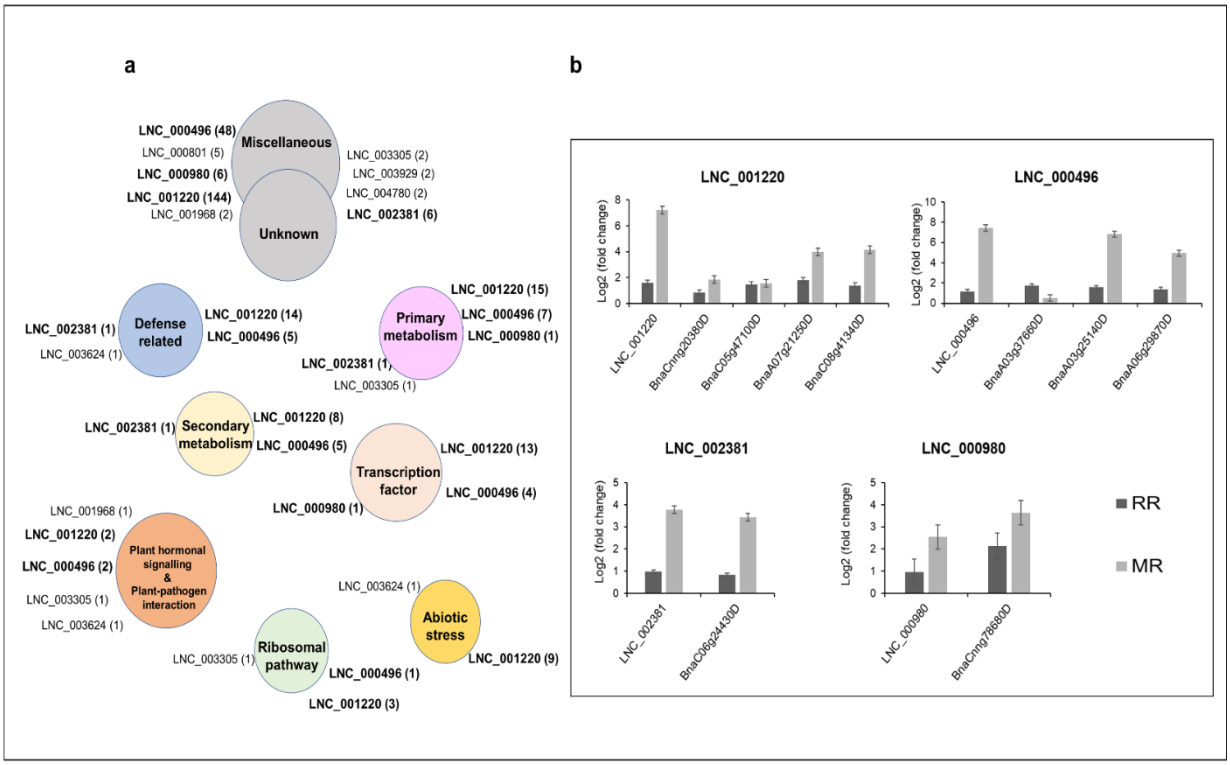


Fig. 4.5 A) List of 10 lncRNAs upregulated in two types of clubroot resistant *Brassica napus* lines, one carrying resistance located on chromosome BnaA03 which was introgressed from winter canola cv. 'Mendel' and the other on BnaA08 which was introgressed from rutabaga, and the different plant defense-related pathways to which the target genes of these lncRNAs belongs; majority of the lncRNAs regulate target genes involved in more than one pathways; number in brackets adjacent to the lncRNAs indicate the total number of genes under the regulation of the lncRNA. Bold font indicates the lncRNAs selected for validation by qRT-PCR. B**)** qRT-PCR validation; the individual graphs display the differentially expressed (DE) lncRNAs and their target genes which are upregulated. [Expression of the DE lncRNAs and their target genes was normalized to *UBC9* (endogenous control). Log_2 (fold change) in R lines was calculated using the formula $2^{-\Delta\Delta Ct} = ((Ct_{\text{gene of interest}} - Ct_{\text{endogenous control}})R - (Ct_{\text{gene of interest}} - Ct_{\text{endogenous control}})S)]$. Three biological replicates and two technical replicates were used. Error bar = standard error]. RR: resistant *B. napus* lines carrying resistance introgressed from rutabaga; MR: resistant *B. napus* lines carrying resistance introgressed from winter canola cv. 'Mendel'.]

4.3.6 *B. napus* lncRNAs predicted to function as precursors or endogenous target mimics (eTM) of microRNAs (miRNAs)

Analysis of the 464 DE lncRNAs (based on Mendel-resistance) using psRNAtarget (Dai et al. 2018) indicated that 10 lncRNAs could possibly serve as target mimics of seven different miRNAs and seven miRNAs belonging to one family. Among these, two lncRNAs, LNC_000146 and LNC_003825, are mimicking the genes targeted by members of the miR156 family (Supplementary Table 4.10). In contrast, a single miR1140 seems to be targeted by two lncRNAs (LNC_000915 and LNC_003019). The above-mentioned 10 lncRNAs could function as eTM and were predicted to sequester the miRNAs and, thus, conserve the mRNAs by disrupting the miRNA-mRNA interactions. We also found that some of the known miRNAs are possibly produced by lncRNAs, such as the miR159 by LNC_003347 and LNC_002274 (Fig. 4.6) and miR168b by LNC_001885 (Supplementary Table 4.11). Thus, it can be anticipated that some of the lncRNAs expressed in response to infection by *P. brassicae* can function as the targets of miRNAs or serve as the precursors of miRNAs.

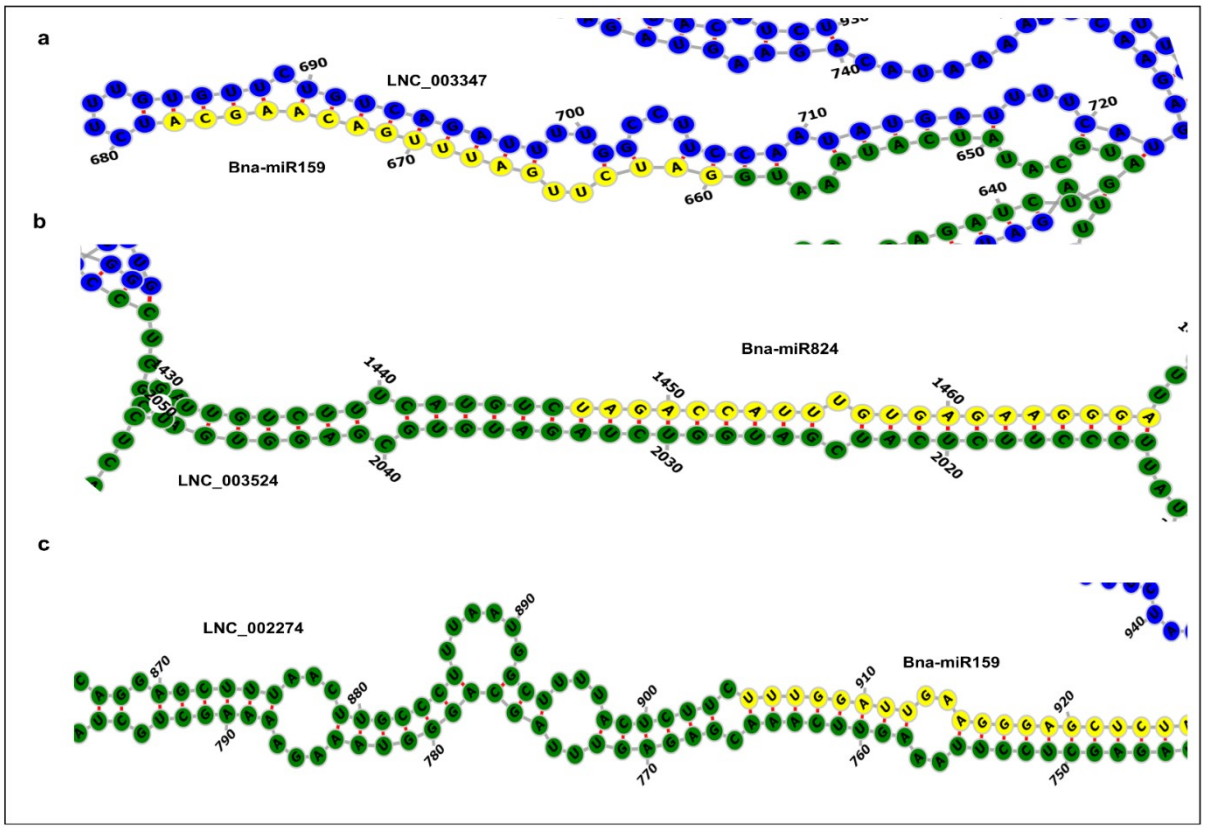


Fig. 4.6 Representative predicted lncRNAs functioning as miRNA precursor sequences. The predicted model displays a possible formation of hair-pin loop in the lncRNA sequence, a characteristic feature of pre-miRNA. A) LNC_003347 is a target mimic of bna-miR159. B) LNC_003524 is a target mimic of bna-miR824. (c) LNC_002274 is a target mimic of bna-miR159. Numbers shows the nucleotide positions of lncRNAs (not the physical position in the genome. Blue: lncRNA sequence; green: miRNA precursor sequence; yellow: miRNA mature sequence.

4.3.7 Identification of Simple Sequence Repeats (SSRs) within the DE lncRNAs

All the 464 DE lncRNA sequences identified by using Mendel-resistance were analyzed for the presence of SSRs. Among these, 196 sequences contained a total of 269 SSR motifs, where 57 sequences contained more than one SSR motif (Fig. 4.7). The proportion of mononucleotide SSRs (70.6 %) was greater than the di-, tri- and tetranucleotide SSRs (Fig 4.8A and B) while no penta- or hexanucleotide SSRs could be detected; 22 were identified as compound SSRs (Fig 8B). Among the dinucleotide repeats, the AT repeats was most frequent

(Fig 8C). Thus, the DE lncRNAs in response to infection by *P. brassicae* were found to contain SSRs which can be used in molecular breeding; the detailed information of the SSRs is presented in Supplementary Table 4.12.

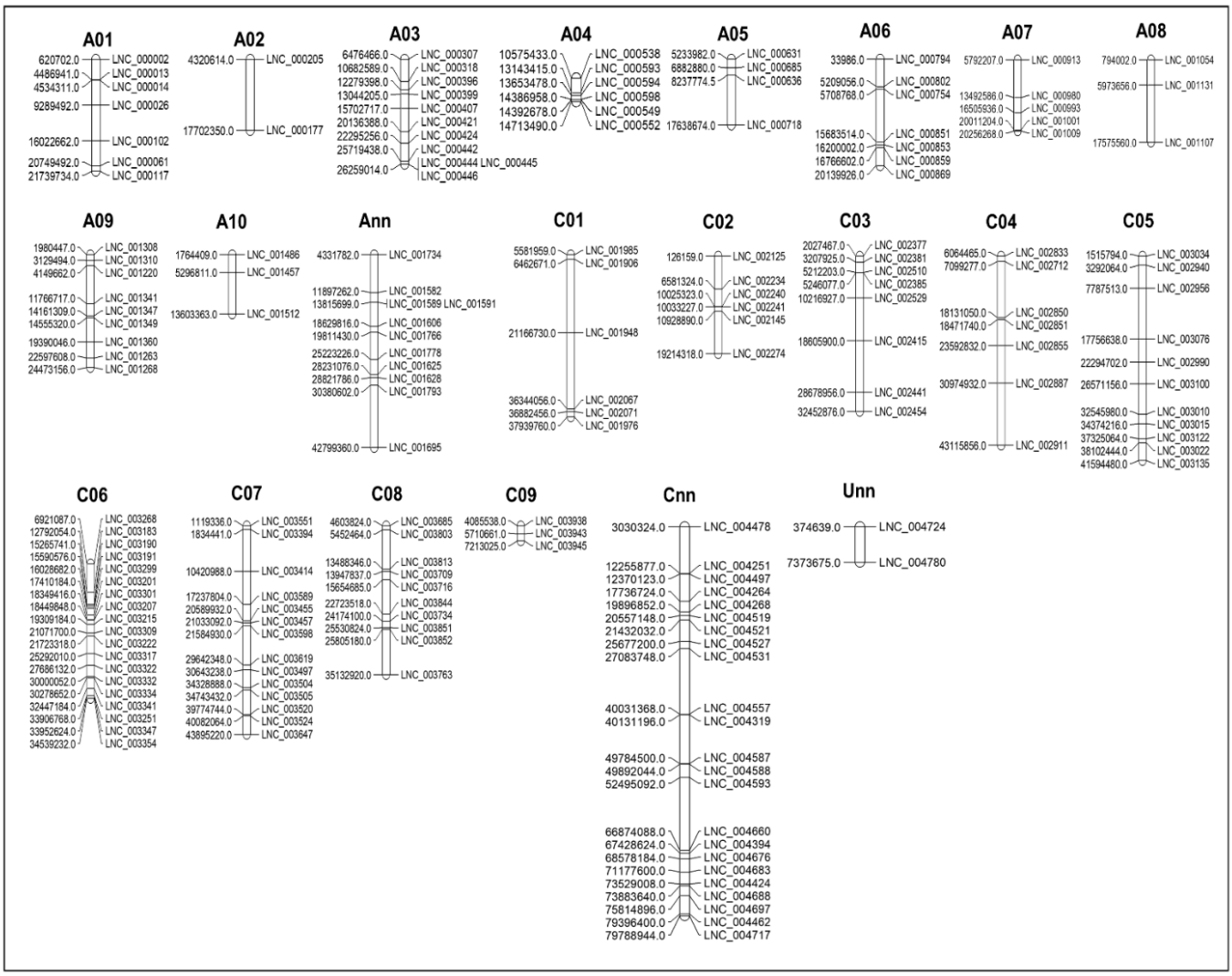


Fig. 4.7 The distribution of the lncRNA based SSRs in 19 chromosomes of *Brassica napus*. The names of the lncRNAs are listed on the right side of the chromosome and their physical position (bp) listed on the left side. The lncRNA of the A genome which could not be positioned on physical map of the *B. napus* genome sequence (Chalhoub et al. 2014) are included in a separate chromosome and marked as 'Ann'; similarly, the C genome lncRNAs are marked as 'Cnn', and those which genomic location is unknown are marked as 'Unn'.

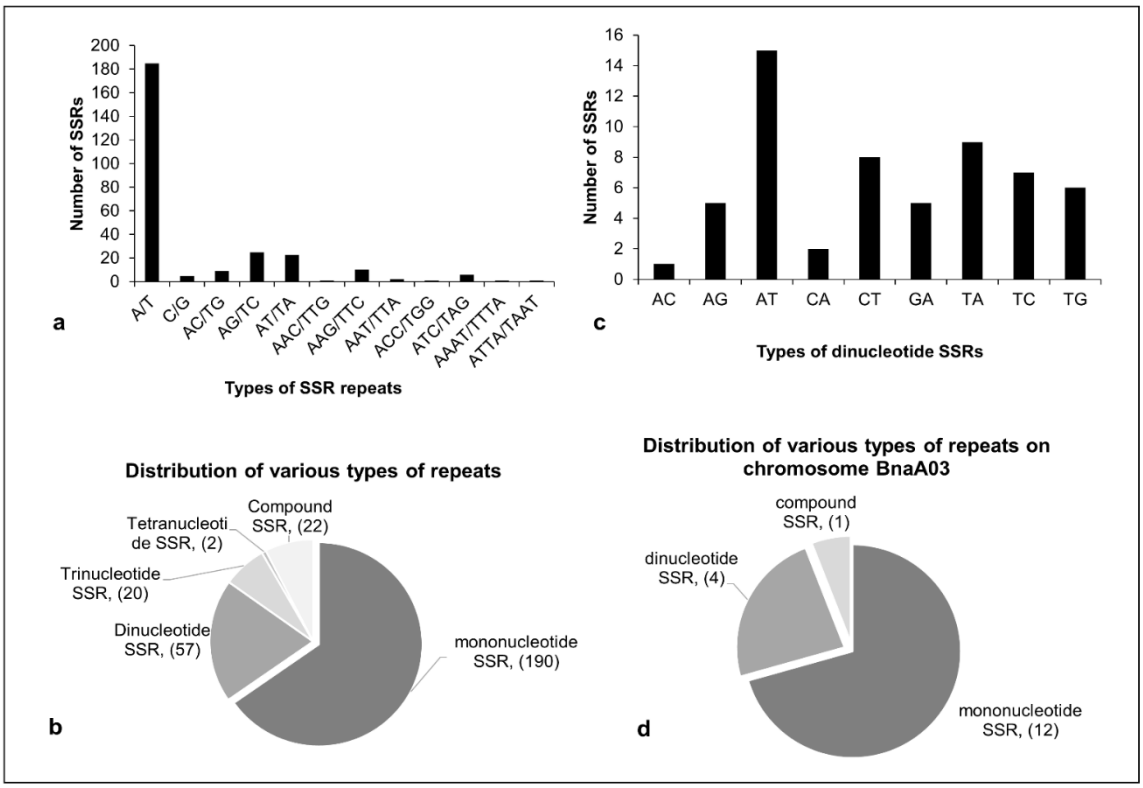


Fig. 4.8 Simple sequence repeats (SSRs) identified within the DE lncRNAs of *Brassica napus*. A) distribution of the SSRs based on the type of the motifs; B) distribution of the SSRs based on the length of the motifs (number of SSRs are shown in brackets); C) types of dinucleotide SSRs and their occurrence; and D) distribution of various types of SSRs on chromosome BnaA03 (number of SSRs are shown in brackets).

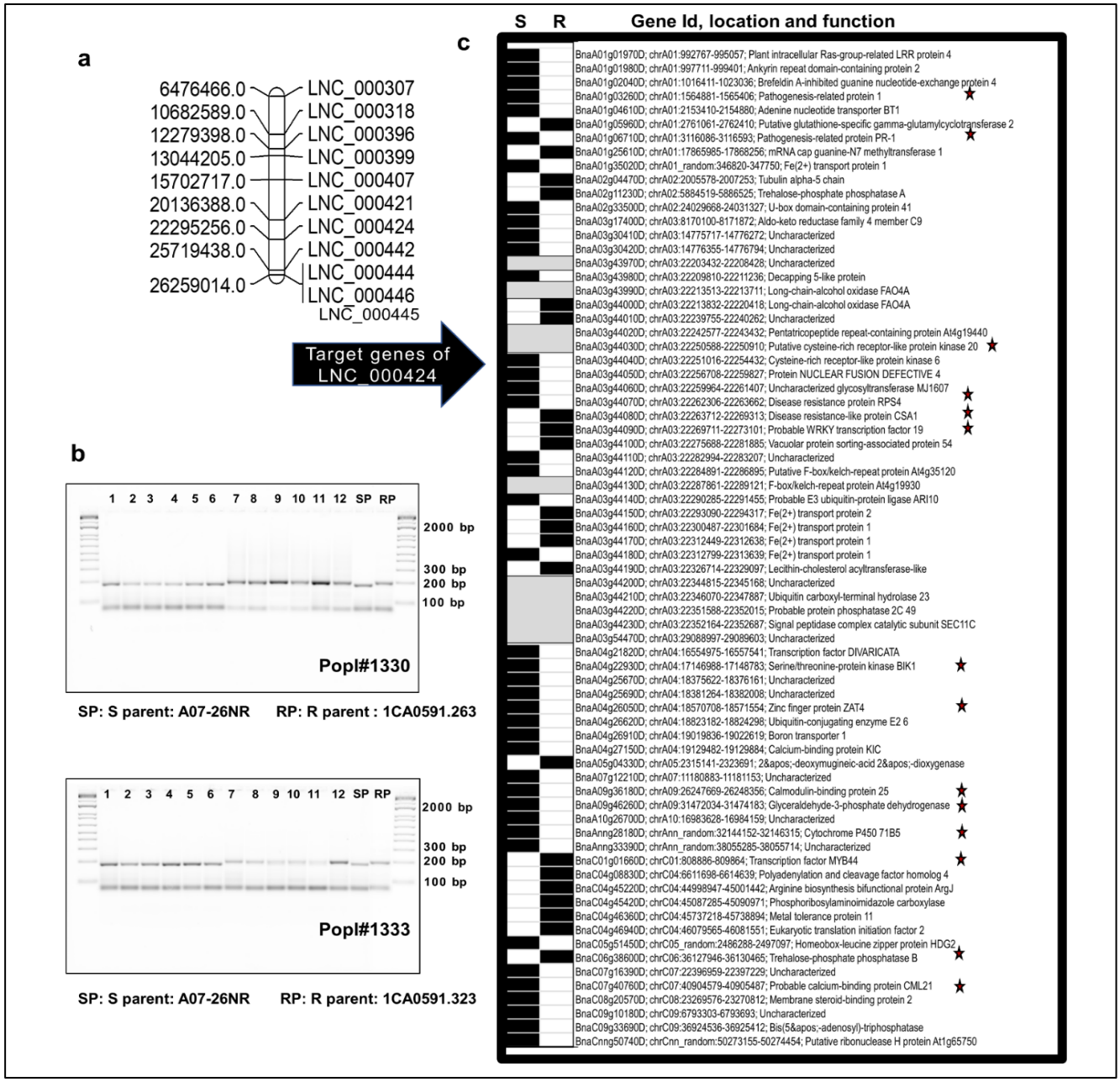


Fig. 4.9 Analysis of simple sequence repeats (SSRs) of LNC_000424. A) The lncRNAs carrying SSRs and their physical position on chromosome BnaA03. B) Gel images depicting partial genotyping of two *Brassica napus* doubled haploid populations (Popl#1330 (Gel lane 1-6: S lines; 7-12: R lines) and Popl#1333(Gel lane 1-6: S lines; 7-12: R lines)) by the SSR marker designed from the LNC_000424. C) Information of the target genes predicted to be regulated by LNC_000424; black and white colour against the S (susceptible lines) and R (resistant lines) indicate low and high expression, respectively, and the grey blocks indicate the target genes

predicted to be cis regulated by the LNC_000424 but showing no expression at 10 dpi; and the stars indicate the target genes involved in plant defense.

4.3.8 Identification and validation of SSRs from the DE lncRNAs of BnaA03

A total of 17 SSRs were identified within 13 lncRNAs located on chromosome BnaA03 (Supplementary Table 4.12); this included 12 mononucleotide, four dinucleotide and one compound SSR (Fig. 4.8D). Of the 13 lncRNAs, the LNC_000318 and LNC_000424 contained two and four SSRs, respectively, while all others contained only one SSR. A total of 18 primer pairs were designed based on seven mononucleotide, four dinucleotide and one compound SSR, and were tested for polymorphism (Supplementary Table 4.12); among these, the marker designed based on the compound SSR ((TA)12(T)12) identified within the LNC_000424 (Supplementary Table 4.11) found to be polymorphic. Genotyping of the 153 DH lines of the two populations, Popl#1330 and Popl#1333, with this marker showed an association with clubroot resistance (Fig 4.9); however, this association could not be established in two lines of the Popl#1330 and one line of the Popl#1333. Thus, this SSR marker can potentially be used for identification of the resistant plants with an accuracy of greater than 98%. Expression of the LNC_000424 was slightly upregulated in two of the three replications as well as average value (fold change of 0.89) of all three replications of the R plants (Fig. 4.3).

4.4 Discussion

Recent studies have demonstrated that lncRNAs are involved in the regulation of various biological processes in plants including growth and development, photo-morphogenesis and responses to biotic and abiotic stresses (Ariel et al. 2014; Ding et al. 2012; Franco-Zorrill et al. 2007; Heo et al. 2009; Swiezewski et al. 2009). In this study, we investigated the possible role of the lncRNAs in mediating resistance to clubroot disease in canola through expression analysis of the lncRNAs in response to *P. brassicae* in plants carrying resistance gene located on chromosome BnaA03, as well as through the development of molecular markers from the lncRNAs and testing the markers in two mapping populations. To identify the lncRNAs possibly playing a fundamental role in resistance to this disease, we compared the DE lncRNAs from this study with the DE lncRNAs that we reported previously (Summanwar et al. 2019) by using canola lines carrying resistance on BnaA08.

To identify the stage of initial infection of the roots of the R and S plants used in this study for RNA sequencing, we conducted histopathological study, and the results are consistent with our earlier observations on canola carrying clubroot resistance gene located on the chromosome BnaA08 (Summanwar et al. 2019). The lncRNA profiling of the root samples collected at 10 dpi identified a total of 464 DE lncRNAs in the roots of the R and S plants, of which 37 were expressed only in the S plants while only three were specific to the R plants. This is in contrast to our observation on the R plants carrying resistance introgressed from rutabaga (Summanwar et al. 2019) where the number of lncRNAs expressed only in the R plants was greater than the number expressed only in the S plants. However, a greater number of the DE lncRNAs showed upregulation in the S plants as compared to the R plants, which is consistent with our previous study (Summanwar et al. 2019). It is, therefore, probable that a greater number of lncRNAs are differentially expressed in the S plants, and this might be due to the disruption of the growth of the S plants by *P. brassicae* infection. Zhang et al. (2018) also reported that a greater number of lncRNAs differentially expressed in susceptible cotton than in the resistant ones in response to infection by *Verticillium dahliae*. In this study, we further explored the lncRNAs regulating the genes involved in PAMP-triggered immunity (PTI) and effector-triggered immunity (ETI), as well as in pathways like phenylpropanoid biosynthetic pathway, and primary and secondary metabolic pathways.

4.4.1 LncRNAs regulating genes involved in plant-pathogen interactions

Both PTI and ETI play an important role in plant immunity (for review, Bigeard et al. 2015), and the pattern recognition receptors (PRRs) are essential for recognizing the molecular signatures associated with pathogens and are known to induce local and systemic immunity (for review, Jones and Dangl 2006). In the present study, we found that several lncRNAs, including LNC_000745, LNC_003846, LNC_001589, LNC_001591, LNC_002078, LNC_001592, LNC_001519, and LNC_000496, were upregulated in the R plants and were predicted to modulate a brassinosteroid insensitive 1-associated receptor kinase 1 (*BAK1*) (BnaA03g52520D) gene. The PRRs, for example flagellin-sensing 2 (*FLS2*), EF-Tu receptor (*EFR*) and *BAK1*, belong to LRR receptor-like serine/threonine-protein kinase family (Chinchilla et al. 2007; Roux et al. 2011; for review, see Jones and Dangl 2006). Our analysis suggests that *FLS2* is co-regulated by eight lncRNAs (LNC_000077, LNC_000994, LNC_002415, LNC_002402, LNC_001085, LNC_003855, LNC_000177 and LNC_003247) and the *EFR* by a single lncRNA

(LNC_001473) detected at 10 dpi in response to infection by *P. brassicae*. Chinchilla et al. (2007) provided evidence that *FLS2* and *BAK1* form a complex on stimulation by flagellin, thus initiating PRR-dependent signalling for innate immunity in *Arabidopsis*. Similarly, Gómez-Gómez and Boller (2000) and Kunze et al. (2004) have demonstrated, respectively, the role of *FLS2* and *EFR* in eliciting innate immunity in this model plant.

In this study, we also detected 13 DE lncRNAs (LNC_002263, LNC_001473, LNC_001382, LNC_000745, LNC_001589, LNC_001591, LNC_002078, LNC_001592, LNC_001519, LNC_000416, LNC_000205, LNC_2891 and LNC_000496) which were predicted to regulate the RPM1-interacting protein 4 (*RIN4*), a regulator of genes encoding disease resistance proteins such as resistance to *Pseudomonas syringae* protein 2 (*RPS4*) and resistance to *P. syringae 3* (*RPM1*) (Supplementary Table 4.13). The NBS-LRR gene family is one of the largest gene families in plant kingdom which is known to play an important role in the cytosolic recognition of pathogenic effectors for the activation of ETI (for review, Noman et al. 2019). The NBS-LRR proteins like *RPM1* and *RPS2* are known to physically interact with *RIN4* to activate the *RPM1*-dependent hypersensitive response (for review, see Belkhadir et al. 2004), where *RIN4* acts as a negative regulator of basal defense responses but functions as a positive regulator of *RPM1*-mediated resistance (Mackey et al. 2002; for review, see Belkhadir et al. 2004). It is known that the reduced expression of *RIN4* may not be compromised by the *RPS2*-mediated resistance (Mackey et al. 2002). Thus, our study suggests the involvement of lncRNAs in the regulation of genes involved in innate immunity of the plants and their importance in PTI and ETI.

4.4.2 LncRNAs regulating genes involved in primary and secondary metabolism

Most of the DE lncRNAs detected in this study were found to regulate target genes involved in primary and secondary metabolic pathways. For instance, the phenylpropanoid pathway is known to be very important for biotic stress responses and is responsible for the biosynthesis of several secondary metabolites that are important during host-pathogen interactions (for review, see Caretto et al. 2015). Genes encoding the enzymes such as phenylalanine ammonia lyase (for review, see Vogt 2010), laccase (Liu et al. 2018; Wang et al. 2015), chalcone synthase (for review, see Dao et al. 2011), cinnamoyl-CoA reductase (Kawasaki et al. 2006), cinnamyl alcohol dehydrogenase (Bagniewska-Zadworna et al. 2014), 4-coumarate-CoA ligase (Liu et al. 2017), caffeoyl shikimate esterase, caffeoyl-CoA *O*-methyltransferase 1

and caffeic acid *O*-methyltransferase (for review,see Xie et al. 2018) play an important role in the phenylpropanoid biosynthetic pathway; some of the DE lncRNAs of BnaA03 identified in our study are predicted to target the genes encoding the aforementioned enzymes (Supplementary Table 4.6). In addition to this, many genes involved in primary metabolic pathways including carbon and amino acid metabolism are also targeted by the DE lncRNAs (Supplementary Table 4.6). Thus, the results from the present study agree well with our previous observations on lncRNAs differentially expressed in canola carrying clubroot resistance gene introgressed from rutabaga (Summanwar et al. 2019) located on BnaA08 (Hasan and Rahman 2016). Taken together, all these observations suggest that the primary and secondary metabolic pathways, including the phenylpropanoid pathway, may play an important role in conferring resistance to clubroot disease in canola. This is also consistent with the results from other genomics and proteomics-based studies where the modulation of primary and secondary metabolic pathways during plant-pathogen interactions has been observed (Patil and Kumudini 2019; Ranjan et al. 2019; Sharma et al. 2007). However, the implication of these pathways in the regulation of clubroot resistance needs to be further investigated through overexpression and/or knock out of the important genes from these pathways.

4.4.3 qRT-PCR validation of the selected lncRNAs and their target genes

We also validated the relative expression of the selected lncRNAs, and their target genes based on their direct or indirect association with chromosome BnaA03 to which the clubroot resistance has been mapped (Fredua-Agyeman and Rahman 2016). For instance, the LNC_000340 was found to be upregulated in the R plants and is predicted to regulate two targets — one of which is an ethylene-responsive transcription factor 109 (*ERF109*) (BnaA03g52830D). Transcription factors of the *ERF* family are known to be involved in mediating responses to biotic and abiotic stress (Bahieldin et al. 2016; O'Donnell et al. 1996). Another lncRNA located on chromosome BnaA03, LNC_000354, was detected only in the S plants and it is predicted to regulate a gene encoding probable protein phosphatase 2C 2 (PP2C2) (BnaA10g04610D) (Fig. 4.4). PP2C2 belongs to the largest protein phosphatase family that are known to be the regulators of the pathways involved in plant defense and development (for review, see Schweighofer et al. 2004). Widjaja et al. (2010) reported that the *PP2C* is required for defense response against *P. syringae* and provide both positive and negative regulation of gene expression involved in plant defense. Similarly, the LNC_000421 was also expressed only in the S plants and was predicted

to regulate four genes; among these, the GATA transcription factor 17 (BnaC07g33440D) belongs to a group of DNA binding proteins widely found in eukaryotes (Gupta et al. 2017). According to Gupta et al. (2017), GATA factor might be involved in the regulation of signalling in response to abiotic stress in rice. They also suggested that the GATA transcription factors can create splicing variants under different environmental conditions. However, the specific role of these lncRNAs, their respective target genes and their roles in the regulation of gene expression remains to be investigated; nonetheless, the results from our study provided some insights into their possible involvement in mediating resistance to clubroot disease.

4.4.4 LncRNAs expressed in canola lines carrying rutabaga- and Mendel-resistance

Comparative analysis of the DE lncRNAs identified in this study and the 530 DE lncRNAs that we identified previously by using clubroot resistant canola lines carrying resistance introgressed from rutabaga (Summanwar et al. 2019) disclosed only 12 (about 2.2%) of these showing a similar trend of upregulation in both types of resistant lines; among these, 10 (about 1.88% of the total) lncRNAs regulated the genes involved in defense-related pathways (Supplementary Table 4.9; Fig. 4.5A) Therefore, it is highly likely that, at least, some of these lncRNAs play a critical role in resistance to clubroot disease. Of the remaining large majority of the DE lncRNAs, expression of some of these lncRNAs might be unique for these two resistances (BnaA03 and BnaA08) while the others might have resulted from the use of a limited number of DH lines (12 lines) from two genetically distinct populations. However, a similar pattern of upregulation of the above-mentioned 12 common lncRNAs and their target genes, and further confirmation of randomly selected four of these by qRT-PCR analysis, suggests that these DE lncRNAs are potentially relevant for the resistance in both genotypes. The expression of the predicted target genes of the four common lncRNAs, LNC_001220, LNC_000496, LNC_000980 and LNC_002381, which are known to be involved in plant defense and/or plant pathogen interaction, were further confirmed by qRT-PCR. In this study, the LNC_000496 was predicted to regulate the genes involved in plant pathogen interaction which included *BAK1* (Chinchilla et al. 2007) and *RIN4* (Belkhadir et al. 2004); a gene coding lipid transfer protein (*LTP*) (BnaA03g25140D), which belongs to pathogenesis related protein family 14 (PR-14) (van Loon and van Strien 1999), also known to be involved in plant defense (Kirubakaran et al. 2008, Zhu et al. 2012); and a disease resistance protein (BnaA06g29870D) which is a probable NBS-LRR type protein (Noman et al. 2019). Similarly, the LNC_001220 has also been predicted to regulate

a disease resistance protein-coding gene (BnaC08g41340D). In addition to these, other target genes of LNC_001220 involved in plant defense included caffeic acid 3-0 methyltransferase (BnaA07g21250D), a gene involved in lignin biosynthesis found to be upregulated in *P. brassicae* infected tissues (Irani et al. 2019); a Ca^{2+} binding protein-coding gene called calmodulin-5 (BnaCnng20380D) (Aldon et al 2018) and an auxin responsive protein-coding gene called indoleacetic acid-induced protein 16 (BnaC05g47100D) (Fu and Wang 2011). A MYB related protein R1, which has been associated with drought tolerance in potato (Shin et al. 2011), was also predicted to be regulated by LNC_000980 in our study. Also, a major latex protein 34 (*MLP34*) (BnaC06g24430D) of *MLP* family of PR-10 proteins was predicted to be regulated by LNC_002381 in this study. Previously, an *MLP* like protein had been demonstrated to have role in protection against pathogen *Verticillium dahliae* in cotton (Yang et al. 2015). Thus, comparative analysis of lncRNA data from the lines carrying Mendel- and rutabaga-resistance suggest that the common 12 lncRNAs and their defense-related target genes might play an important role in mediating resistance to *P. brassicae*; however, functional validation of these common lncRNAs would provide further insight to these lncRNAs.

4.4.5 Development of lncRNA-based molecular markers

LncRNAs have been found to have high sequence conservation at the intra-species and sub-species levels (Deng et al. 2018). For example, Zhao et al. (2018) reported that about 84% of the lncRNAs are conserved in cotton (*Gossypium spp.*), and Shen et al. (2018) reported that the lncRNAs regulating lipid biosynthesis in *B. napus* are conserved in other members of the family Brassicaceae. Markers developed based on the conserved sequence regions can be used in molecular breeding with high efficiency; for example, Jaiswal et al. (2019) successfully developed SSR markers based on lncRNAs in five *Capsicum* spp. (*C. annuum, C. chinense, C. frutescens, C. pubescens* and *C. baccatum*). In our study, in addition to the identification of DE lncRNAs in the R and S plants, we also investigated the potential use of the information of the DE lncRNAs from chromosome BnaA03, harbouring resistance to clubroot disease (Fredua-Agyeman and Rahman 2016), in molecular breeding.

Molecular markers can be considered as functional markers if they are developed from the gene sequences (Lau et al. 2015); this type of markers offer several advantages over random DNA markers for application in plant breeding. Different types of gene-based markers have been reported in plants, for example, EST-SSR markers in Lei bamboo (Cai et al. 2019), gene-tagged

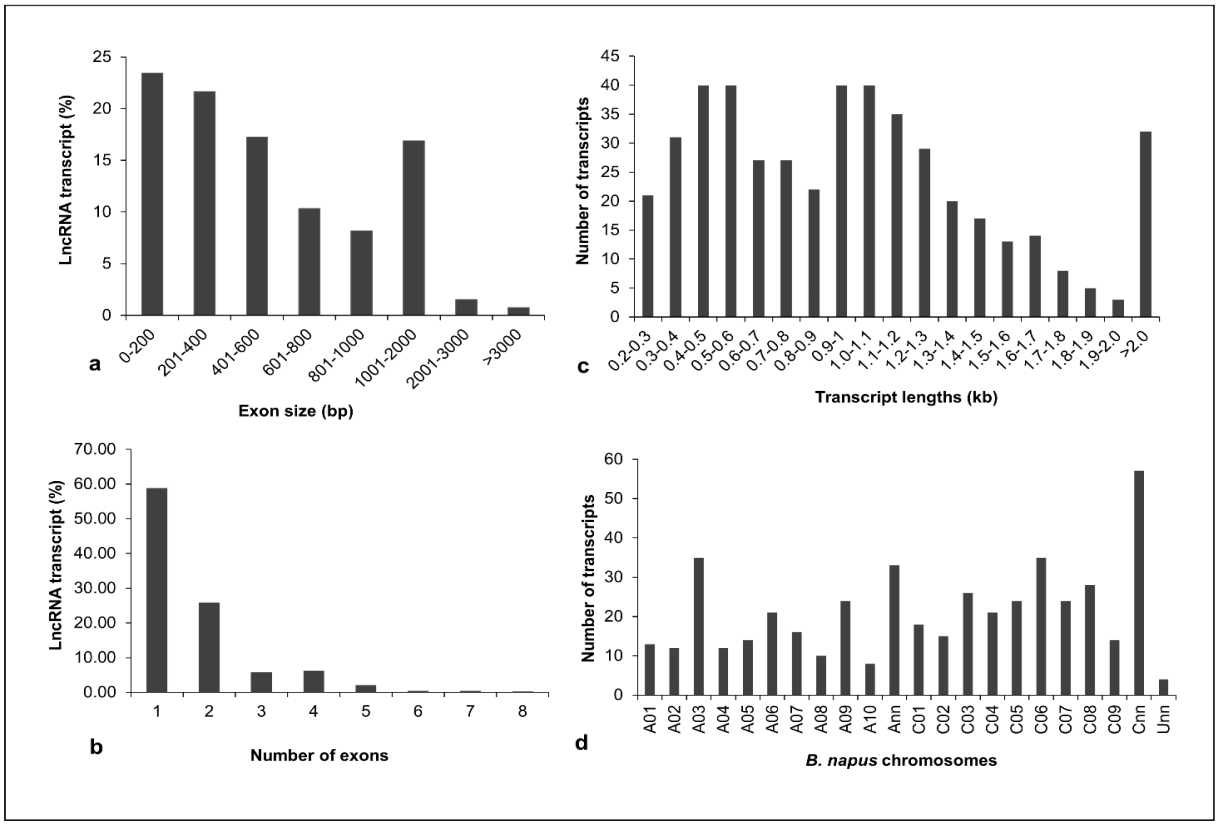
SSR markers in apple (Zhen et al. 2018), inter-retroelement amplified polymorphism (IRAP) markers in *Aegilops tauschii* (Saeidi,et al. 2008), and miRNA-SSR markers in rice (Ganie and Mondal 2015; Mondal and Ganie 2014) and in Brassica (Fu et al. 2013). To date, increasing evidence supports that the lncRNAs are involved in the regulation of functional genes; therefore, markers developed based on lncRNAs can be considered as functional markers.

In the present study, the development of an SSR marker from the lncRNA LNC_000424 and establishment of its association with clubroot resistance has paved the way for exploring the Brassica lncRNAs for use in breeding. This lncRNA is located at 22,295,257–22,297,928 bp position of BnaA03 of the *B. napus* genome (Genoscope *B. napus* Genome Browser, http://www.genoscope.cns.fr/brassicanapus/; retrieved on 25 November, 2019), and the SSR occurred at the 22,295,388–22,295,423 bp position; the primer pairs of the SSR marker from this lncRNA exhibiting association with clubroot resistance located at 22,295,341 and 22,295,531 bp position of BnaA03. However, the molecular markers reported by Fredua-Agyeman and Rahman (2016) linked to this resistance with about 2% recombination have been positioned at 40,936,414–41,929,968 bp of BnaA03 of the *B. napus* genome or at 24,376,817–24,684,311 bp of BraA03 of the *B. rapa* genome. The difference between the position of the lncRNA-based SSR marker identified in the present study and the position of the markers reported by Fredua-Agyeman and Rahman (2016) might have resulted from the updates of the Brassica genome sequence database. Through bioinformatic analysis, we predicted that the lncRNA LNC_000424 trans-regulates 45 target genes and cis-regulates 27 target genes (Supplementary Table 4.14; Fig 4.9); some of the target genes trans-regulated by this lncRNA are involved in plant defense (Supplementary Table 4.14; Fig 4.9). However, the failure of the marker designed from this lncRNA to detect the resistance phenotype in about 2% cases indicates that this lncRNA might not be involved in the regulation of clubroot resistance in Brassica. It is possible that this lncRNA is located about 2 cM away from the clubroot resistance gene, and this might have resulted the observed recombination between the marker and the resistance phenotype. Further investigation would be needed to confirm the possible role of this lncRNA in mediating resistance to clubroot disease in *B. napus*. Nonetheless, the SSR marker developed based on the LNC_000424 can be used in marker assisted breeding including pyramiding of multiple genes to develop clubroot resistant canola cultivars; however, recombination between the marker and the resistance phenotype needs to be taken into account.

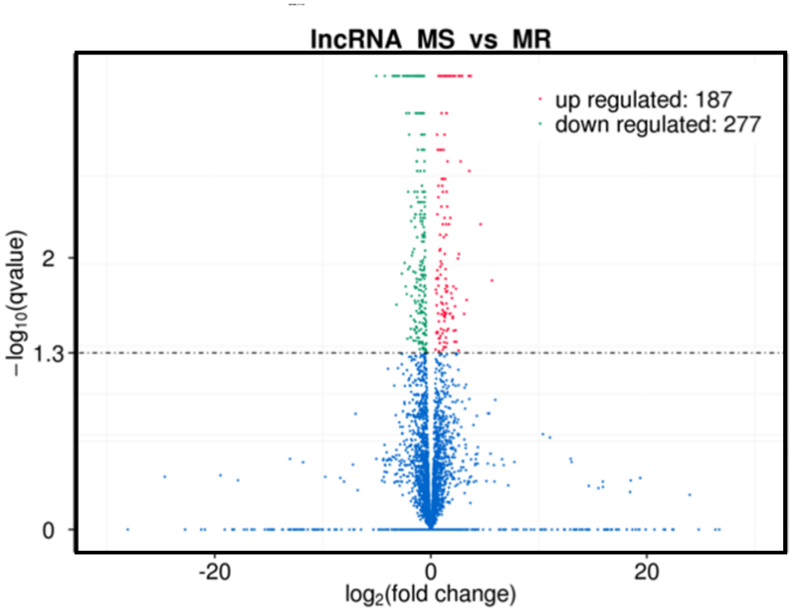
4.5 Conclusion

We have identified 464 DE lncRNAs in the roots of clubroot resistant canola plants challenged with *P. brassicae* pathotype 3. KEGG pathway enrichment analysis showed that the majority of the target genes regulated by these lncRNAs are involved in plant-pathogen interaction, plant hormone signaling, and in primary-secondary metabolic pathways. Bioinformatic analysis of these lncRNAs showed that some of these can function as endogenous target mimics of miRNAs or miRNA precursors. Comparison of the DE lncRNAs from this study and those we reported previously by using different resistant lines identified about 2.2% of the DE lncRNAs which might play a fundamental role in resistance to this disease. We also identified SSRs in about 40% of the DE lncRNAs where a marker from the lncRNA LNC_000424 of chromosome BnaA03 showed an association with clubroot resistance in 98% of the DH lines. Thus, the results from this study, including the development of lncRNA-based markers, extended our knowledge of the potential utility of the lncRNAs in the breeding of canola.

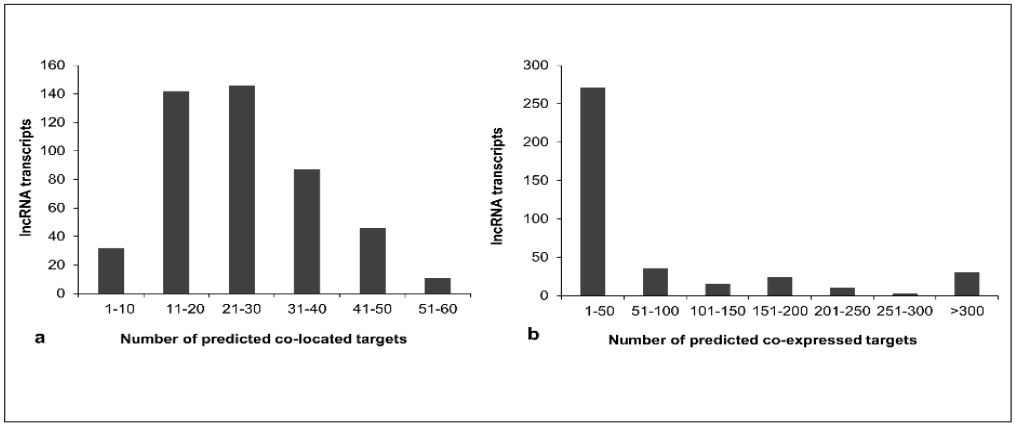
Supplementary Figures



Supplementary Fig. 4.1 Characterization of the 464 DE lncRNAs in *Brassica napus* in response to *Plasmodiophora brassicae* pathotype 3 A) size distribution of exons; B) number of exons per transcript of the lncRNAs; C) length distribution of the lncRNAs; and D) distribution of the lncRNAs on chromosomes



Supplementary Fig. 4.2 Volcano plot showing the differential expression of *Brassica napus* lncRNAs post infection with *Plasmodiophora brassicae* pathotype 3 with threshold q-value<0.05**.**



Supplementary Fig. 4.3 Differentially expressed lncRNAs and their predicted targets in *Brassica napus*. A) number of predicted co-located target genes; and B) number of predicted co-expressed target genes.

Chapter 5: Comparative transcriptome analysis of canola carrying clubroot resistance from 'Mendel' or Rutabaga and the development of molecular markers

5.1 Introduction

Brassica oilseed crops, such as *Brassica napus* (AACC, 2*n* = 38), *B. juncea* (AABB, 2*n* = 36) and *B. rapa* (AA, 2*n* = 20), provide about 15% of the total global vegetable oil supply (for review, see Rahman et al. 2013). Among these crops, *B. napus* is cultivated most extensively in the world, with Canada, China and the European union being the largest producers (for review, see Carré 2014). Among the different threats to production of *B. napus*, clubroot disease, caused by *Plasmodiophora brassicae* Woronin, is an important one (for review, see Howard et al. 2010). In addition to the Brassica oilseeds, clubroot disease can also cause significant yield loss in vegetable crops belonging to the family Brassicaceae (for review, see Howard et al. 2010). The potential for this disease-causing extreme loss in many Brassica crops has necessitated research for the identification and characterization of genes involved in mediating clubroot resistance as well as the development of markers for use in breeding (for review, see Rahman et al. 2014).

Plants being sessile in nature recognize and respond to pathogens and resist infection, whenever possible, by their genetic makeup. In most instances, the interaction between plants and pathogens involve the recognition of specific molecular signals. For instance, pathogens possess pathogen associated molecular patterns (PAMPs), which are recognized by the pattern recognition receptor (PRR) proteins in host cell (for review, see Jones and Dangl 2006). Recognition of the PAMPs by the PRRs results in the activation and deployment of defense responses that may ultimately lead to resistance against the pathogen (for review, see Jones and Dangl 2006). Some of the molecular factors responsible for resistance include genes associated with PAMPs-triggered immunity (PTI) and effector-triggered immunity (ETI). The activation of PTI or ETI activates a cascade of signalling events leading to the induction of genes encoding proteins responsible for resistance (for review, see Bolton 2009). Such genes include pathogenesis-related proteins, and proteins involved in the formation of reactive oxygen species, cell wall modifications, cell cycle alterations, calcium and hormone signalling, expression of various transcriptional factors, and mitogen-activated protein kinases (MAPKs) (Siemens et al. 2006, Chen et al. 2016, Zhang et al. 2016, Irani et al. 2018, Luo et al. 2018, Mei et al. 2019).

A number of reports that describe plant responses to *P. brassicae* at a molecular level have contributed to an enhanced understanding of the specific processes that may be involved in

mediating those responses. For instance, analysis of phytohormone profile in leaves, roots, and galls from *B. napus* infected with *P. brassicae* revealed that salicylic acid (SA) is involved in defense response (Prerostova et al. 2018). Jahn et al. (2013) reported the importance of auxin signalling pathways in the establishment of root galls by *P. brassicae* in *Arabidopsis*. *P. brassicae* is also observed to interfere with the hormonal homeostasis of the host plant, specifically affecting endogenous cytokinin (CK) levels, which are linked to cell divisions and root hypertrophy leading to formation of galls. The genes involved in CK biosynthesis have been reported to be differentially expressed during *P. brassicae* infection and gall formation in root tissues (for review, see Robin et al. 2019). It has also been reported that CK metabolism and its regulation is important during *P. brassicae* infection of plants belonging to the Brassicaceae (Siemens et al. 2006, Rolfe et al. 2016, Malinowski et al. 2016, Kong et al. 2018, Ciaghi et al. 2019). In addition, Irani et al. (2019) reported an increased expression of the genes involved in phenylpropanoid biosynthetic pathway in both roots and shoots of *Arabidopsis* plants infected with *P. brassicae*. Although the aforementioned studies have improved our knowledge about host responses to *P. brassicae* infection, our understanding of these molecular processes is far from being complete.

In a number of studies, efforts have been made to characterize the molecular-genetic basis of resistance to *P. brassicae* in Brassica. These studies have revealed that major genes conferring resistance to clubroot are frequently found to be located on the Brassica A genome (Hasan and Rahman 2016; for review, see Rahman et al. 2014) while resistance in the C genome is controlled by quantitative trait loci (QTL) (Farid et al. 2020) . Consequently, the resistance of the A genome has been used frequently to develop clubroot resistant Brassica oilseed (for review, see Rahman et al. 2014) and vegetable crops (for review, see Hirai 2006). For instance, the first clubroot resistant European winter canola cv. 'Mendel' has been reported to carry resistance located on the chromosome BnaA03 (Fredua-Agyeman and Rahman 2016) and this resistance has been introgressed into Canadian canola (for review, see Rahman et al. 2014) and used extensively for the development of clubroot resistant spring canola cultivars. It is also well known that multiple clubroot resistance genes are located on the chromosome BnaA03 (Fredua-Agyeman and Rahman 2016). Yu et al. (2016) integrated information from the chromosome region of BraA03 of *B. rapa* carrying clubroot resistance with transcriptome analysis and

identified putative candidate genes to be involved in resistance and developed molecular markers for use in breeding.

The other important resistance of the A genome which has been used in breeding is located on BnaA08 (Hasan and Rahman 2016; for review, see Rahman et al. 2014). However, limited study has so far been conducted to understand the molecular basis of this resistance. The genomic region of BnaA08, conferring resistance to clubroot, carry two genes, *Crr1a* and *Crr1b,* of which *Crr1a* encode Toll-Interleukin-1 receptor/nucleotide-binding site/leucine-rich repeat (TIR-NB-LRR) protein conferring resistance to *P. brassicae* isolate Ano-01 from Japan (Hatakeyama et al. 2013). The genomic region of BnaA08 introgressed from rutabaga (*B. napus* var. *napobrassica*) into Canadian canola found to confer resistance to *P. brassicae* pathotypes 2, 3, 5, 6 and 8 (Hasan and Rahman 2016) as well as to some of the recently identified pathotypes (unpublished data). To the best of our knowledge, no study has so far been conducted to compare the molecular basis of resistance of BnaA08 of rutabaga with BnaA03 of 'Mendel' – despite these two resistances have been used extensively in the breeding of spring canola cultivars (Rahman et al. 2014). A comparative study of these two resistances would extend our knowledge of the molecular genetic basis of clubroot resistance in *Brassica*; such study has so far not been reported in literature.

Advances in modern, "omics" technologies, including transcriptomics and proteomics have extended our knowledge of the molecular changes that occur in a plant following pathogen infection. For instance, a number of next-generation sequencing (NGS) studies (Chang et al. 2019, Zhu et al. 2019) and proteome studies (Devos et al. 2006, Sharma et al. 2007, Cao et al. 2008, Liang et al. 2008, Song et al. 2016) have been used to characterize plant responses to pathogens as well as for characterization of the resistance responses in plants (Chen et al. 2016, Zhang et al. 2016, Jia et al. 2017, Irani et al. 2018, Ciaghi et al. 2019, Summanwar et al. 2019, Li et al. 2020). NGS approaches, combined with the knowledge of the genomic regions involved in the control of resistance, can lead to the identification of putative candidate genes involved in mediating clubroot resistance. This information can also be used to develop gene-based molecular markers for use in breeding (Andersen et al. 2003, Poczai et al. 2013). Such gene-based markers have the inherent advantage of eliminating the potential recombination events, which often occur for the markers designed from the flanking region of the genes (Andersen et al. 2003, Poczai et al. 2013). Thus, not only do gene expression studies provide important

insights into the resistance mechanisms, but they can also be used in breeding important crops, such as canola, for clubroot resistance.

The objectives of this study were to (i) identify the genes that are differentially expressed in spring *B. napus* canola lines carrying clubroot resistance introgressed from rutabaga (BnaA08) and the lines carrying resistance introgressed from 'Mendel' (BnaA03), as compared to the lines lacking resistance, in response to infection by *P. brassicae*; (ii) perform a comparative analysis of the changes in gene expression due to resistance sources (BnaA08 vs. BnaA03) and identify the genes which could be involved in resistance to this disease. Additionally, we report the development of gene-based molecular markers from the differentially expressed gene sequences to provide molecular tools for the development of clubroot resistant Brassica crop cultivars.

5.2 Materials and methods

5.2.1 Plant materials

Two doubled-haploid (DH) populations were used in this study. The first DH population (Rt) was developed from F_1 plants of Rutabaga-BF × UA AlfaGold. The parent Rutabaga-BF is a clubroot resistant inbred line derived from rutabaga cv. Brookfield through self-pollination of single plants (Hasan and Rahman 2016). The other parent UA AlfaGold is an open-pollinated spring-type *B. napus* canola cultivar (Rahman 2017) which is susceptible to clubroot. The details of the development of the DH population and their phenotyping for resistance to *P. brassicae* pathotype 3 have been described by Hasan and Rahman (2016).

The second DH population (Md), consisting of the populations #1333 and #1330, was developed from F_1's of the crosses 1CA0591.263 × A07-26NR and 1CA0591.323 × A07-26NR, respectively. The parents 1CA0591.263 and 1CA0591.323 carry clubroot resistance, and were developed from cross between the winter canola cv. 'Mendel' and a Canadian spring canola line (Rahman et al. 2011), while the parent A07-26NR is susceptible to clubroot. The details of the development of the aforementioned DH populations and their resistance to *P. brassicae* pathotype 3 have been reported by Fredua-Agyeman and Rahman (2016).

5.2.2 Plant materials for RNA-seq analysis

A total of 192 plants which consisted of eight plants each of the 12 resistant (RtR) and 12 susceptible (RtS) DH lines of the DH population Rt were grown in a greenhouse (22/15 °C day/night and 16 h photoperiod) in 72-cell trays filled with Sunshine Professional Growing Mix

(Sun Gro Horticulture, USA). The seedlings were inoculated with *P. brassicae* single spore isolate SACAN-ss1 (pathotype 3) at 10 days after germination using the pipette method (Voorips and Visser 1993). The details of inoculum preparation, inoculation of the plants, and root tissue collection for RNA extraction are described in Summanwar et al. (2019). After inoculation, the soil was kept saturated with water to ensure successful infection and gall development. Similarly, 192 seedlings of 12 resistant (MdR) and 12 susceptible (MdS) lines of the DH population #1333 carrying Mendel-resistance were grown and inoculated to obtain root tissue for RNA extraction. The two experiments were repeated three times representing three biological replicates.

5.2.3 RNA isolation, library preparation and sequencing

In our earlier study (Summanwar et al. 2019), we observed that the initial infection and disease symptom development occur 10 days after inoculation; therefore, root tissues of the susceptible and resistant lines collected at this stage were used for RNA-seq analysis. In order to accomplish this, root tissues of three seedlings from each of the 12 RtS and 12 RtR lines as well as 12 MdS and 12 MdR lines was used for RNA isolation, and the samples were sent to Novogene (USA) for library preparation and sequencing.

5.2.4 Mapping to the reference genome and transcriptome assembly

RNA-sequencing was performed on Illumina HiSeq platform, which generated paired-ended (150 × 2 bp) reads. Novogene's in-house Perl transcripts were used for removal of the low-quality reads, adapter sequences and unrecognized bases. The Q20, Q30 and GC content was determined from the clean reads. The clean reads were indexed using Bowtie v2.0.6 (Langmead and Salzberg 2012) and aligned to the *B. napus* reference genome v5.0 (Chalhoub et al. 2014) by using TopHat v2.0.9 (Trapnell et al. 2012). The assembly of the mapped reads was carried out using Scripture (beta2) (Guttman et al. 2010) and Cufflinks (v2.1.1) (Trapnell et al 2010). Expression analysis of the mRNAs for each sample was calculated based on Fragments Per Kilobase pair of exon model per Million fragments mapped (FPKM) by taking the length of the fragments and count of the mapped reads into account and using Cuffdiff (v2.1.1) (Trapnell et al. 2010). The genes were considered differentially expressed if the adjusted *p*-value was < 0.05. To acquire the Gene Ontology (GO) annotations, enrichment analysis was carried out using GOSeq R package (GOSeq, topGO, hmmscan; (Young et al. 2012)). GO terms with corrected *p*-value less than 0.05 were considered significantly enriched by the differentially expressed genes.

For further annotation of the DEGs from the significantly enriched pathways, KEGG Orthology-Based Annotation System KOBAS v2.0 software (http://www.genome.jp/kegg/, (Mao et al. 2005)) was used.

5.2.5 Phenotyping of canola seedlings treated using trehalose

Seeds of each of seven RtS and RtR DH lines were surface sterilized using 10% bleach for 15 min and 70 % ethanol for 1 min, followed by washing three times with sterile distilled water (SDW). The sterilized seeds were placed on Whatman filter paper Grade 2 (GE Healthcare; USA) in Petri dishes containing SDW (T0 sample) and trehalose sugar solution of 30 mM, 60 mM, 90 mM, and 120 mM concentration for seed priming. In total, eight seeds from each of the susceptible and resistant DH lines were grown for each trehalose treatment, and the experiment was repeated three times which constituted three replicates. The Petri dishes were sealed using Parafilm (Fisher Scientific, USA) and placed in an incubator at 22/15 °C temperature (day/night) with a 16 h photoperiod for five days or until germination. The germinated seedlings were inoculated with the pathogen by dipping the roots in a resting spore suspension (1×10^7 spores/ml) of *P. brassicae* isolate SACAN-ss1 (pathotype 3) for 20 sec and were transferred to soil (Sunshine Professional Growing Mix). The plants were grown for 45 days in the greenhouse and phenotyped for resistance to clubroot disease as described by Hasan and Rahman (2016), and disease severity index was calculated using the method of Horiuchi and Hori (1980) as modified by Strelkov et al. (2006):

$$DSI(\%) = \frac{\sum(n\times0+n\times1+n\times2+n\times3)}{N\times3} \times 100$$

where, n is the number of plants in each class, N is the total number of plants, and 0, 1, 2 and 3 are the disease symptom severity classes. A DSI of 0 indicates the absence of visible galls in any of the plants, i.e., highly resistant, while a DSI of 100 indicates severe galling in all plants, i.e., extreme susceptibility.

5.2.6 Quantitative real time polymerase chain reaction (qRT-PCR) analysis

The expression pattern of 12 candidate genes known to be involved in plant defense responses and identified in both DH populations (Rt and Md) were validated using qRT-PCR. All qRT-PCR primers used in this study are listed in Supplementary Table 5.1. Prior to cDNA synthesis using iScript™ cDNA Synthesis Kit (Bio-Rad, USA), RNA from root tissues of the

inoculated susceptible and resistant plants of both DH populations were treated with DNase. All qRT-PCR reactions were performed using PowerUP SYBR Green Master Mix from Applied Biosystems (Thermo Fisher Scientific, USA) on QuantStudio 6 Flex Real-Time PCR System (Thermo Fisher Scientific, USA). Two technical replicates of each biological replicate were analyzed. Ubiquitin conjugating enzyme 9 (*UBC9*), a housekeeping gene, was used as an endogenous control (Joshi et al. 2016). Fold change of the candidate genes were calculated using the formula:

$2^{-\Delta\Delta Ct}$= ((Ct gene of interest – Ct endogenous control) R –
(Ct gene of interest – Ct endogenous control) S) (Schmittgen et al. 2008)

For the evaluation of gene expression following trehalose treatment and pathogen challenge, six seedlings of each of the three RtS and RtR lines were used in each biological replicate, and the experiment was repeated three times. The seeds were sterilized and germinated on Whatman filter paper Grade 2 in Petri dishes as described previously. The Petri dishes were sealed with Parafilm (Fisher Scientific, USA) and placed in an incubator at 22/15 °C temperature (day/night) with a 16 h photoperiod for five days or until germination. The seedlings were infected by dipping the roots in a resting spore suspension (1×10^7 spores/ml) of *P. brassicae* isolate SACAN-ss1 (pathotype 3) for 20 seconds and were transferred to soil. The roots of the treated seedlings were harvested at 10 dpi and flash frozen until qRT-PCR analysis.

5.2.7 Development of SSRs from the DEGs

The DEGs from the RNA-seq carried out between the RtS and RtR were analyzed and those from BnaA08 showing upregulation in the resistant plants (<-1fold change; Supplementary Table 5.2) were selected for the development of molecular marker. For this, the candidate genes sequences were downloaded from the *Brassica napus* genome resource Genoscope (http://www.genoscope.cns.fr/brassicanapus/ (Chalhoub et al. 2014)) and scanned for the presence of simple sequence repeat (SSR) motifs using the web-based microsatellite finder tool MISA-web (https://webblast.ipk-gatersleben.de/misa/ (Beier et al. 2017)). Twenty-four primer pairs were generated for 17 SSRs from 15 DEGs located on BnaA08 and were used to investigate their association with resistance to *P. brassicae* pathotype 3 using the DH population Rt reported by Hasan and Rahman (2016). For this, genomic DNA of five resistant and five susceptible DH lines and their parents were used for identification of the polymorphic markers.

PCR was carried out as described by Fredua-Agyeman and Rahman (2016) and the PCR conditions were as described by Green et al. (2018). The polymorphic markers from BnaA08 were used to genotype the DH population Rt by gel electrophoresis.

Similarly, the DEGs obtained from the RNA-seq of MdS and MdR lines were analyzed and those from BnaA03 showing upregulation in the resistant plants (<-1fold change; Supplementary Table 5.2) were selected for the development of SSR markers. In this case, 47 primer pairs were designed from the flanking sequences of 34 DEGs and nine primer pairs were designed based on sequences of nine DEGs. All these 56 SSR markers were tested for polymorphism, and the polymorphic markers were tested on the two DH populations #1330 and #1333 (Fredua-Agyeman and Rahman 2016) using the ABI sequencer 3730 (Applied Biosystems, USA). Markers when present in the genomic DNA were scored "+" and when absent were scored "-"only clear bands with sharp peaks were included in this analysis. Recombination frequency (RF) for the markers with clubroot resistance was calculated based on the number of recombinant lines divided by the total number of lines.

5.3 Results

5.3.1 RNA-seq analysis

An NGS experiment was conducted to identify the differentially expressed genes (DEGs) in *B. napus* plants carrying clubroot resistance introgressed from rutabaga from which we previously reported the identities of DE lncRNAs (Summanwar et al. 2019). In this study, we analyzed the data further to identify the DEGs with protein coding potential from the RtS and RtR plants in response to *P. brassicae* infection. As reported previously (Summanwar et al. 2016), our consolidated NGS approach (i.e., including DE lncRNAs and DE protein coding genes) generated 461 million paired-end reads from six strand-specific libraries of which 79.8% were clean reads and 73.5% were unique reads. Similarly, from another consolidated sequencing experiment using *B. napus* plants carrying clubroot resistance introgressed from European winter canola cv. 'Mendel', we established the identities of protein-coding DEGs between MdS and MdR plants in response to *P. brassicae* infection. This consolidated experiment generated >445 million paired-end reads, including 79.9% clean reads and 73.9% unique reads which were mapped to the *B. napus* reference genome.

Our results indicated that a total of ~7,500 genes were DE between the RtS and RtR lines as well as between the MdS and MdR lines; among these, ~3,500 were upregulated and ~4,000 were

downregulated in both the RtR and MdR lines (Supplementary Fig. 5.1A and 5.1B). Among these, 64 genes were uniquely expressed in the RtR plants, whereas 38 were expressed only in the RtS plants. Similarly, 40 genes were expressed only in the MdR plants and 86 were uniquely expressed in the MdS plants. In order to further investigate which DEGs may be responsible for clubroot resistance in both types of resistant lines, we focused on the genes that were observed to be upregulated in the RtR and MdR plants and their roles in plant defense related pathways.

5.3.2 Functional annotation of the genes upregulated in RtR and MdR lines using GO enrichment and KEGG pathway analysis

The genes upregulated in RtR and MdR lines were assigned to three GO classes viz. biological process, cellular component, and molecular process. The major sub-categories of the biological process, which included over 100 genes of RtR and MdR, were "single-organism metabolic process", "oxidation-reduction process", "transmembrane transport", and "single-organism transport" process (Supplementary Fig. 5.2A), and for the molecular process were "oxidoreductase activity", "transporter activity", and "transmembrane transporter activity" (Supplementary Fig. 5.2B). The most commonly represented five sub-categories of cellular component of the upregulated genes of RtR and MdR were "cell periphery", "extracellular region", "external encapsulating structure", "cell wall" and "apoplast" (Supplementary Fig. 5.2C). KEGG pathway enrichment analysis indicated that these genes were involved in primary metabolism, biosynthesis of secondary metabolites and phenylpropanoids suggesting that these metabolic processes may be involved in the resistant canola lines in response to infection by *P. brassicae.* Genes involved in plant hormone signalling pathways and plant-pathogen interaction were more represented in the MdR plants, while the genes from carbon and amino acid metabolic pathways were more upregulated in the RtR plants (Supplementary Table 5.3).

5.3.3 DEGs potentially involved in resistance to *P. brassicae*

5.3.3.1 Primary metabolism

Approximately 8-9% of the genes involved in primary and secondary metabolism were upregulated in the RtR and MdR lines. While comparing the two types of resistant lines, a greater number of genes (83) involved in primary metabolic pathways, such as the pentose phosphate pathway (PPP), glycolysis/gluconeogenesis, and starch and sucrose metabolism, were upregulated in the RtR plants as compared to the MdR plants (57 genes) **(**Fig. 5.1). However,

some genes from these pathways, including phosphoglycerate kinase, phosphophenolpyruvate carboxykinase and acid beta-fructofuranosidase were upregulated in both types of resistant lines. Similarly, a greater number of genes involved in amino acid metabolism, mevalonic acid pathway (MVA) and methylerythritol 4-phosphate (MEP) were upregulated in the RtR plants (85) as opposed to the MdR plants (40) (Fig. 5.1). In MdR lines, a greater number of genes involved in secondary metabolism, for instance phenylpropanoid, lignin, flavonoid and stilbene biosynthesis, were upregulated as compared to the RtR lines (Fig. 5.1), however, some genes from these pathways including trans-cinnamate 4-monooxygenase, 4-coumarate--coA ligase-like 9, flavonol synthase 3, caffeic acid 3-O-methyltransferase, flavone 3'-O-methyltransferase 1, 4-coumarate--CoA ligase-like 9 and peroxidase 2 were upregulated in both RtR and MdR (Fig. 5.1; Supplementary Table 5.4). We also observed that the lipoxygenase *(LOX)* genes were upregulated only in the MdR lines (Fig. 5.1). Log2 (fold change) acquired from the RNA-seq experiment for the genes upregulated in RtR and MdR as shown in Fig. 5.1 are reported in Supplementary Table 5.4.

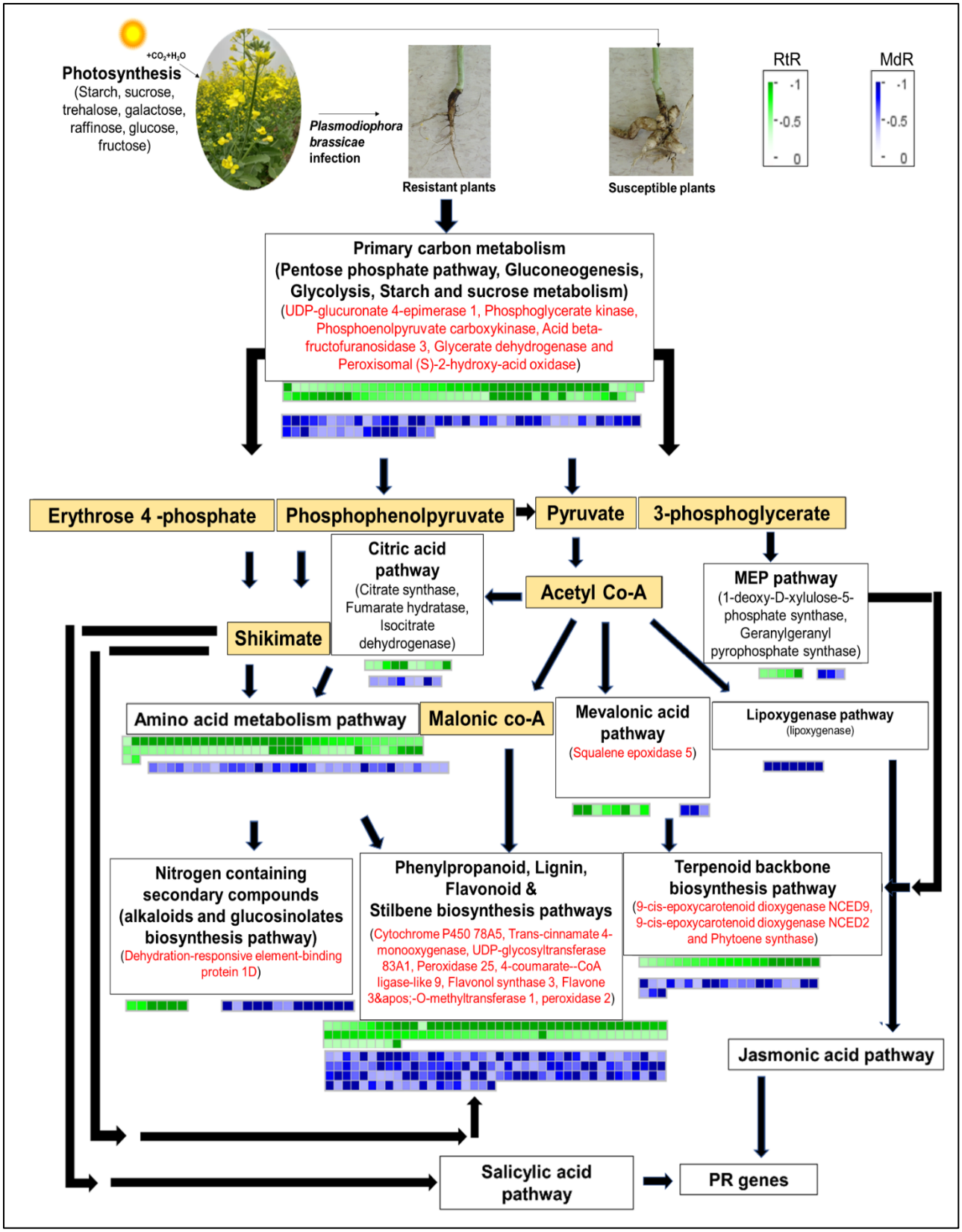

+CO2+H2O
Photosynthesis
(Starch, sucrose, trehalose, galactose, raffinose, glucose, fructose)
Plasmodiophora brassicae infection
Resistant plants
Susceptible plants
RtR
-1
-0.5
0
MdR
-1
-0.5
0
Primary carbon metabolism (Pentose phosphate pathway, Gluconeogenesis, Glycolysis, Starch and sucrose metabolism)
(UDP-glucuronate 4-epimerase 1, Phosphoglycerate kinase, Phosphoenolpyruvate carboxykinase, Acid beta-fructofuranosidase 3, Glycerate dehydrogenase and Peroxisomal (S)-2-hydroxy-acid oxidase)
Erythrose 4 -phosphate
Phosphophenolpyruvate
Pyruvate
3-phosphoglycerate
Citric acid pathway
(Citrate synthase, Fumarate hydratase, Isocitrate dehydrogenase)
Acetyl Co-A
MEP pathway
(1-deoxy-D-xylulose-5-phosphate synthase, Geranylgeranyl pyrophosphate synthase)
Shikimate
Amino acid metabolism pathway
Malonic co-A
Mevalonic acid pathway
(Squalene epoxidase 5)
Lipoxygenase pathway
(lipoxygenase)
Nitrogen containing secondary compounds (alkaloids and glucosinolates biosynthesis pathway)
(Dehydration-responsive element-binding protein 1D)
Phenylpropanoid, Lignin, Flavonoid & Stilbene biosynthesis pathways
(Cytochrome P450 78A5, Trans-cinnamate 4-monooxygenase, UDP-glycosyltransferase 83A1, Peroxidase 25, 4-coumarate--CoA ligase-like 9, Flavonol synthase 3, Flavone 3'-O-methyltransferase 1, peroxidase 2)
Terpenoid backbone biosynthesis pathway
(9-cis-epoxycarotenoid dioxygenase NCED9, 9-cis-epoxycarotenoid dioxygenase NCED2 and Phytoene synthase)
Jasmonic acid pathway
Salicylic acid pathway
PR genes

Fig. 5.1 An overview of the primary metabolic pathways and their relationship with the secondary metabolic pathways, plant hormone signalling, and pathogenesis related genes involved in plants in response to pathogen attack. Boxed heat maps shows expression profile of the DEGs upregulated in the resistant lines carrying clubroot resistance of rutabaga (RtR) or clubroot resistance of winter canola cv. Mendel (MdR) (Yellow box, intermediate metabolites; white box, pathways; names in bold font and in brackets, pathway names; names in normal font and in brackets, important genes of the pathway; red font, common genes upregulated in both RtR and MdR lines. Heat maps: green, genes upregulated in RtR; blue, genes upregulated in MdR; Log2(fold change) of all the genes are given in Supplementary Table 5.3. The log2(fold change) = log2(susceptible lines (FPKM)/resistant lines (FPKM)); therefore, lower the value, higher the upregulation in the resistant lines)

5.3.3.2 Plant-pathogen interaction

a) Defense related genes and cell wall related proteins:

Genes belonging to plant-pathogen interaction and plant hormone signalling pathways were upregulated in both RtR and MdR lines. The MapMan package (Thimm et al. 2004) was used to display the host-pathogen interaction pathways for the DEGs. The genes which are known to be involved in plant defense, such as the genes encoding transcription factors, peroxidases, kinases, cell wall proteins and disease resistance proteins were identified in both RtR and MdR lines (Fig. 5.2A and 2B) indicating that, at least some of them might be involved in resistance to clubroot disease. Genes encoding disease resistance proteins, for example BnaA03g45000D, BnaA03g44110D, BnaA08g20170D, BnaA09g14980D and BnaAnng18850D, were specifically upregulated in RtR, i.e., their expression was influenced by the rutabaga-resistance (<-1) **(**Fig. 5.2A; Supplementary Table 5.5), while the genes BnaC06g23970D, BnaA09g44940D, BnaA04g28100D, BnaC02g47400D, BnaC05g47550D, BnaA07g26100D, BnaCnng11220D, BnaC06g34090D, BnaC06g34040D, BnaAnng28770D, BnaC01g40270D and BnaA07g28550D encoding disease resistance proteins were specifically upregulated in the MdR lines, i.e., their expression was influenced by the Mendel-resistance (<-1). A number of genes encoding defensin-like proteins, such as BnaA01g35960D, BnaC01g29140D, BnaA01g35950D, BnaC01g29230D, BnaA05g03400D and BnaC01g29250D, were induced in the MdR lines (Fig. 5.2B; Supplementary Table 5.5) but not in the RtR lines (Fig. 5.2A; Supplementary Table 5.5). Thus, it is apparent that a greater number of disease resistance genes and defensins were expressed in the MdR lines as compared to the RtR lines; some of these genes might play an important role in disease resistance, specifically, in the MdR plants.

A greater number of genes encoding cell wall related proteins were upregulated in the RtR lines (144) than in the MdR lines (102) (Fig. 5.2A and 5.2B). Some of the genes associated with cell wall functions, which were significantly upregulated (<-1) in the RtR and MdR lines, included those encoding enzymes involved in hemicellulose synthesis, cell wall precursor synthesis, pectin esterases and acetylesterases for cell wall modification and various pectate lyases and polygalaturonases for cell wall degradation. The two genes BnaC09g05460D and BnaA03g30660D, which are involved, respectively, in cell wall degradation and modification were upregulated (<-1) in both RtR and MdR lines (Fig. 5.2C; Supplementary Table 5.5).

b) Signalling

Plant hormones are critical for modulating the developmental processes and signalling networks in response to various environmental cues including biotic and abiotic stresses (Bari and Jones 2009, Derksen et al. 2013). Relatively a greater number of genes involved in auxin (RtR:35 vs. MdR:33), brassinosteroid (RtR:13 vs. MdR:7) and SA (RtR:9 vs. MdR:7) signaling were upregulated in the RtR plants; in contrast, a greater number of genes involved in abscisic acid (RtR:8 vs. MdR:16), ethylene (RtR:29 vs. MdR:63), and JA (RtR:6 vs. MdR:22) signalling were upregulated in the MdR lines **(**Fig. 5.2A,B, Supplementary Table 5.5) reflecting the differences between the RtR and MdR plants with respect to the expression of plant hormone signalling genes.

The difference between the MdR and RtR lines was also observed with respect to the expression of signalling proteins like receptor kinases, G proteins and Ca+2 signalling (Fig. 5.2A and 5.2B, Supplementary Table 5.5). For instance, 36 signalling related receptor kinases were upregulated in the MdR lines (<-1) as opposed to only nine in the RtR lines (<-1) (Supplementary Table 5.5). Similarly, a greater number of genes encoding G-proteins (RtR:25 vs. MdR:41) as well as those playing a role in calcium signalling (RtR:38 vs. MdR:72) were upregulated in the MdR lines (Supplementary Table 5.5). Thus, our study provides valuable information of the pathways involved in resistance to clubroot disease in canola carrying resistance introgressed from rutabaga and cv. Mendel. A global view of the genes upregulated in the RtR and MdR lines in response to infection by *P. brassicae* is presented in Fig. 5.2 and the details of the genes are given in Supplementary Table 5.5.

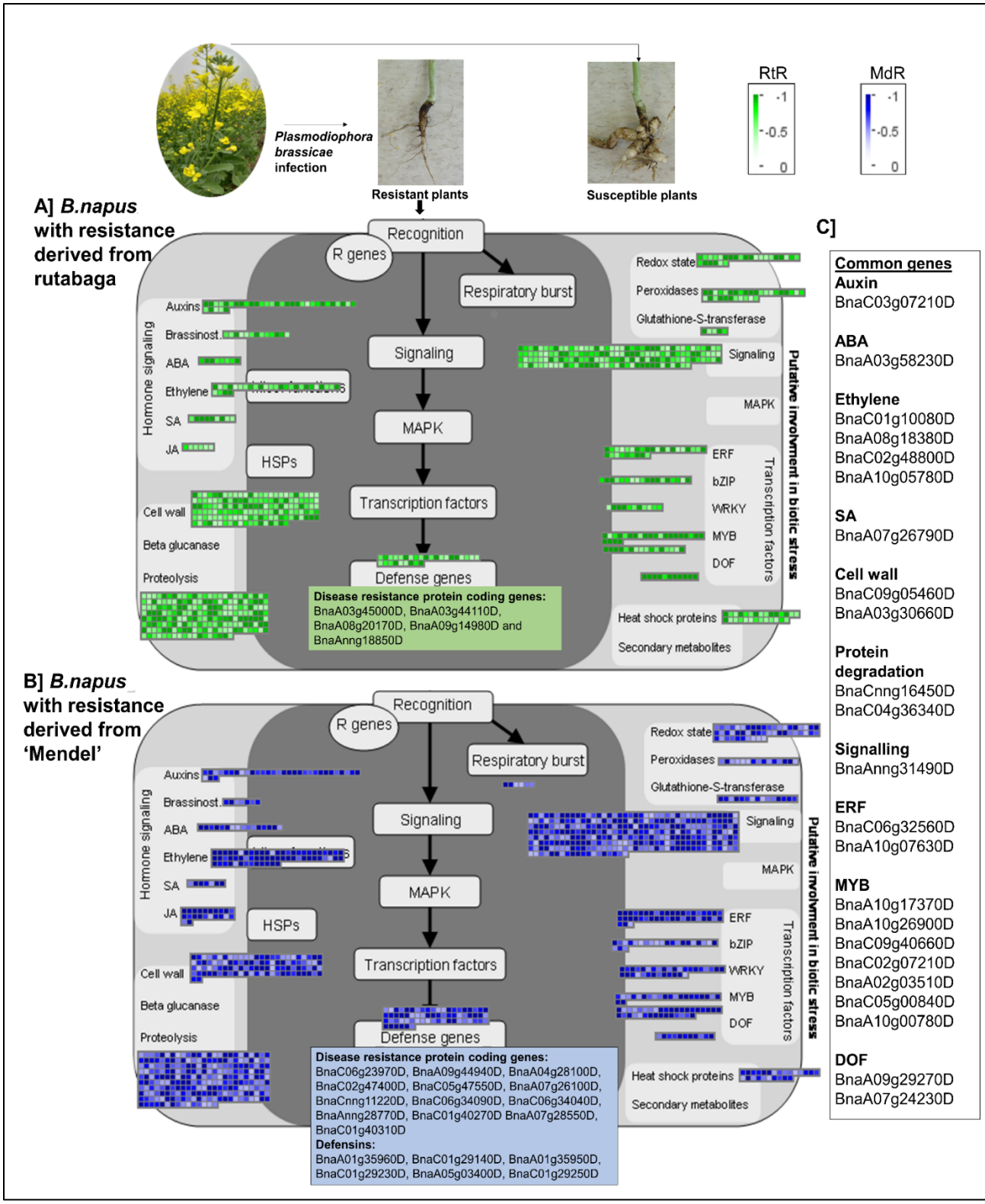

Plasmodiophora brassicae infection
Resistant plants
Susceptible plants
RtR
MdR
-1
-0.5
0
A] B.napus with resistance derived from rutabaga
Recognition
R genes
Respiratory burst
Signaling
MAPK
Transcription factors
Defense genes
HSPs
Hormone signaling
Auxins
Brassinost.
ABA
Ethylene
SA
JA
Cell wall
Beta glucanase
Proteolysis
Redox state
Peroxidases
Glutathione-S-transferase
Signaling
MAPK
ERF
bZIP
WRKY
MYB
DOF
Transcription factors
Putative involvment in biotic stress
Heat shock proteins
Secondary metabolites
Disease resistance protein coding genes:
BnaA03g45000D, BnaA03g44110D, BnaA08g20170D, BnaA09g14980D and BnaAnng18850D
B] B.napus with resistance derived from 'Mendel'
Disease resistance protein coding genes:
BnaC06g23970D, BnaA09g44940D, BnaA04g28100D, BnaC02g47400D, BnaC05g47550D, BnaA07g26100D, BnaCnng11220D, BnaC06g34090D, BnaC06g34040D, BnaAnng28770D, BnaC01g40270D BnaA07g28550D, BnaC01g40310D
Defensins:
BnaA01g35960D, BnaC01g29140D, BnaA01g35950D, BnaC01g29230D, BnaA05g03400D, BnaC01g29250D
C]
Common genes
Auxin
BnaC03g07210D
ABA
BnaA03g58230D
Ethylene
BnaC01g10080D
BnaA08g18380D
BnaC02g48800D
BnaA10g05780D
SA
BnaA07g26790D
Cell wall
BnaC09g05460D
BnaA03g30660D
Protein degradation
BnaCnng16450D
BnaC04g36340D
Signalling
BnaAnng31490D
ERF
BnaC06g32560D
BnaA10g07630D
MYB
BnaA10g17370D
BnaA10g26900D
BnaC09g40660D
BnaC02g07210D
BnaA02g03510D
BnaC05g00840D
BnaA10g00780D
DOF
BnaA09g29270D
BnaA07g24230D

Fig. 5.2 MapMan analysis showing putative genes upregulated in *Brassica napus* lines carrying clubroot resistance gene introgressed from rutabaga and winter canola cv. Mendel, and their involvement in different pathways in response to infection by *Plasmodiophora brassicae*. **A)** Genes upregulated in resistant lines carrying rutabaga-resistance (RtR); disease resistant protein-coding genes specific to this resistance are given in the green panel. **B)** Genes upregulated in resistant lines carrying Mendel-resistance (MdR); disease resistant proteins coding genes and defensins specific to this resistance are given in the blue panel. **C)** List of common genes upregulated in both RtR and MdR lines. Heat maps: green, genes upregulated in RtR; and blue, genes upregulated in MdR. Log2(fold change) of all the genes are given in Supplementary Table 5.5. The log2(fold change) = log2(susceptible lines (FPKM)/resistant lines (FPKM)); therefore, lower the value, higher the upregulation in resistant lines. (ABA: abscisic acid; SA: salicylic acid; JA: jasmonate; Brassinost: brassinosteroid; HSP: heat shock protein; MAPK: mitogen-activated protein kinase; PR-proteins: pathogenesis-related proteins; ERF: ethylene-responsive factors; bZIP: basic leucine zipper; DOF: DNA-binding with one finger)

5.3.4 Quantitative real time-PCR (qRT-PCR) validation of expression

To confirm the results of the RNA-seq, 24 DEGs identified in this study that may play a role in plant defense in the RtR and MdR lines were selected for the qRT-PCR assay. The selected genes belong to plant hormone signalling and phenylpropanoid pathways, encode defensins, PR genes, and the genes involved in plant-pathogen interaction. In all these cases, qRT-PCR results were consistent with gene expression data from RNA-seq analysis (Fig. 5.3).

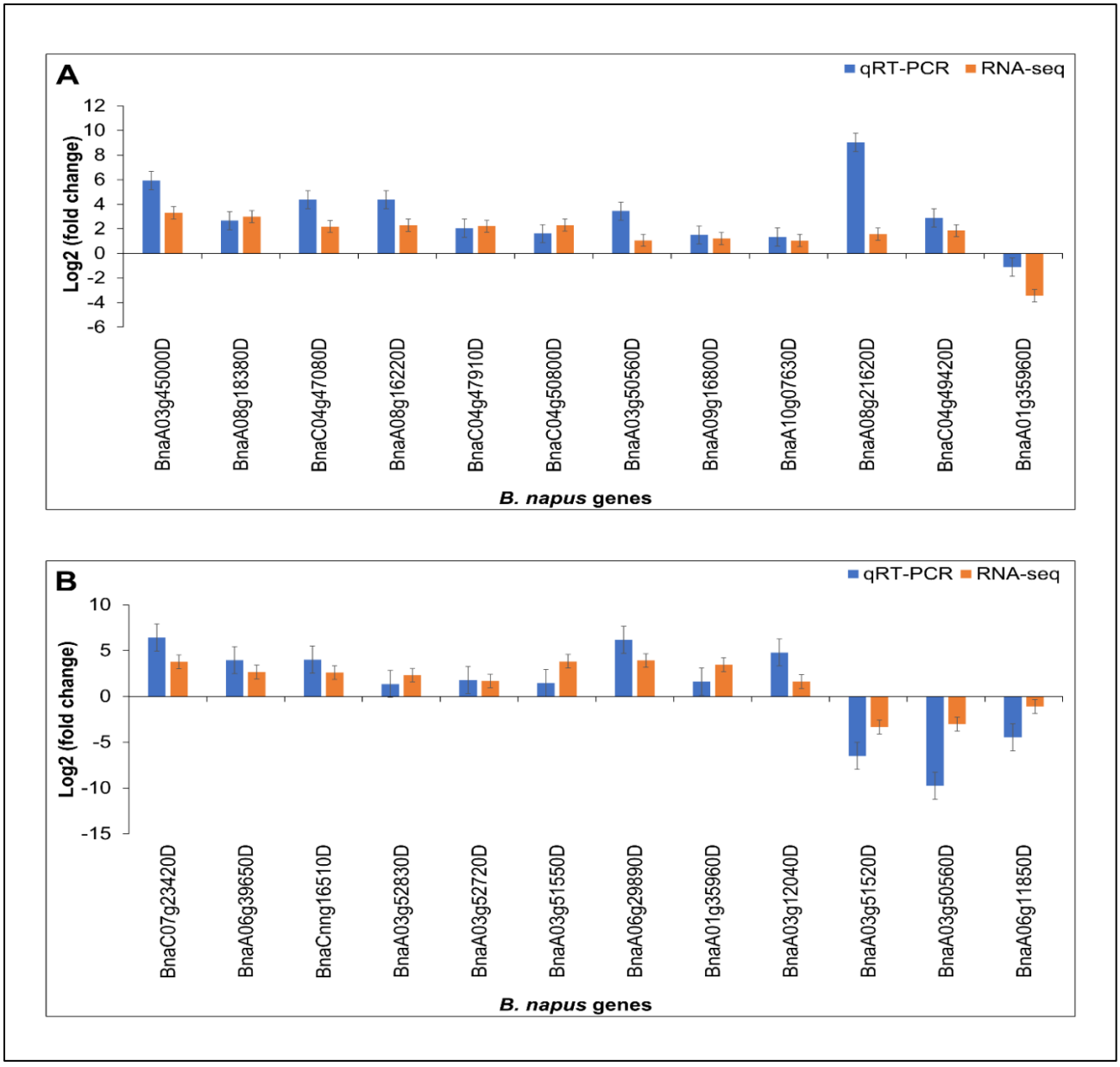


Fig. 5.3: **A) qRT-PCR validation of the differentially expressed genes (DEGs) identified by using A)** susceptible (RtS) and resistant (RtR) *Brassica napus* lines carrying clubroot resistance gene of rutabaga (*B. napus* var. *napobrassica*), and **B)** susceptible (MdS) and resistant (MdR) *B. napus* lines carrying clubroot resistance gene of winter canola cv. Mendel. All the DEGs selected for validation are known to be involved in plant defense. [Expression of the genes was normalized to *UBC9* (endogenous control). Log_2 (fold change) in R lines was calculated using the formula $2^{-\Delta\Delta Ct} = ((Ct_{\text{gene of interest}} - Ct_{\text{endogenous control}})R - (Ct_{\text{gene of interest}} - Ct_{\text{endogenous control}})S)]$. Three biological replicates and two technical replicates were used. Error bar = standard error].

5.3.4.1 Genes upregulated in both RtR and MdR lines and their validation using qRT-PCR

A total of 513 genes were observed to be upregulated in both RtR and MdR lines, (Supplementary Table 5.6). We selected five genes which were involved in plant defense-related pathways and displaying similar expression pattern in both RtR and MdR lines for further validation by qRT-PCR. This included the genes encoding ethylene-responsive transcription factor 109 (BnaA03g52830D), lipid transfer protein EARLI 1(BnaA06g38220D), disease resistance-like protein "constitutive shade-avoidance 1" (BnaC03g08650D), protein REVEILLE 1 (BnaC09g40660D) and RING-H2 finger protein (BnaCnng16450D). Expression profiles of these genes as gleaned from the qRT-PCR data were similar to that of the RNA-seq data (Fig. 5.4).

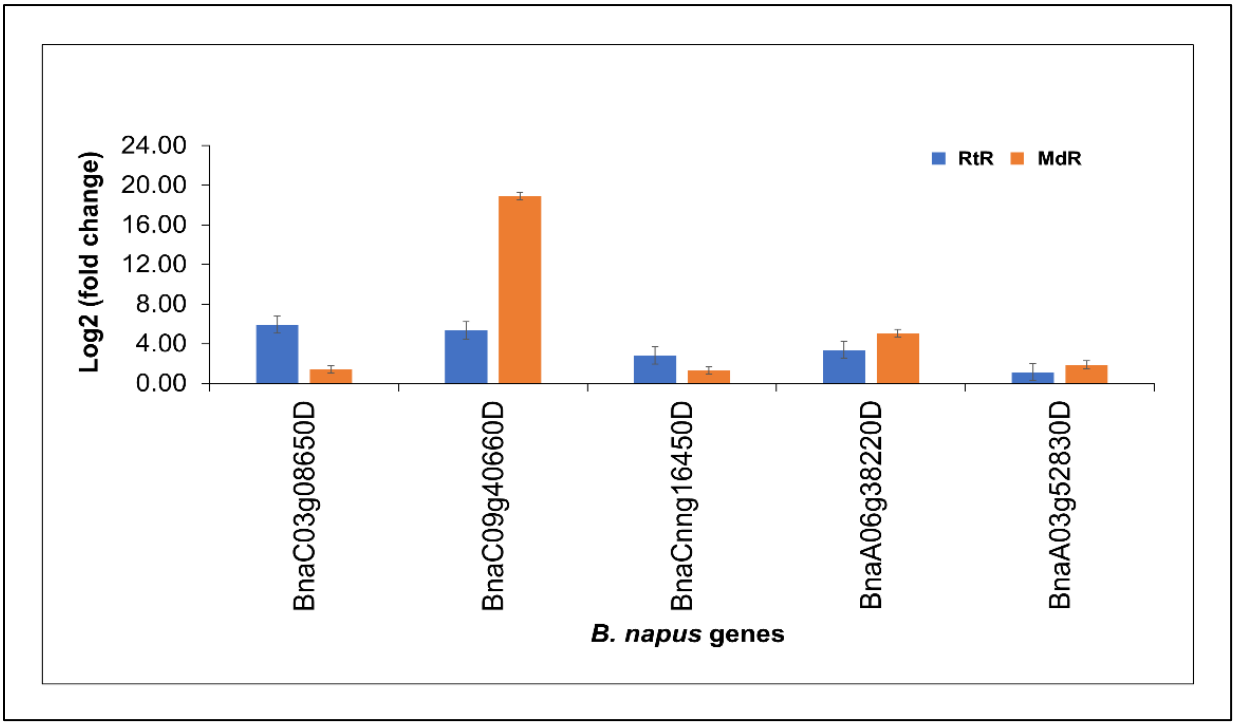


Fig. 5.4: qRT-PCR validation of the differentially expressed genes involved in plant disease response which are upregulated in clubroot resistant *Brassica napus* canola lines carrying resistance introgressed from rutabaga (RtR) or winter canola cv. Mendel (MdR). [Expression of the genes was normalized to *UBC9* (endogenous control). Log_2 (fold change)) in R lines was calculated using the formula $2^{-\Delta\Delta Ct} = ((Ct_{\text{gene of interest}} - Ct_{\text{endogenous control}})R - (Ct_{\text{gene of interest}} - Ct$

endogenous control)S)]. Three biological replicates and two technical replicates were used. Error bar = standard error].

5.3.5 Comparison between BnaA08 (RtR) and BnaA03 (MdR) resistance

We annotated a total of 44 genes from the chromosome BnaA08 of the RtR lines, which were upregulated and found to be involved in host-pathogen interaction (Fig. 5.5A, Supplementary Table 5.7); one of them is the PR disease resistance RPP5-like gene (BnaA08g20170D). This gene was upregulated only in RtR but not in MdR lines (Fig. 5.5A, Supplementary Table 5.7). Similarly, a total of 92 genes, which were upregulated and found to be involved in host-pathogen interaction, were annotated from the chromosome BnaA03 of the MdR lines (Fig. 5.5B, Supplementary Table 5.7). For the genes encoding disease resistance proteins, a single gene BnaA08g20170D was upregulated in the RtR line (Fig. 5.5A), while three genes viz. BnaA03g46200D, BnaA03g48510D and BnaA03g03830D were upregulated in the MdR lines (Fig. 5.5B). Furthermore, among these 92 genes, the most upregulated gene (fold change: -5.78042) was found to be an aspartic proteinase-like protein 1 (BnaA03g39150D) from BnaA03 of the MdR lines (Supplementary Table 5.7). However, the aforementioned genes which expressed in the MdR lines, but not in the RtR lines, can be considered to be specific to MdR. Some of the genes from JA, brassinosteroid and auxin signalling pathway were upregulated in the MdR lines (Fig. 5.5B), whereas not a single gene from these pathways were expressed in the RtR lines (Fig. 5.5A). Overall, a greater number of genes involved in plant response to biotic stress were found to be expressed in the MdR lines as compared to the RtR lines (Fig. 5.5A and 5.5B). This suggests that at 10 dpi, the clubroot resistance of the MdR lines was influenced by BnaA03, however, the resistance in the RtR lines could also be influenced by the defense related genes from other chromosomes.

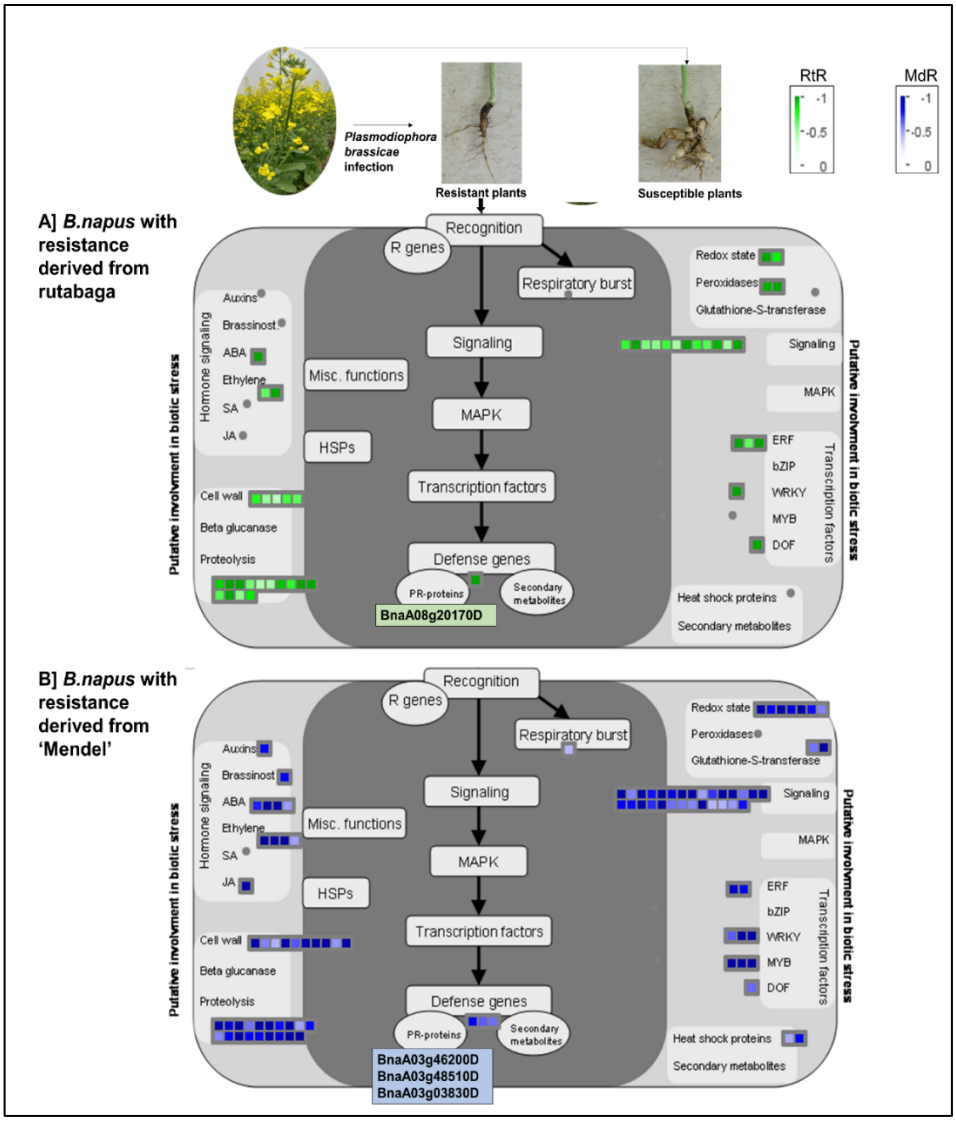


Fig. 5.5 MapMan analysis showing putative genes from clubroot resistant *Brassica napus* lines carrying rutabaga-resistance (RtR; BnaA08) and resistant lines carrying Mendel-resistance (MdR; BnaA03) and their role in different pathways involved in response to infection by *Plasmodiophora brassicae*. A) DEGs from BnaA08 upregulated in RtR; disease resistant genes specific to rutabaga-resistance are given in the green panel. B**)** DEGs from BnaA03 upregulated in MdR; disease resistant genes specific to Mendel-resistance are given in the blue panel. Heat maps: green, upregulated in RtR; blue, upregulated in MdR. Log2(fold change) of all the genes are given in Supplementary Table 5.5. The log2(fold change) = log2(susceptible lines (FPKM)/resistant lines (FPKM)); therefore, lower the value, higher the upregulation in resistant lines. (ABA: abscisic acid; SA: salicylic acid; JA: jasmonate; Brassino: brassinosteroid; HSP: heat shock protein; MAPK: mitogen-activated protein kinase; PR-proteins:

pathogenesis-related proteins; ERF: ethylene-responsive factors; bZIP: basic leucine zipper; DOF: DNA-binding with one finger)

We also identified several orthologous genes, which were upregulated in either BnaA08 of RtR or BnaA03 of MdR lines and were involved in biotic stress response. Six orthologous genes, viz. probable WRKY transcription factor 19, serine carboxypeptidase-like 29, serine carboxypeptidase-like 42, probable xyloglucan endotransglucosylase/hydrolase protein 18, ethylene-responsive transcription factor called "cytokinin response factor 4 (*CRF4*)", and ethylene-responsive transcription factor ERF060, were identified in both BnaA08 and BnaA03 (Fig. 5.6). Among these, four were upregulated in RtR and two in MdR, however, only the *CRF4*, which is an ethylene-responsive transcription factor, was upregulated in both types of resistant lines (Fig. 5.6). The *CRF4* (BnaA08g13940D) of BnaA08 is located at 11,929,168 – 11,930,155 bp, while the *CRF4* (BnaA03g48910D) of BnaA03 is located at 25,137,421 – 25,138,399 bp position of the *B. napus* genome. Thus, the upregulation of *CRF4* in both resistant lines indicate that this gene might play a fundamental role in resistance to *P. brassicae*.

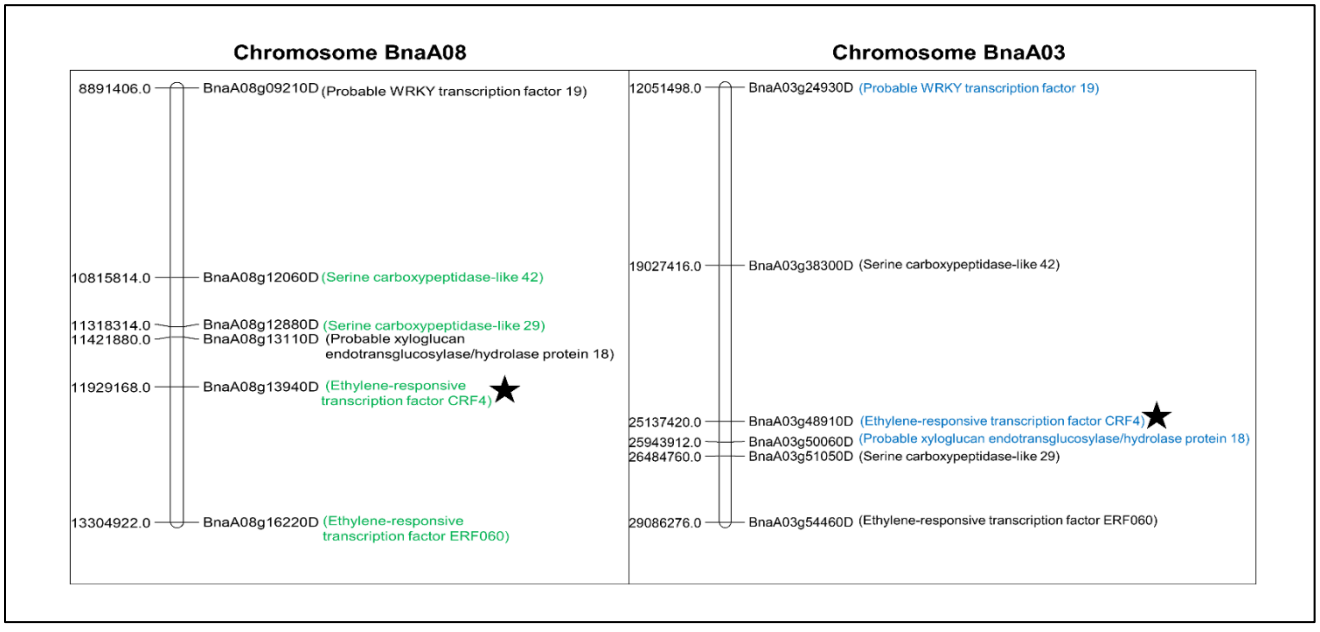


Fig. 5.6 Differentially expressed orthologous genes involved in plant defense found on chromosomes BnaA08 of RtR lines and BnaA03 of MdR lines of *Brassica napus*, carrying clubroot resistance derived, respectively, from rutabaga and winter canola cv. Mendel.

Green font, upregulated in RtR lines; blue font, upregulated in MdR lines; asterisk: upregulated gene found on chromosome BnaA08 of RtR lines and chromosome BnaA03 of MdR lines.

5.3.6 Evaluation of canola lines treated with trehalose for resistance to clubroot

As mentioned earlier, we observed that several genes involved in primary metabolic pathways were upregulated in the resistant lines, and some of these genes might be, at least, to some extent responsible for upregulation of the plant defense genes. Among the different primary metabolites, sugars are known to enhance oxidative burst at the early stage of infection by pathogen, increase lignification in cell walls, and induce the genes encoding pathogenesis-related (PR) proteins (for review, see Morkunas et al. 2014). It is well established that the carbon from the primary metabolic pathway generally flows to the secondary metabolic pathway for synthesis of different secondary metabolites such as phenols, lignins, flavones, alkaloids, and stilbenes (for review, see Bolton 2009, Huang et al. 2019). To understand the role of sugars in resistance to *P. brassicae*, we examined the effect of priming the susceptible and resistant DH lines with trehalose for resistance to clubroot disease. Of the four concentrations we used in this study, pooled data of the seven susceptible DH lines treated with 30 mM and 60 mM trehalose showed a partial resistance to this disease with DSI of about 45%, while the treatment with 120 mM trehalose and SDW did not change the phenotype of the DH lines, i.e. they were completely susceptible with DSI of 100% (Fig. 5.7A). The differences between the DSI values for the trehalose treatments were statistically significant ($P<0.05$). A similar ($P<0.05$) trend for the DSI value was also found among for the seven lines individually (data not shown). Priming of the resistant lines with trehalose did not exert any negative effect on the growth and development of the plants, as well as for resistance to clubroot disease (data not shown).

5.3.7 Quantification of defense responsive genes in response to priming with trehalose

The expression profile of six plant defense genes, viz. chitinase (*CHI*), pathogenesis-related 1(*PR1*), pathogenesis-related 2 (*PR2*), phenylalanine ammonia lyase (*PAL*), lipoxygenase (*LOX*) and peroxidase (*PER*) in response to priming of the three DH lines with trehalose is presented in Fig. 5.7. Among the six genes, a consistent expression of the genes in the three DH lines was found for *PR2*, *LOX* and *PAL.* The *PR2* gene was significantly upregulated (fold change: S1: 6.93; S2: 4.3; S3: 14.66) in T30 treatment as compared to the T0 treatment ($P<0.05$).

The *LOX*, a gene critical in lipoxygenase and JA pathway (Fig. 5.1), was significantly upregulated (fold change: S1: 16.27; S2: 7.99; S3: 4.65) in both T30 and T60 treatments ($P<0.05$) (Fig. 5.7B), while the gene *PAL* was upregulated in T30 treatment ($P<0.05$). Thus, the treatment with trehalose resulted an increased expression of some of the defense related genes under T30 and T60 treatments in the S lines (Fig. 5.7A, B). On the contrary, it is evident from the results observed in T90 and T120 treatment that the expression of these defense-related genes is statistically similar to the expression observed in the T0 treatment; however, expression of all six plant defense genes, on average, was lower in the T120 treatment as compared to the T30 and T60 treatments.

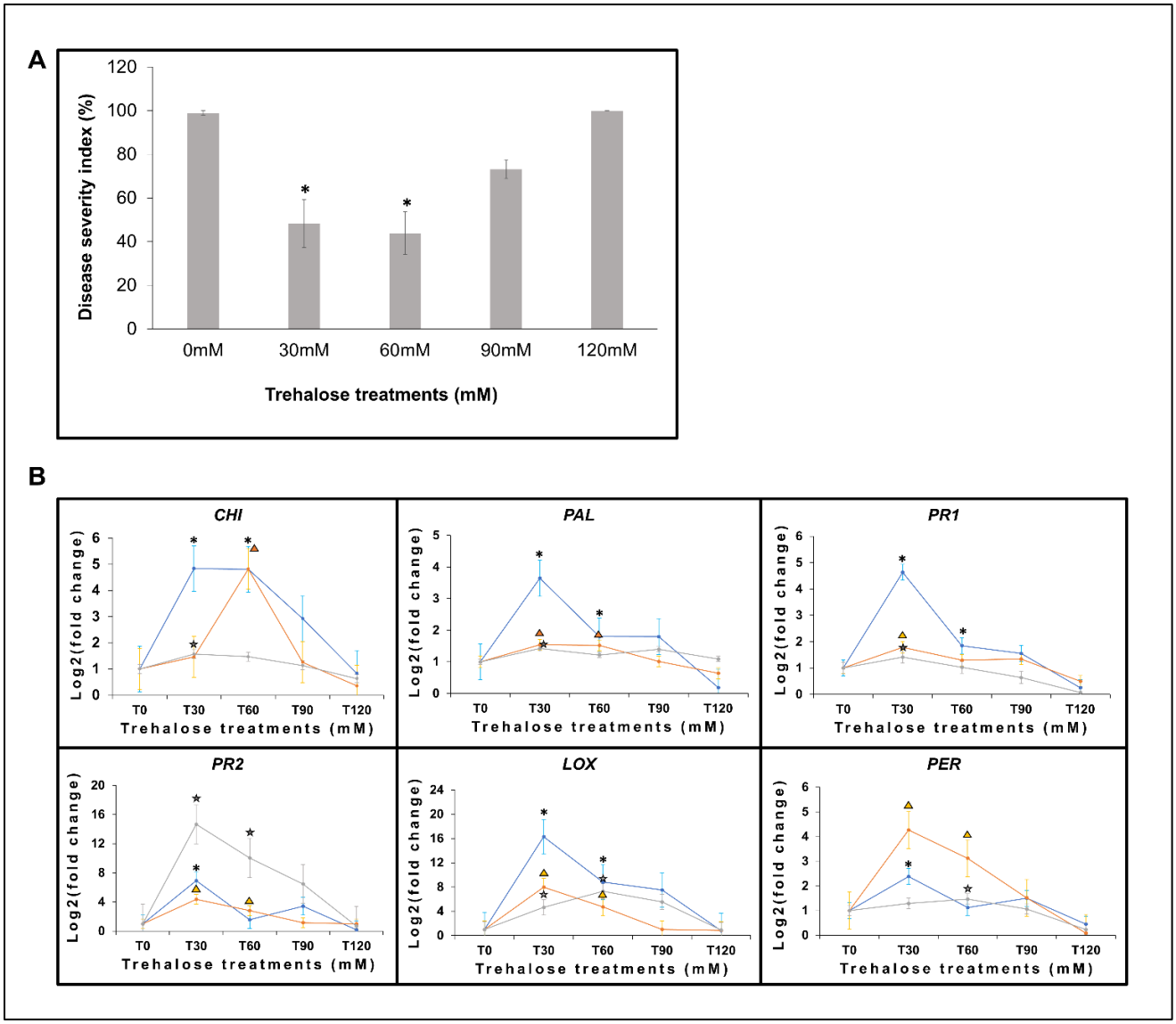


Fig. 5.7 Results of the clubroot susceptible doubled haploid (DH) *Brassica napus* lines treated with trehalose at concentrations of 0 (T0), 30 (T30), 60 (T60), 90 (T90) and 120 (T120) mM. A) Disease severity index (DSI %); pooled data of the seven DH lines presented. Statistical significance as compared to T0 was calculated using *t*-test and is indicated by asterisk. B) Expression analysis of chitinase (*CHI*), pathogenesis-related 1(*PR1*), pathogenesis-related 2 (*PR2*), phenylalanine ammonia lyase (*PAL*), lipoxygenase (*LOX*) and peroxidase (*PER*) in three susceptible DH lines (S1, blue; S2, orange; and S8, grey) treated with trehalose of the above-mentioned five concentrations. [Statistical significance was calculated using *t*-test; asterisks indicate the T30 and T60 are statistically different from T0 in S1; triangle indicate the T30 and T60 are statistically different from T0 in S2; and star indicate the T30 and T60 are statistically different from T0 in S3.]

5.3.8 Development of SSR markers from the DEGs of BnaA08 (RtR) and BnaA03 (MdR)

To develop gene-based molecular markers associated with resistance, we screened all the 84 genes from BnaA08 upregulated (<-1) in RtR for the presence of SSRs. Of the 84 DEGs, 15 contained a total of 17 SSRs; these 15 DEGs located at the following positions: five located at 157,029 – 6,790,400 bp, two at 10,299,331 – 10,318,442 bp, seven at 12,256,423 – 14,314,535 bp and one located at BnaA08_random 1,822,640 – 1,824,044 bp. A total of 24 SSR primer pairs were designed based on the 17 SSRs and were tested for polymorphism between the resistant and susceptible parents of the Rt mapping population. Among these, the SSR marker designed from the DEG BnaA08g03250D based on (AGC)5 sequence repeat exhibited polymorphism. Genotyping of the 81 DH lines of this population with this marker showed an association with clubroot resistance; however, 2.4% recombination occurred between the markers and resistance (Supplementary Table 5.8). Representative genotyping results are shown in Fig. 5.8A.

In case of Mendel-resistance, 119 DEGs from BnaA03 were upregulated (<-1) in the MdR lines; however, nine DEGs found to contain SSRs. A total of nine primer pairs were designed from the nine DEGs; these DEGs located at the following positions: three located at 12,392,309 – 15,561,056 bp, four at 18,474,534 – 22,509,061 bp and two located at 28,491,805 – 28,617,965 bp. All 56 markers, which included 47 markers designed from flanking sequences of 34 DEGs and nine designed from the nine DEGs, were tested for polymorphism between the resistant and susceptible parents of the two mapping populations #1330 and #1333. Among these, two SSR markers designed from the DEGs BnaA03g41300D and BnaA03g44400D, based on the sequence repeats (TCC)6 and (CAT)5, respectively, exhibited polymorphism. Genotyping of the 96 DH lines of the two mapping populations showed an association of these markers with

resistance. However, the marker from BnaA03g41300D showed about 3.12% recombination, while the marker from BnaA03g44400D showed about 1.04% recombination **(**Supplementary Table 5.9). Representative chromatograms from ABI DNA analyzer are shown in Fig. 5.8 B & C**.** In case of the 47 primer pairs designed from the flanking sequences of 34 DEGs upregulated in MdR, none of these show polymorphisms between the resistant and susceptible parents.

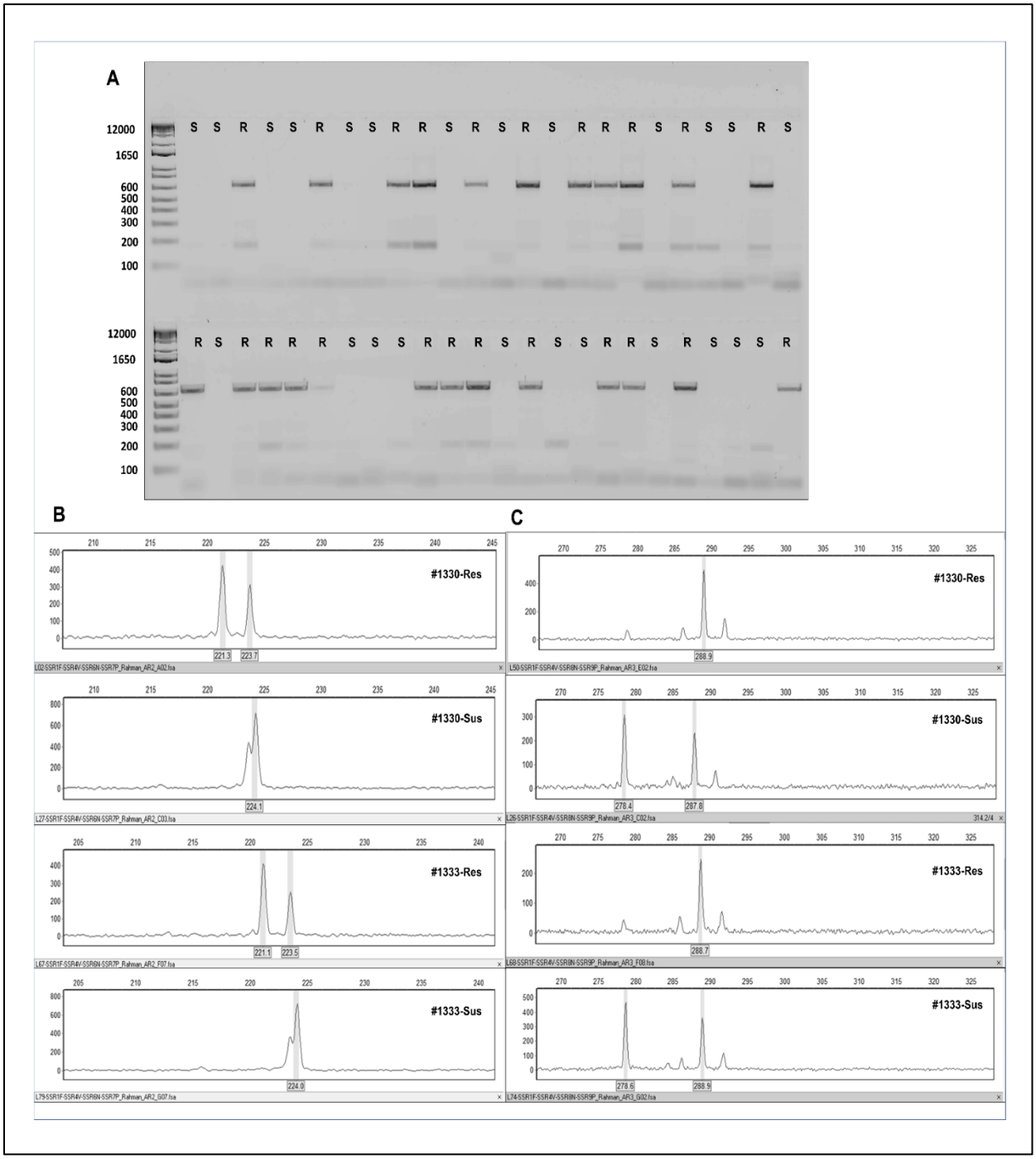


Fig. 5.8 Representative genotyping results of the gene-based SSR markers. A) Gel images depicting genotyping of *Brassica napus* doubled haploid population, carrying Rutabaga resistance, by the SSR marker designed from the BnaA08g03250D of BnaA08. B) ABI DNA analyzer electropherogram peaks depicting genotyping of two *B. napus* doubled haploid populations (#1330 and 1333) by the SSR marker designed from the BnaA03g41300D of BnaA03. C) ABI DNA analyzer electropherogram peaks depicting genotyping of two *B. napus* doubled haploid populations (#1330 and 1333) by the SSR marker designed from the BnaA03g44400D of BnaA03. [B & C) x-axis indicate the size of DNA fragments in base pairs, while the y-axis indicates the intensity of the fluorescent signal]

5.4 Discussion

RNA-seq has been implemented by several researchers mostly by using single clubroot resistance gene to study the transcriptomic changes occur during interaction between the *Brassica* host and *P. brassicae* and to identify the putative candidate genes (Chen et al. 2016, Irani et al. 2018, Huang et al. 2019, Li et al. 2020). In this study, we conducted a comparative analysis of the DEGs upregulated in the lines carrying rutabaga-resistance located on chromosome BnaA08 and the DEGs upregulated in the lines carrying Mendel-resistance located on BnaA03 to identify the candidate genes to be involved in clubroot resistance, as well as to identify the genes which might play a fundamental role in resistance and, thus, to extend our knowledge of the molecular-genetic basis of resistance. To our knowledge, such study has so far not been conducted in canola. We also used these DEGs for the development of molecular markers for use in breeding. Furthermore, based on the identification of the primary and secondary metabolic pathways upregulated in the resistant lines, we studied the importance of these pathways for clubroot resistance in susceptible plants after priming with carbohydrates, like trehalose.

Genes encoding different transcription factors, disease resistance proteins, peroxidases, heat shock proteins, cell wall proteins as well as those involved in different biosynthetic pathways such as primary metabolism, secondary metabolism and plant hormone signalling showed differential expression in the RtR and MdR lines after *P. brassicae* infection and, therefore, might be involved in conferring resistance to clubroot disease. Plants have the ability to recognize and respond to pathogen attacks by deploying defense responses (for review, see Bolton 2009). These defense responses result in an increased energy demand, including for carbon skeletons and amino acids, which are the products from primary metabolic pathways like the glycolytic pathway, PPP, sucrose metabolism and TCA-cycle, and used by the secondary

metabolic pathways like the phenylpropanoid pathway (for review, see Bolton 2009). These primary and secondary metabolic pathways are known to respond to pathogen infection (Danson et al. 2000, Mutuku and Nose 2010, 2012, Hren et al. 2009) and were also found to be upregulated in the RtR and MdR lines used in this study in response to *P. brassicae* infection at 10 dpi. The building blocks used in the biosynthesis of plant secondary metabolites (SMs) like nitrogen/sulphur containing compounds, phenolic compounds and terpenoids (for review, see (Croteau et al. 2000, Ncube and Staden 2015) are derived from acetyl-coA (acetate pathway), shikimic acid (shikimate pathway), mevalonic acid (MVA pathway) and 1-deoxyxylulose-5-phosphate (MEP pathway) (Fig. 5.1).

Plant SMs are synthesized from the modified aromatic amino acids via the shikimate pathway (for review, see Vogt 2010, Parthasarathy et al. 2018). For instance, 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase (*DAHP*) catalyzes the first step of the shikimate pathway (for review, see Santos-Sánchez et al. 2019). Subsequently, reactions in the phenylpropanoid pathway are catalyzed by the enzymes encoded by phenylalanine ammonia-lyase (*PAL),* peroxidase *(PER)*, 4-coumaroyl CoA-Ligase (*4CL*), trans-cinnamate 4-monooxygenase, cinnamyl alcohol dehydrogenase, and caffeic acid 3-O-methyltransferase gene families (for review, see Caretto et al. 2015, Tohge et al. 2013). *PAL* and *4CL* are important intermediate enzymes responsible for the synthesis of stilbenes, flavones, flavonols, lignins, tannins, coumarins and anthocyanins (Zhang et al. 2017; for review, see Vogt 2010, Parthasarathy et al. 2018). *PER* catalyzes the redox reaction of various substrates resulting in the lignification and cross-linking of cell wall proteins (Hilaire et al. 2001, Govind et al. 2016). Upregulation of these genes in RtR and MdR lines suggests that increased lignification of the cell wall may be important to prevent or slow down pathogen invasion and spread (Lahlali et al. 2017, Badstöber et al. 2020) for resistance to *P. brassicae*. Furthermore, plant SMs like terpenes, generated by the MVA and MEP pathways, produce cytosolic and plastidial terpenes, respectively (Bick and Lange et al. 2003, Laule et al. 2003) (Fig. 5.1**)**. The 3-hydroxy-3-methylglutaryl-coenzyme A reductase (*HMGR*) from the MVA pathway and geranylgeranyl pyrophosphate synthase (*GGPPS*) from the MEP pathway showed some degree of induction in the resistant plants at 10 dpi. Previously, overexpression of *HMGR* (Oba et al. 1985, Kondo et al. 2003, Wang et al. 2012) and *GGPPS* (Henriquez et al. 2016, Takemoto et al. 2018) have been reported to enhance biotic stress tolerance in *Solanaceae* plants. In the current study, induction of

these genes due to infection by *P. brassicae* suggests their role in clubroot resistance in Brassica as well. Thus, the concomitant increase in the expression of the genes involved in carbon metabolism, acetate pathway, shikimate pathway, MVA, MEP and phenylpropanoid pathways suggests that altered carbon metabolism plays an important role in mediating the resistance to *P. brassicae* in the RtR and MdR lines.

The modulation of primary metabolism in plants also influences the SA and JA responses during the defense responses against the pathogens (Lemarié et al. 2015). Isochorismate from the shikimate pathway is a precursor of SA and is important for the induction of local and systemic acquired resistance (Wildermuth et al. 2001, Torrens-Spence et al. 2019, Rekhter et al. 2019). Likewise, JA is an important phytohormone, produced by the lipoxygenase pathway in which the acetyl-coA from the glycolytic pathway serves as the precursor (for review, see Dudareva et al. 2013). It is also well known that both SA and JA are responsible for the induction of PR genes (Lemarié et al. 2015, Mei et al. 2006); for review, see (Okada et al. 2014, Ruan 2019) (Fig. 5.1). In this study, we observed SA pathway activation in both RtR and MdR lines at 10 dpi, providing support to the previous studies Siemens et al. 2006, Lemarié et al. 2015, Ning et al. 2019). We also observed the lipoxygenase (*LOX)* and most of the genes of JA pathway like jasmonate ZIM domain-containing protein (*JAZ*) and allene oxide cyclase (*AOC*) (for review, see (Wasternack and Song 2017) were expressed only in the MdR lines indicating that JA may play a role along with SA in mediating clubroot resistance in MdR (Lemarié et al. 2015, Fu et al. 2019, Liu et al. 2020). Thus, it appears that the induction of the specific genes involved in SA and JA pathways is different in clubroot resistant *B. napus* lines for mediating resistance to *P. brassicae* which is also evident from other studies (Lemarié et al. 2015). However, this needs to be verified through mutational analyses of the genes of SA and JA pathway as well as analysis of the levels of endogenous SA and JA in clubroot resistant lines in response to challenge by the pathogen.

Primary metabolism can be modulated to induce resistance mechanisms in plants before the impending pathogen attack. This process of priming defense responses to various pathogens leads to an increased expression of the genes involved in mediating resistance (for review, see Bolton 2009) and ensures that energy costs associated with plant defense mechanisms are efficiently managed (for review, see Bolton 2009). In plants, such modulation is achieved by priming these pathways with small carbohydrates, which have been suggested to play multiple roles. These include, the biosynthesis of plant structural compounds, acting as signalling

molecules that regulate the plant immune system, enhancing the oxidative burst during the early stages of infection, increasing the lignification of cell walls, inducing the genes encoding PR proteins, and the stimulation of flavonoid biosynthesis for plant defense responses (Gómez-Ariza et al. (2007); for review, see (Morkunas and Ratajczak 2014). Different carbohydrates, including sucrose, trehalose, raffinose, α-D-glucopyranosyl-(1–1)- α–D-glucopyranoside and galactinol can activate plant defense responses for resistance towards various pathogens (Trouvelot et al. 2014, Singh et al. 2011, Lyu et al. 2013, Delorge et al. 2014, Mostofa et al. 2015). For example, it has been demonstrated that treatment of rice plants with sucrose induces the accumulation of *PR1a, PR1b* and *PR5* in leaves resulting in an enhanced resistance to the fungus *Magnaporthe oryzae* (Gómez-Ariza et al. 2007). Furthermore, this study also showed that the plants can regulate their sucrose level in response to pathogen infection and the altered sucrose level can act as endogenous molecular signal to activate the defense response (Gómez-Ariza et al. 2007). Consistent with the aforementioned studies, our study also suggests that carbohydrates play a role in the elicitation of plant defense responses against pathogens.

Trehalose is known to have a robust effect on carbon metabolism in plants. Previous studies have shown that *P. brassicae* increases the accumulation of sugars such as glucose, fructose, and trehalose in infected plant roots (Brodmann et al. 2002, Bae et al. 2005). Li et al. (2018) also observed a significant increase in sugar and fructose contents as well as upregulation of *SWEET* genes in clubroot-susceptible *B. rapa* as compared to clubroot-resistant *B. rapa*. They suggested that increased sugar translocation and *P. brassicae* infection is closely related as *P. brassicae* triggers the expression of the *SWEET* genes causing sugar translocation from the sugar producing source organs to the gall tissue. This suggests that trehalose can act as a virulence factor by altering carbohydrate metabolism (Brodmann et al. 2002, Bae et al. 2005). However, in this study, an increased resistance in the clubroot susceptible *B. napus* lines, inoculated with *P. brassiae*, was observed due to exogenous application of trehalose (Fig. 5.7A) indicating trehalose might play a role in elicitation of plant defense response as well. Enhanced resistance of plants to biotic stress following exogenous application of trehalose has also been reported by other researchers, for instance in *Arabidopsis* against green peach aphid and in wheat against *Blumeria graminis* (Singh et al. 2011, Reignault et al. 2001, Tayeh et al. 2014). Our study further indicated the upregulation of defense related genes like *PR1, PR2, LOX* and *PAL* in the susceptible DH lines in response to trehalose (Fig. 5.7B). Similar results have also been reported

by Tayeh et al. (2014) in wheat where an increased expression of *PR1, LOX* and *PAL* was found after priming with trehalose. Therefore, in our study, the seed treatment using lower concentrations of trehalose might have functioned as an elicitor resulting in the activation of plant defenses in canola before the pathogen attack. However, higher concentrations of trehalose such as 90 mM and 120 mM resulted in 73-100% clubroot disease. This decrease in resistance could also be correlated to the decrease in expression of the defense related genes. Indeed, a correlation was observed between the expression trends of *PR1, LOX* and *PAL* genes and the partial resistance in the T30 and T60 treatment plants and increased susceptibility in the T90 and T120 treatment plants. On the contrary, Govind et al. (2016) had shown induction of *PAL* and *PER* enzyme activity in pearl millet treated with trehalose concentration of 200 mM and later challenged by *Sclerospora graminicola*. The inconsistency in our study from Govind et al. (2016) could be due to the differences in the duration and method of application of trehalose and the plant species and the pathogen. Previous studies with *Arabidopsis* seedlings have also showed that trehalose at a higher concentrations, such as 100 mM, interfere with carbon allocation by increasing starch in source organs, and this result in insufficient metabolizable carbon available for root growth (Wingler et al. 2000, Schluepmann et al. 2004). The trehalose toxicity had also been linked to starch and anthocyanin accumulation and root growth inhibition (Wingler et al. 2000, Schluepmann et al. 2004). Thus, it could be speculated that the insufficient amount of metabolizable sugar must have affected the energy status of the plants (for review, see Morkunas et al. 2012), affecting the gene expression and enzyme activities. Therefore, the plants treated with higher concentrations of trehalose showed a similar expression of the defense related genes as the pattern observed in T0 plants. Thus, the knowledge from previous studies and our results suggests that modulation of primary and secondary metabolic pathways as well as induction of defense related genes by carbohydrates are capable of eliciting defense response, and such responses might be critical for clubroot resistance as observed in this study. Additionally, it is imperative to study the response of *B. napus* to other sugars, such as glucose, fructose and raffinose and other osmotic adjusters such as PEG, to understand among the different sugars if trehalose is specifically capable of eliciting clubroot resistance in susceptible canola.

5.4.1 Comparison between BnaA08 (RtR) and BnaA03 (MdR) resistance

The clubroot resistance in the RtR line was mapped to BnaA08 (Fredua-Agyeman and Rahman 2016) and in MdR, it was mapped to BnaA03 (Hatakeyama et al. 2013). This is the first study to analyze and show comparisons between the Mendel (BnaA08; MdR)- and rutabaga (BnaA03; RtR)-resistances influencing clubroot resistance in *B. napus*. The chromosome BnaA08 of the *Brassica* A genome has been reported to carry a major clubroot resistance locus *Crr1* (Suwabe et al. 2003, 2006), which was later found to include two genes, *Crr1a* and *Crr1b* (Hatakeyama et al. 2013). We searched the genomic region of *Crr1a* (GenBank accession no. AB605024; (Hatakeyama et al. 2013)) in the *B. napus* genome (Genoscope *B. napus* Genome Browser, http://www.genoscope.cns.fr/brassicanapus/; retrieved on 5 May, 2020) and positioned this gene at BnaA08: 9,455,911 – 9,461,175 bp. In our study, the most upregulated gene from BnaA08 was the BnaA08g20170D located at chrA08:15,299,746 –15,304,559 bp; this encodes for a disease resistance RPP5-like protein which is a TIR-NB-LRR type R protein (Parker et al. 1997), and these proteins have been implicated to play a role in disease resistance ((Gassmann et al. 1999, Yi and Richards 2007); for review, see Marone et al. 2013). BnaA08g20170D positioned at ~ 6 Mb away from *Crr1a*; however, both loci encode TIR-NBS-LRR type proteins which have been known to confer clubroot resistance in *B. rapa* (Hatakeyama et al. 2013). Based on this, it can be anticipated that the BnaA08g20170D is *Crr1b*, the minor gene contributing resistance at 10 dpi. Expression of this gene was specific to the RtR lines, and therefore, it is highly likely that this gene is an important contributor for resistance to *P. brassicae*.

In case of the chromosome BnaA03, the clubroot resistance locus *CRa* (GenBank accession no. AB751516; (Ueno et al. 2012)) is located at 22,862,712 –22,873,697 bp of the *B. napus* genome (Genoscope *B. napus* Genome Browser, http://www.genoscope.cns.fr/brassicanapus/; retrieved on 5 May, 2020). The two TIR-NBS-LRR type genes, BnaA03g46200D (BnaA03:23,666,076 – 23,668,891 bp) and BnaA03g48510D (BnaA03:24,885,706 – 24,888,274 bp), which were upregulated only in the MdR lines were found to be located near *CRa*. Expression of these two genes specifically in the DH lines of #1333 indicates that they might play an important role in resistance to clubroot disease in the MdR lines. Among the different genes of BnaA03, the gene aspartic proteinase-like protein 1(*AP*) (BnaA03g39150D: BnaA03:19,505,783 – 19,506,967 bp) showed the greatest level of expression in the MdR lines. Previously, Xia et al. (2004) reported that an *AP* produces a peptide

elicitor which is involved in resistance to *Pseudomonas syringae* in *Arabidopsis*. Based on the level of upregulation that we found for BnaA03g39150D in the MdR lines and the involvement of this gene in disease resistance, it is possible that this gene play a role in clubroot resistance in the MdR lines; however, functional validation of this gene will be needed.

We also identified the orthologous gene *CRF4* which was upregulated in both the RtR (BnaA08) and MdR (BnaA03) lines. *CRF4* belongs to the subfamily of ERF transcription factors and is a component of the cytokinin signaling pathway (Rashotte et al. 2006) and is positioned ~ 2.4Mb from the *Crr1a* of BnaA08 (RtR) and ~2.2Mb from the *CRa* of BnaA03 (MdR). The cytokinin signaling pathway has been deemed to play an important role in curbing the clubroot disease development (Siemens et al. 2006, Ciaghi et al. 2019). Kwon et al. (2016) suggested that *CRF2* enhances the production of SA resulting an autoimmune response. Overexpression of *CRF5* in *Arabidopsis* has also been found to enhance resistance to *Pseudomonas syringae* through activation of a large number of GCC-box pathogenesis related genes (Liang et al. 2010). Given the fact that *CRF4* was expressed in both types of resistant lines and is positioned closed to the genomic region known to confer clubroot resistance; it is highly likely that this gene plays a fundamental role in clubroot disease resistance; functional study with this gene will be needed to delineate its exact role.

5.4.2 Development of gene-based molecular markers

Molecular markers for a trait developed from the sequence of the genes are called functional markers (Lau et al. 2015). These types of markers are expected to show zero recombination and, therefore, can be used in molecular breeding with greatest efficiency. Several researchers have reported different types of gene-based markers in different plants, for instance, EST-SSR marker in *Morinda officinalis* (Liao et al. 2019), gene-tagged SSR markers in wheat and rice (Singh et al. 2018, Molla et al. 2019), inter-retroelement amplified polymorphism (IRAP) markers in *Crocus sativus* (Alsayied et al. 2015), and miRNA-SSR markers in Brassica (Fu et al. 2013) and rice (Ganie and Mondal 2015). In the present study, we attempted to develop SSR markers from the exons of the DEGs expressed from BnaA08 and BnaA03 and are upregulated, respectively, in the RtR and MdR lines and are involved in primary metabolism and biotic stress response pathway. However, the SSR markers from these DEGs (BnaA08g03250D of BnaA08 of RtR, and BnaA03g41300D and BnaA03g44400D of BnaA03 of MdR) showed about 1.04 to 3.12% recombination with the resistance phenotype. The DEG BnaA08g03250D of

BnaA08 positioned at 2,668,129 –2,671,679 bp of the *B. napus* reference genome (Chalhoub et al. 2014) which is about ~6.7 Mb upstream from the locus *Crr1a* (GenBank accession no. AB605024; (Hatakeyama et al. 2013)). This gene encodes a 3,4-dihydroxy-2-butanone kinase and is known to be involved in the riboflavin pathway (Herz et al. 2002, Buffon et al. 2016), and riboflavin metabolism has been suggested to be involved in plant disease resistance (Dong and Beer 2000, Zhang et al. 2009). However, the occurrence of 1.8% recombination between the marker and the resistance phenotypes suggests that this gene might not be the candidate gene for clubroot resistance.

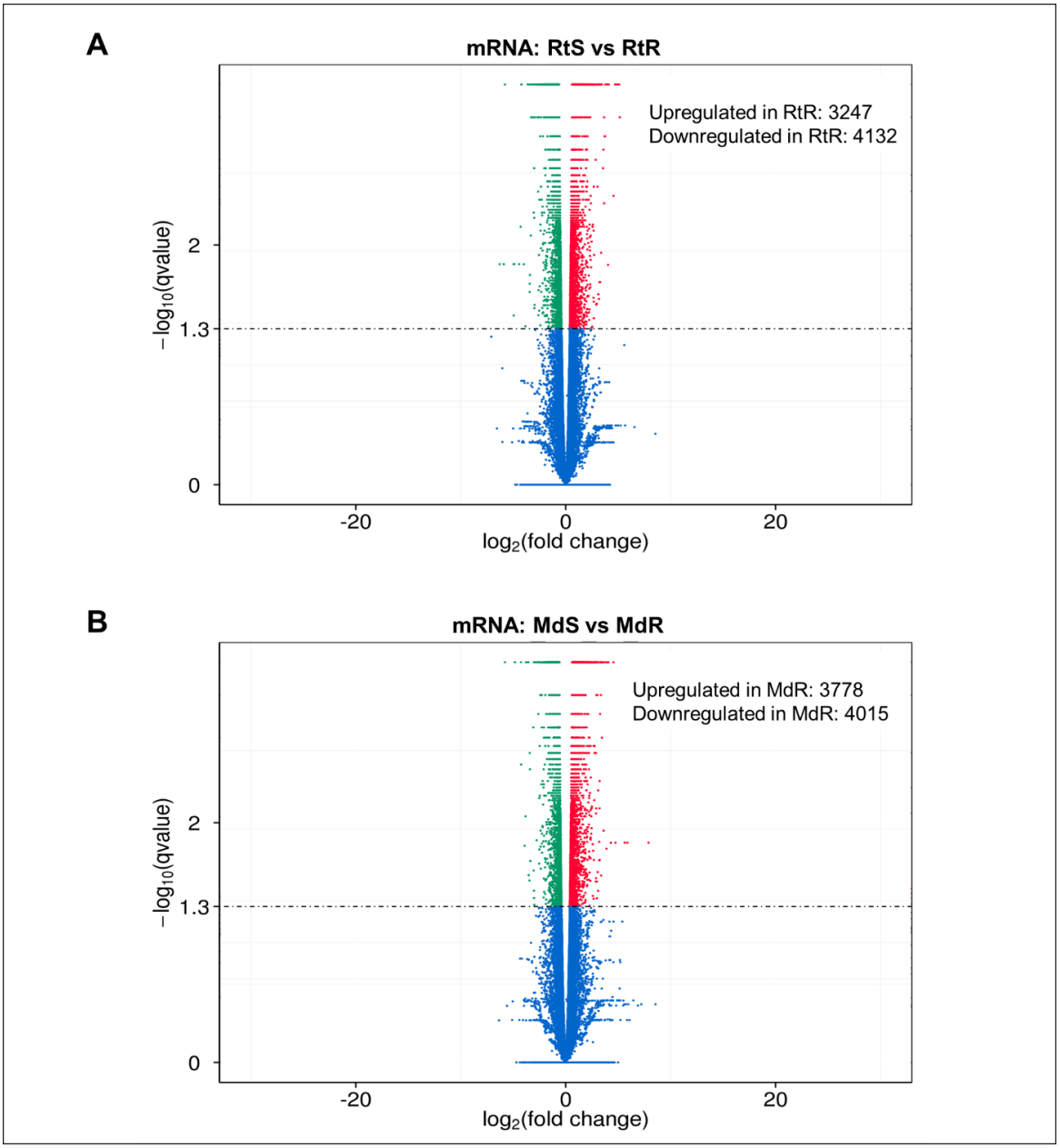
Similarly, we developed two SSR markers from BnaA03g41300D and BnaA03g44400D. These two genes have been predicted to be involved in plant-pathogen interaction via protein degradation pathway as annotated by using Mapman v3.5.1 (Supplementary Table 5.7), and the protein degradation pathway has been considered important in many signalling systems including defense (Marino et al. 2012). However, the SSR marker of the gene BnaA03g41300D, which is located at 20,680,546 – 20,683,271 bp, and BnaA03g44400D, which is located at 22,506,155 – 22,509,061 bp, of BnaA03 showed 3.12% and 1.04% recombination, and these two genes are located about ~2.1 Mb and 0.35 Mb away from the *CRa* (GenBank accession no. AB751516; (Ueno et al. 2012)) locus of BnaA03, respectively. The gene BnaA03g41300D encodes a gene aspartic proteinase-like protein (*AP*), which has been known to be involved in biotic resistance (Xia et al. 2004) and is located close to another *AP*-coding gene (BnaA03g39150D), which was also strongly expressed in the MdR lines (Supplementary Table 5.7). The other gene of BnaA03, BnaA03g44400D, encoding allantoate deiminase has been known to be involved in purine catabolism for the purpose of recycling carbon and nitrogen and is, therefore, important in the synthesis of primary and secondary metabolites (Todd and Polacco 2006); for review, see (Stasolla et al. 2003, Zrenner et al. 2006) which have been reported to play a role in plant disease resistance (for review, see (Bolton 2009)). However, the lack of perfect association of these markers with the resistance phenotype indicate that they might not be the candidate genes for clubroot disease resistance. Given that fact that these three genes were significantly upregulated in the resistant lines and their reported role in plant defense, it is possible that they are somehow involved in resistance to *P. brassicae*; however, functional validation of these genes would be needed. Nevertheless, the three SSR markers that we

developed based on the three DEGs for the two clubroot resistance loci can be used in marker assisted breeding with an efficiency of about 97-99%.

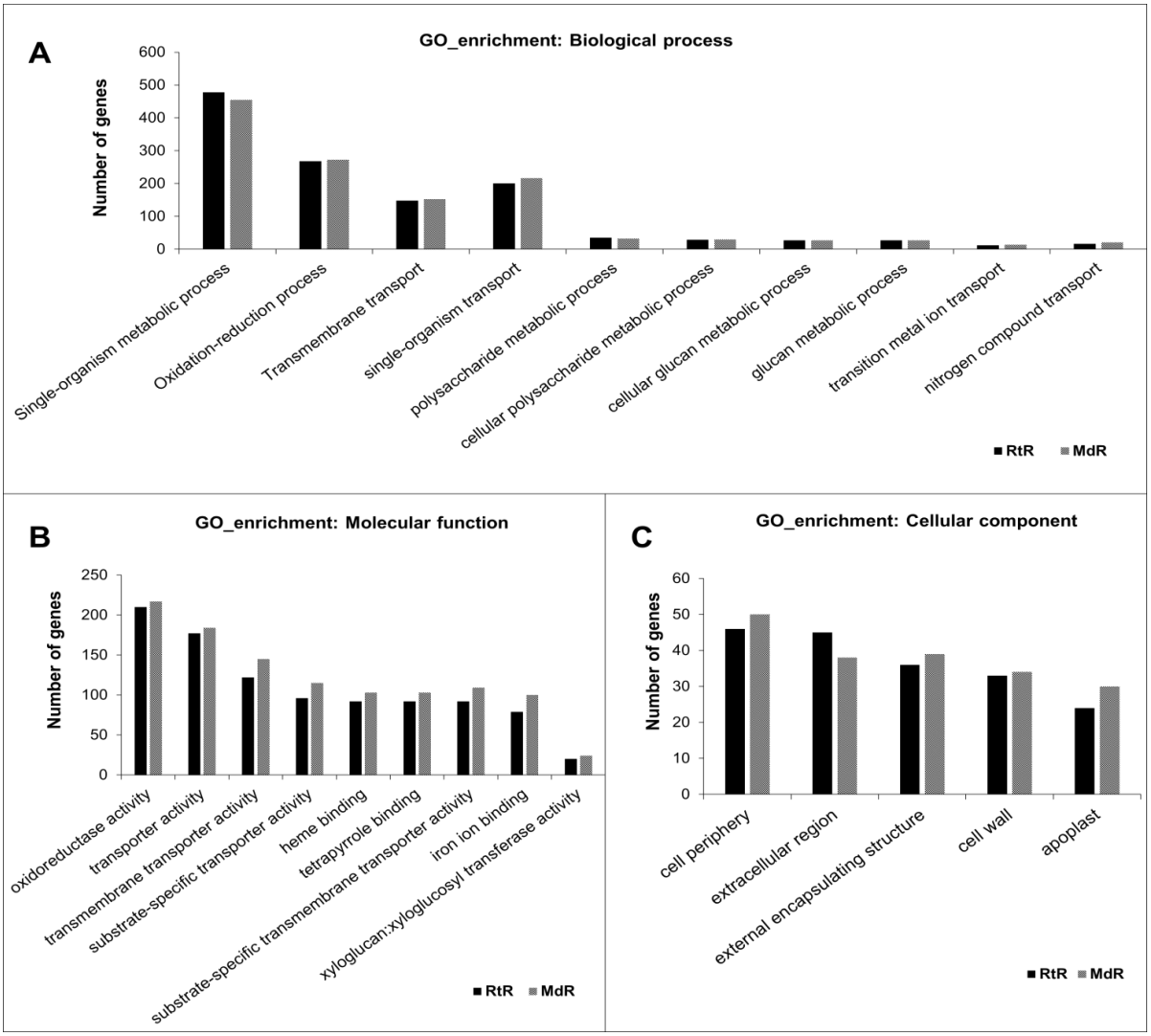
5.5 Conclusion

We have identified ~3500 DEGs upregulated in RtR and MdR canola lines challenged with *P. brassicae* pathotype 3. Approximately, 8-9% genes were upregulated in RtR and MdR lines and were involved in primary and secondary metabolic pathways such as carbohydrate metabolism, amino acid metabolism, TCA-cycle, MVA and MEP pathway, and phenylpropanoid pathway. The importance of primary metabolism in modulating clubroot resistance was further evidenced by priming canola plants with small carbohydrates like trehalose to elicit early plant defense. Furthermore, on comparing the BnaA08 (RtR) and BnaA03 (MdR), we found the gene *CRF4* was upregulated in both RtR and MdR lines and located ~2.3Mb away from the *Crr1a* (BnaA08 of RtR) and *CRa* (BnaA03 of MdR). We also developed SSR markers from the DEGs involved in primary and secondary metabolic pathways and upregulated in the RtR and MdR lines. Thus, the results from this study provided not only the candidate genes to be involved in clubroot resistance, but also extended our knowledge of the molecular-genetics of resistance, which could be important for further understanding of the mechanism of clubroot resistance. In addition, the gene-based molecular markers developed in this study could be used in the breeding of canola.

Supplementary Figures



Supplementary Fig 5.1 Volcanic plot showing differential expression of mRNAs in roots of *Brassica napus* plants at 10 days after inoculation with *Plasmodiophora brassicae* pathotype 3: (A) Expression in susceptible lines lacking rutabaga-resistance (RtS) and resistant lines carrying rutabaga-resistance (RtR), and, B) Expression in susceptible lines lacking Mendel-resistance (MdR) and resistant lines carrying Mendel-resistance (MdR). Threshold q-value <0.05 indicates significant differences in expression.



Supplementary Fig 5.2 Gene Ontology enrichment analysis of the genes upregulated in RtR and MdR lines of *Brassica napus*. A) Common biological process related terms, B) Common molecular function related terms, and C) Common cellular component related terms. RtR and MdR indicates the lines carrying clubroot resistance introgressed from rutabaga and winter canola cv. Mendel.

Chapter 6: General Discussion

The progress in Next Generation Sequencing (NGS) has fast-tracked gene discovery and quantitative trait loci (QTL)- mapping of disease resistance traits in economically important crops (Nguyen et al. 2019). NGS has enabled the identification of QTLs using methods such as whole genome resequencing, high-density SNP arrays for faster mapping, bulk segregant RNA-seq for gene discovery, QTL-seq for gene identification (Jaganathan et al. 2020). Among these, RNA-seq can be used in analysis of differentially expressed genes (DEG) across different developmental stages, time points and under multiple conditions at a high-resolution, is one of the most common application of the RNA-seq. RNA-seq provides a global view of the protein coding (~< 2%) and non-coding genes (~> 98%) of the genome (Wang et al. 2009, Wilusz et al. 2009). Differentially expressed genes (DEGs) from RNA-seq experiment, are the genes whose expression levels are considered significantly different across two or more conditions (McDermaid et al. 2018). Such an analysis can provide considerable understanding into the molecular mechanisms of a plant, which contributes towards its phenotypic differences including plant growth and developmental patterns, biotic and abiotic stress responses. Such a study also aids in the identification of gene candidates which could be used as potential candidates for crop improvement or for the development of markers, for their use in marker- assisted breeding. These markers within genes may have an effect on gene expression, thus playing an important role in contributing to the differential expression profiles of that gene (Lawson and Zhang 2006). The absence of recombination and its complete linkage to the desired phenotype would prevent the loss and false selection in markers assisted breeding, thus making it an effective marker. On the contrary, random polymorphic markers increase the chances of transferring unwanted genes along with the target gene, due to its linkage with undesirable traits (Amom and Nongdam 2017). Thus, the development of such gene-based markers for genes controlling clubroot resistance will facilitate the introgression of disease resistance into elite cultivars, for the development of durable disease resistance.

The overall broad research objectives presented in this thesis aided the identification of clubroot-responsive DE lncRNAs and mRNAs from spring canola carrying resistance introgressed from rutabaga (RtR) (*B. napus* var. *napobrassica*) as well as from spring canola carrying resistance introgressed from winter canola cv. ‘Mendel’ (MdR). The knowledge encouraged the progress in understanding the molecular basis of their resistance to clubroot and

helped in understanding the differences in responses in RtR and MdR plants against *P. brassicae.* This knowledge also facilitated the identification of candidate genes for future functional studies as well as the development of RNA- and lncRNA-based markers for marker assisted breeding.

The lncRNA response to biotic stress has been observed in many plants such as wheat, canola and rice against diseases like powdery mildew, white mold, and rice blast, respectively. However, considering the economic importance of canola in Canada and the economic losses due to clubroot, the study of lncRNAome of canola against clubroot had not been attempted previously. LncRNAs are the emerging source of potential targets which can be used for crop protection (Summanwar et al. 2020 and references there in). The DE lncRNAs discussed in this dissertation allowed us to recognize potential candidates for future functional studies involved in developing clubroot resistant canola cultivars. For instance, we found LNC_003848 positioned at BnaC08 and trans regulated a target gene Transparent Testa 12 located on BnaA08. We observed the gene was positioned ~2.7 Mb away from the *Crr1a* locus. According to (Hasan and Rahman 2016), the *Crr1a* locus is responsible for conferring resistance to multiple *P. brassicae* pathotypes. The distance of Transparent Testa 12 from the *Crr1a* locus, suggests that LNC_003848 could be regulating important gene/s which are critical in mediating resistance against clubroot by an unknown mechanism.

Furthermore, considering the importance of the clubroot resistance loci of BnaA03 and their association with the clubroot resistance trait in canola (Fredua-Agyeman and Rahman 2016), we had focused on developing the SSR markers from the DE lncRNAs upregulated and located on BnaA03. We observed the LNC_000424 containing a compound SSR was showing a strong association with the clubroot resistance trait. This marker showed ~1.9% recombination and therefore further functional study will aid in understanding its role in mediating clubroot resistance. Moreover, this lncRNA regulates defense related genes such as *PR1* (Breen et al. 2017), WRKY transcription factor 19 (Phukan et al. 2016) and MYB transcription factor 44 (Ambawat et al. 2013), serine/threonine protein kinase *BIK1* (Lu et al. 2010), and calcium binding protein *CML21* (Aldon et al. 2018) and disease resistance protein coding genes such as *RPS4, CSA1* (Faigón-Soverna et al. 2006; Gassmann et al. 1999); and therefore, offers to be a good candidate for further investigation.

The clubroot resistance from the winter canola cultivar ‘Mendel’ and rutabaga has been important because of their response to multiple pathotypes (Hasan and Rahman 2016, Fredua-

Agyeman and Rahman 2016). The winter canola 'Mendel' and rutabaga are important sources of clubroot resistance (Hasan and Rahman 2016; Fredua-Agyeman and Rahman 2016), there are no studies comparing the lncRNAome of these cultivars as well as the resistance loci in response to clubroot. Therefore, the comparison of the DE lncRNAs upregulated in both RtR and MdR lines further advanced our knowledge about the common mechanisms in play against clubroot disease. These common DE lncRNAs were predicted to regulate same target genes involved in plant defense pathways, in both RtR and MdR lines genotypes. This also suggests that both DE lncRNAs and their respective target genes played a role in decelerating the disease progression in both the resistant plants at 10 dpi. For instance, LNC_000496 is predicted to regulate *BAK1* and *RIN4*, genes involved in plant pathogen interaction (Belkhadir et al. 2004; Chinchilla et al. 2007), *LTP*, a gene encoding protein belonging to PR-14 family (Van Loon and Van Strien 1999) and a NBS-LRR type disease resistant protein. This lncRNA like others is predicted to co-regulate multiple genes involved in plant defense in both genotypes. An induction of expression of such lncRNAs in both, RtR and MdR lines, suggests that they are potentially relevant for the clubroot resistance phenotype in both genotypes. Thus, a further study observing the expression of these lncRNAs from the onset of the disease until the gall development stage, will provide us with an insight into the importance of its expression for establishing clubroot resistance. Further, functional studies will explain the mechanism in play for establishing the resistance against clubroot in canola.

LncRNAs execute their functions by different modes of actions, which include their interactions with different RNA and DNA molecules as well as proteins (Summanwar et al. 2020 and references therein). Many lncRNAs sequester miRNAs by RNA-RNA interaction, thus negatively regulating their expression and protecting the original target genes of miRNAs from degradation. For instance, LNC_001163 was upregulated only in RtR lines and was predicted to mimic the gene *AGL16*, a target of the bna-miR824. Studies in abiotic stress (Szaker et al. 2019; Xiang et al. 2019) have connected the negative regulation of *AGL16* by bna-miR824 as being important for mediating responses against abiotic stresses. As opposed to these results, the bna-miR824 seems to be negatively regulated by LNC_001163, resulting in an increase in expression of *AGL16*. In addition, LNC_001163 is positioned ~2.9Mb away from the *Crr1a* locus on BnaA08. The loss/gain-of -function studies can further explain the involvement of this lncRNA in mediating clubroot resistance.

In addition to DE lncRNAs, we also made comparisons between the mRNAs upregulated in RtR and MdR lines. Many studies have focused on clubroot disease development in *Arabidopsis, B. rapa, B. oleracea and B. napus* (Chen et al. 2015; Zhang et al. 2016; Jia et al. 2017; Zhao et al. 2017; Irani et al. 2018; Ji et al. 2018; Jia et al. 2017; Peng et al. 2019; Prerostova et al. 2018; Zhang et al. 2016, Li et al. 2020). Li et al. (2020) suggested that the resistant genotypes sense the pathogen earlier than susceptible genotypes initiating the basal resistance which reaches its maximum expression at 96 hpi. This could have been the reason for not finding many R-genes at 10 dpi, in our study. However, we did not find any R-genes which were commonly expressed between both the genotypes at 10 dpi, which suggests the R-genes are specific to the cultivars, agreeing with the statement of Li et al. (2020), who conjectured that many specific responses of the R-genes revealed gradually. According to Li et al. (2020), the low number of DEGs in resistant lines may result from a basal resistance reaction. However, this research group observed a greater number of DEGs in clubroot resistant lines at later time points. This also suggests that the R- genes from this study are potential candidates which can be used for crop improvement using cisgenesis or gene pyramiding. In addition, on comparing the (BnaA08; RtR) and (BnaA03; MdR) resistance, we found the cytokinin responsive factor (*CRF4)* upregulated from both genotypes. The ~2.3 Mb location of this gene from *Crr1a* on BnaA08 and *CRa* on BnaA03, makes it an ideal candidate for studying its function in clubroot resistance. The cytokinin signaling pathway has been known to play an important role in limiting the clubroot disease development (Ciaghi et al. 2019; Siemens et al. 2006). *CRF2* has been shown to enhance the production of SA as well as activate a large number of GCC-box pathogenesis related genes (Kwon 2016; Liang et al. 2010) resulting in an increased resistance to *Pseudomonas syringae in Arabidopsis.* Given the fact that *CRF4* was expressed in both types of resistant lines and is positioned closed to the genomic region known to confer clubroot resistance; it is highly likely that this gene plays a fundamental role in clubroot disease resistance. Future functional study will be needed to delineate its exact role.

The overall transcriptomic landscape (mRNA and lncRNAs of the RtR and MdR lines), also advanced our knowledge about the important pathways underpinning the clubroot resistance from rutabaga as well as 'Mendel' resistance. Defense responses in plants have been linked to increased energy demands for carbon and amino acids, which are supplied by primary metabolic pathways to secondary metabolic pathways like the phenylpropanoid pathway. In 2010, Mukutu

and Nose demonstrated that *R. solani* infection activated the glycolytic pathway, pentose phosphate pathway (PPP), and tricarboxylic acid cycle **(**TCA)-cycle for the induction of phenylpropanoid pathway. Some of the important gene families of these pathways, which respond to pathogen infection according to previous studies, including glycolytic enzyme like phosphoglycerate kinase (*PGK*); TCA-cycle enzyme like succinate dehydrogenase; PPP enzyme like ribose-5-phosphate isomerase and sucrose metabolism enzymes like sucrose synthase (Danson et al. 2000; Hren et al. 2009; Mutuku and Nose 2012; Mutuku and Nose 2010), were observed to be upregulated in both resistant lines in response to *P. brassicae*. These pathways provide precursors such as phosphoenolpyruvate and erythrose-4-phosphate to the shikimate pathway. The gene encoding enzyme phopho-2-dehydro-3-deoxyheptonate aldolase (*DAHP*) responsible for the first step in shikimate pathway utilizes these precursors from the carbon metabolic pathways (Santos-Sánchez et al. 2019). *DAHP* was also upregulated in both RtR and MdR lines.

Furthermore, the phenylalanine from the shikimate pathway is a precursor of phenylpropanoid biosynthesis pathway and some of the important downstream genes include *PAL*, *PER* and 4-coumaroyl CoA-Ligase (*4CL*) which were upregulated in RtR and MdR (Caretto et al. 2015). The *4CL*, is a branch point enzyme, and critical for the downstream production of metabolites such as stilbenes, flavones, flavonols, lignins, tannins, coumarins and anthocyanins (Vogt 2010). Lignin biosynthesis genes such as trans-cinnamate 4-monooxygenase, cinnamyl alcohol dehydrogenase and caffeic acid 3-O-methyltransferase (Tohge et al. 2013) were expressed in both resistant lines. The shikimate pathway is also responsible for the production of aromatic amino acids like tyrosine, and tryptophan. These amino acids are precursor to many alkaloids. In our study, we found more upregulation in genes responsible for the production of phenolics than alkaloids and glucosinolates. These secondary metabolites are known to have a role in broad-spectrum defense response against pathogens when accumulated in plant tissues (Dixon et al. 2002; Wagner et al. 2019; Zaynab et al. 2018).

Similarly, the biosynthesis of SA depends on the secondary metabolic pathways such as the cinnamic acid pathway and the isochorismate pathway, both of which requires the shikimate pathway (Backer et al. 2019). The role of SA has been documented in clubroot resistance previously (Lemarie et al. 2015; Ning et al. 2019; Siemens et al. 2006). Interestingly, in addition to the secondary metabolic pathways, SA pathway can be activated by the gene *RPP5*- like gene,

also known as Suppressor of npr1-1, constitutive 1 (*snc1*). The *snc1* is known to activate SA dependent resistance pathways via *PAD4* (Zhang et al. 2003). Moreover, the plants expressing *snc1* confer constitutive resistance in the absences of cell death, which has been generally associated with R-gene mediated resistance (Li et al. 2001). Previously, *snc1* has been found responsible for constitutively express PR genes and provide resistance against both *Pseudomonas syringae* pv *maculicola* ES4326 and *Peronospora parasitica* Noco2 in *Arabidopsis* (Zhang et al. 2003). Similarly, Fu et al. (2019) found that *snc1* expression is downregulated in the clubroot susceptible *B. rapa* plants. This is the first study of *B. napus*, where the *snc1*, *PAD4* and *PR1* genes were found upregulated in RtR plants in response to *P. brassicae*. Yi and Richards (2007) found that transgenic overexpression of *snc1* resulted in resistance against another race of *Pseudomonas syringae* via the accumulation of SA. Considering this gene is positioned on BnaA08, we hypothesized that this could be the *Crr1b* locus as mentioned by (Hatakeyama et al. 2013), the minor locus contributing towards clubroot resistance. Similarly, SA- dependent activation of defense responses has also been connected with the induction of aspartic proteases (Xia et al. 2004). Among the genes located on BnaA03, upregulated in MdR lines, aspartic proteinase-like protein 1(*AP*) was significantly upregulated. Thus, upregulation of the aforementioned genes in RtR and MdR lines suggest that the activation of SA pathway is critical for clubroot defense. Future studies should functionally validate its function in mediating clubroot resistance.

The knowledge from our study was exploited to modulate primary and secondary metabolic pathways for the induction of defense related genes by priming plants using carbohydrates such as trehalose. Such responses were found to be induced on treating the susceptible plants, thus making them partially resistant to *P. brassicae*. Previous studies have shown that *P. brassicae* increases the accumulation of sugars like glucose, fructose and trehalose at the secondary infection or gall stage (Bae et al. 2005; Brodmann et al. 2002; Li et al. 2018). The concurrent increase in the expression of *SWEET* genes was documented by Li et al. (2018) in clubroot susceptible roots, suggesting that *P. brassicae* induces the expression of *SWEET* genes for sugar translocation between the source and organs and the gall tissue. Thus, this explains the induction of defense related genes in lower trehalose treatments, as trehalose acts as an elicitor.In contrast, as the concentration of trehalose increases, there was a noticeable decline in resistance until all plants were completely susceptible to clubroot under 120 mM treatment.

Higher concentrations of trehalose have been known to cause trehalose toxicity, which has been linked to the insufficient metabolizable sugar, thus affecting the energy status of the plants. This study thus demonstrates the importance of primary metabolism in regulation of clubroot resistance.

6.1 Conclusion and future perspectives

In summary, the research described in this dissertation has provided new insights into the RNAome and ncRNAome of clubroot-resistant canola by discussing the specific responses of the RtR and MdR lines against clubroot, giving us information on the important genes/pathways essential for slowing the disease progression in both RtR and MdR lines at 10 dpi. In addition, results presented in this project explain the importance of the of primary metabolism in clubroot resistance, as it regulates the production of metabolites essential for secondary metabolism, plant hormone signalling and expression of defense-related genes (Rojas et al. 2014). We also showed the effect of regulation of defense responses by modulating primary metabolic pathways using trehalose treatment on clubroot resistance. Additionally, the DE lncRNAs and DEGs from these cultivars, responsive to clubroot were studied and this provided us with candidate lncRNAs and genes whose mechanism can be studied using loss/gain-of -function studies. The DE lncRNAs and DEGs from this study were also mined for SSR markers. We found one DE lncRNA (LNC_000424) and 3 DEGs (BnaA08g03250D from BnaA08 of RtR, and BnaA03g41300D and BnaA03g44400D from BnaA03 of MdR) to contain SSRs which showed strong association with clubroot resistance. These SSRs were able to distinguish the resistant cultivars from the susceptible with 97-99% accuracy. Although they did not show perfect association, these markers can used for marker assisted breeding. The results from this project have thus advanced our knowledge about the pathways important for clubroot resistance and will be useful for the development of clubroot resistant cultivars. Future studies should be aimed at studying the role of the various lncRNAs (LNC_003848, LNC_000424, LNC_001163) and mRNAs (*CRF4*, *RPP5*-like R-gene (*snc1*), *AP*), discussed in this project in plant defense, as they could be potential candidates for breeding/generating clubroot resistant canola cultivars.

Bibliography

Alam, M. Z., Carpenter-Boggs, L., Mitra, S., Haque, M., Halsey, J., Rokonuzzaman, M., ... & Moniruzzaman, M. (2017). Effect of salinity intrusion on food crops, livestock, and fish species at Kalapara Coastal Belt in Bangladesh. Journal of Food Quality, 2017, 1-23.

Aldon, D., Mbengue, M., Mazars, C., & Galaud, J. P. (2018). Calcium signalling in plant biotic interactions. International Journal of Molecular Sciences, 19(3), 665, 1-19.

Alonso, C., Ramos-Cruz, D., & Becker, C. (2019). The role of plant epigenetics in biotic interactions. New Phytologist, 221(2), 731-737.

Alsayied, N. F., Fernández, J. A., Schwarzacher, T., & Heslop-Harrison, J. S. (2015). Diversity and relationships of *Crocus sativus* and its relatives analysed by inter-retroelement amplified polymorphism (IRAP). Annals of Botany, 116(3), 359-368.

Amante-Bordeos, A., Sitch, L. A., Nelson, R., Dalmacio, R. D., Oliva, N. P., Aswidinnoor, H., & Leung, H. (1992). Transfer of bacterial blight and blast resistance from the tetraploid wild rice *Oryza minuta* to cultivated rice, *Oryza sativa*. Theoretical and Applied Genetics, 84(3-4), 345-354.

Ambawat, S., Sharma, P., Yadav, N. R., & Yadav, R. C. (2013). MYB transcription factor genes as regulators for plant responses: an overview. Physiology and Molecular Biology of Plants, 19(3), 307-321.

Amom, T., & Nongdam, P. (2017). The use of molecular marker methods in plants: a review. International Journal of Current Research and Review, 9(17), 1-7.

Andersen, J. R., & Lübberstedt, T. (2003). Functional markers in plants. Trends in Plant Science, 8(11), 554-560.

Angel, A., Song, J., Dean, C., & Howard, M. (2011). A Polycomb-based switch underlying quantitative epigenetic memory. Nature, 476(7358), 105-108.

Ariel, F., Jegu, T., Latrasse, D., Romero-Barrios, N., Christ, A., Benhamed, M., & Crespi, M. (2014). Noncoding transcription by alternative RNA polymerases dynamically regulates an auxin-driven chromatin loop. Molecular Cell, 55(3), 383-396.

Ariel, F., Lucero, L., Christ, A., Mammarella, M. F., Jegu, T., Veluchamy, A., ... & Benhamed, M. (2020). R-Loop mediated trans action of the *APOLO* long noncoding RNA. Molecular Cell.

Asano, T., Kageyama, K., & Hyakumachi, M. (2000). Germination of surface-disinfected resting spores of *Plasmodiophora brassicae* and their root hair infection in turnip hairy roots. Mycoscience, 41(1), 49-54.

Au, P. C. K., Dennis, E. S., & Wang, M. B. (2017). Analysis of argonaute 4-associated long non-coding RNA in *Arabidopsis thaliana* sheds novel insights into gene regulation through RNA-directed DNA methylation. Genes, 8(8), 198.

Badstöber, J., Ciaghi, S., & Neuhauser, S. (2020). Dynamic cell wall modifications in brassicas during clubroot disease. BioRxiv.

Bae, H., Herman, E., Bailey, B., Bae, H. J., & Sicher, R. (2005). Exogenous trehalose alters *Arabidopsis* transcripts involved in cell wall modification, abiotic stress, nitrogen metabolism, and plant defense. Physiologia Plantarum, 125(1), 114-126.

Bagniewska-Zadworna, A., Barakat, A., Łakomy, P., Smoliński, D. J., & Zadworny, M. (2014). Lignin and lignans in plant defence: insight from expression profiling of cinnamyl alcohol dehydrogenase genes during development and following fungal infection in *Populus*. Plant Science, 229, 111-121.

Bahieldin, A., Atef, A., Edris, S., Gadalla, N. O., Ali, H. M., Hassan, S. M., ... & El-Domyati, F. M. (2016). Ethylene responsive transcription factor ERF109 retards PCD and improves salt tolerance in plant. BMC Plant Biology, 16(1), 216.

Baldrich, P., Campo, S., Wu, M. T., Liu, T. T., Hsing, Y. I. C., & Segundo, B. S. (2015). MicroRNA-mediated regulation of gene expression in the response of rice plants to fungal elicitors. RNA Biology, 12(8), 847-863.

Baliyan, N., Malik, R., Rani, R., Mehta, K., Vashisth, U., Dhillon, S., & Boora, K. S. (2018). Integrating marker-assisted background analysis with foreground selection for pyramiding bacterial blight resistance genes into Basmati rice. Comptes Rendus Biologies, 341(1), 1-8.

Bardou, F., Ariel, F., Simpson, C. G., Romero-Barrios, N., Laporte, P., Balzergue, S., ... & Crespi, M. (2014). Long noncoding RNA modulates alternative splicing regulators in *Arabidopsis*. Developmental Cell, 30(2), 166-176.

Bari, R., & Jones, J. D. (2009). Role of plant hormones in plant defence responses. Plant Molecular Biology, 69(4), 473-488.

Bari, R., Pant, B. D., Stitt, M., & Scheible, W. R. (2006). *PHO2*, microRNA399, and *PHR1* define a phosphate-signaling pathway in plants. Plant Physiology, 141(3), 988-999.

Beier, S., Thiel, T., Münch, T., Scholz, U., & Mascher, M. (2017). MISA-web: a web server for microsatellite prediction. Bioinformatics, 33(16), 2583-2585.

Bick, J. A., & Lange, B. M. (2003). Metabolic cross talk between cytosolic and plastidial pathways of isoprenoid biosynthesis: unidirectional transport of intermediates across the chloroplast envelope membrane. Archives of Biochemistry and Biophysics, 415(2), 146-154.

Belkhadir, Y., Subramaniam, R., & Dangl, J. L. (2004). Plant disease resistance protein signaling: NBS–LRR proteins and their partners. Current Opinion in Plant Biology, 7(4), 391-399.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the royal statistical society. Series B (Methodological), 289-300.

Bennett, R. A., Thiagarajah, M. R., King, J. R., & Rahman, M. H. (2008). Interspecific cross of *Brassica oleracea var. alboglabra* and *B. napus*: effects of growth condition and silique age on the efficiency of hybrid production, and inheritance of erucic acid in the self-pollinated backcross generation. Euphytica, 164(2), 593-601.

Bhatia, G., Goyal, N., Sharma, S., Upadhyay, S. K., & Singh, K. (2017). Present scenario of long non-coding RNAs in plants. Non-coding RNA, 3(2), 16.

Bigeard, J., Colcombet, J., & Hirt, H. (2015). Signaling mechanisms in pattern-triggered immunity (PTI). Molecular Plant, 8(4), 521-539.

Blevins, T., Podicheti, R., Mishra, V., Marasco, M., Wang, J., Rusch, D., ... & Pikaard, C. S. (2015). Identification of Pol IV and RDR2-dependent precursors of 24 nt siRNAs guiding de novo DNA methylation in *Arabidopsis*. Elife, 4, e09591.

Boch, J., & Bonas, U. (2010). *Xanthomonas* AvrBs3 family-type III effectors: discovery and function. Annual Review of Phytopathology, 48, 419-436.

Bogdanove, A. J., Schornack, S., & Lahaye, T. (2010). TAL effectors: finding plant genes for disease and defense. Current Opinion in Plant Biology, 13(4), 394-401.

Bolouri Moghaddam, M. R., & Van den Ende, W. (2012). Sugars and plant innate immunity. Journal of Experimental Botany, 63(11), 3989-3998.

Bolton, M. D. (2009). Primary metabolism and plant defense—fuel for the fire. Molecular plant-microbe Interactions, 22(5), 487-497.

Borges, F., & Martienssen, R. A. (2015). The expanding world of small RNAs in plants. Nature reviews Molecular Cell Biology, 16(12), 727-741.

Bourke, A. (1993). 'The visitation of god?' The potato and the great Irish famine. Lilliput Press Ltd.

Brant, E. J., & Budak, H. (2018). Plant small non-coding RNAs and their roles in biotic stresses. Frontiers in Plant Science, 9, 1038.

Breen, S., Williams, S. J., Outram, M., Kobe, B., & Solomon, P. S. (2017). Emerging insights into the functions of pathogenesis-related protein 1. Trends in Plant Science, 22(10), 871-879.

Brodmann, D., Schuller, A., Ludwig-Müller, J., Aeschbacher, R. A., Wiemken, A., Boller, T., & Wingler, A. (2002). Induction of trehalase in *Arabidopsis* plants infected with the trehalose-producing pathogen *Plasmodiophora brassicae*. Molecular Plant-Microbe Interactions, 15(7), 693-700.

Buffon, G., Blasi, E. A., Adamski, J. M., Ferla, N. J., Berger, M., Santi, L., ... & Sperotto, R. A. (2016). Physiological and molecular alterations promoted by *Schizotetranychus oryzae* mite infestation in rice leaves. Journal of Proteome Research, 15(2), 431-446.

Burki, F., Kudryavtsev, A., Matz, M. V., Aglyamova, G. V., Bulman, S., Fiers, M., ... & Pawlowski, J. (2010). Evolution of Rhizaria: new insights from phylogenomic analysis of uncultivated protists. BMC Evolutionary Biology, 10(1), 377.

Cai, K., Zhu, L., Zhang, K., Li, L., Zhao, Z., Zeng, W., & Lin, X. (2019). Development and characterization of EST-SSR markers from RNA-Seq data in *Phyllostachys violascens*. Frontiers in Plant Science, 10, 50.

Cai, Q., Qiao, L., Wang, M., He, B., Lin, F. M., Palmquist, J., ... & Jin, H. (2018). Plants send small RNAs in extracellular vesicles to fungal pathogen to silence virulence genes. Science, 360(6393), 1126-1129.

Campalans, A., Kondorosi, A., & Crespi, M. (2004). *Enod40*, a short open reading frame–containing mRNA, induces cytoplasmic localization of a nuclear RNA binding protein in *Medicago truncatula*. The Plant Cell, 16(4), 1047-1059.

Canola Council of Canada. Retrieved May 23, 2020. www.canolacouncil.org.

Cao, Y., Fan, G., Zhai, X., & Dong, Y. (2018). Genome-wide analysis of lncRNAs in *Paulownia tomentosa* infected with phytoplasmas. Acta Physiologiae Plantarum, 40(3), 49.

Cao, T., Srivastava, S., Rahman, M. H., Kav, N. N. V., Hotte, N., Deyholos, M. K., & Strelkov, S. E. (2008). Proteome-level changes in the roots of *Brassica napus* as a result of *Plasmodiophora brassicae* infection. Plant Science, 174(1), 97-115.

Caretto, S., Linsalata, V., Colella, G., Mita, G., & Lattanzio, V. (2015). Carbon fluxes between primary metabolism and phenolic pathway in plant tissues under stress. International Journal of Molecular Science, 16(11), 26378-26394.

Carpenter, M. A., Goulden, D. S., Woods, C. J., Thomson, S. J., Kenel, F., Frew, T. J., ... & Timmerman-Vaughan, G. M. (2018). Genomic selection for ascochyta blight resistance in pea. Frontiers in Plant Science, 9, 1878.

Carré, P., & Pouzet, A. (2014). Rapeseed market, worldwide and in Europe. *Ocl*, *21*(1), D102.

Castaings, L., Bergonzi, S., Albani, M. C., Kemi, U., Savolainen, O., & Coupland, G. (2014). Evolutionary conservation of cold-induced antisense RNAs of *FLOWERING LOCUS C* in *Arabidopsis thaliana* perennial relatives. Nature Communications, 5(1), 1-9.

Chakraborty, S., Tiedemann, A. V., & Teng, P. S. (2000). Climate change: potential impact on plant diseases. Environmental pollution, *108*(3), 317-326.

Chakraborty, S., & Newton, A. C. (2011). Climate change, plant diseases and food security: an overview. Plant Pathology, 60(1), 2-14.

Chalhoub, B., Denoeud, F., Liu, S., Parkin, I. A., Tang, H., & Wang, X. (2014). Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. Science, 345 (6199), pp.950-953.

Chang, A., Lamara, M., Wei, Y., Hu, H., Parkin, I. A., Gossen, B. D., ... & Yu, F. (2019). Clubroot resistance gene *Rcr6* in *Brassica nigra* resides in a genomic region homologous to chromosome BnaA08 in *B. rapa*. BMC Plant Biology, 19(1), 224.

Chao, T. C., Zhang, Q., Li, Z., Tiwari, S. K., Qin, Y., Yau, E., ... & Karris, M. A. Y. (2019). The long noncoding RNA *HEAL* regulates HIV-1 replication through epigenetic regulation of the HIV-1 promoter. mBio, 10(5), e02016-19.

Chauhan, S., Yogindran, S., & Rajam, M. V. (2017). Role of miRNAs in biotic stress reactions in plants. Indian Journal of Plant Physiology, 22(4), 514-529.

Chekanova, J. A. (2015). Long non-coding RNAs and their functions in plants. Current Opinion in Plant Biology, 27, 207-216.

Chen, X. (2009). Small RNAs and their roles in plant development. Annual Review of Cell and Developmental Biology, 25, 21-44.

Chen, M., Tan, Z., Zeng, G., & Peng, J. (2010). Comprehensive analysis of simple sequence repeats in pre-miRNAs. Molecular Biology and Evolution, 27(10), 2227-2232.

Chen, D., Yuan, C., Zhang, J., Zhang, Z., Bai, L., Meng, Y., ... & Chen, M. (2012). PlantNATsDB: a comprehensive database of plant natural antisense transcripts. Nucleic Acids Research, 40(D1), D1187-D1193.

Chen, J., Pang, W., Chen, B., Zhang, C., & Piao, Z. (2015). Transcriptome Analysis of *Brassica rapa* near-isogenic lines carrying clubroot-resistant and -susceptible alleles in response to *Plasmodiophora brassicae* during early infection. Frontiers in Plant Science, 6, 1183.

Chen, J., Quan, M., & Zhang, D. (2015). Genome-wide identification of novel long non-coding RNAs in *Populus tomentosa* tension wood, opposite wood and normal wood xylem by RNA-seq. Planta, 241(1), 125-143.

Chen, L. L., & Yang, L. (2015). Regulation of circRNA biogenesis. RNA Biology, 12(4), 381-388.

Chen, G., Cui, J., Wang, L., Zhu, Y., Lu, Z., & Jin, B. (2017). Genome-wide identification of circular RNAs in *Arabidopsis thaliana*. Frontiers in Plant Science, 8, 1678.

Chinchilla, D., Zipfel, C., Robatzek, S., Kemmerling, B., Nürnberger, T., Jones, J. D., ... & Boller, T. (2007). A flagellin-induced complex of the receptor *FLS2* and *BAK1* initiates plant defence. Nature, 448(7152), 497-500.

Chiou, T. J., Aung, K., Lin, S. I., Wu, C. C., Chiang, S. F., & Su, C. L. (2006). Regulation of phosphate homeostasis by microRNA in *Arabidopsis*. The Plant Cell, 18(2), 412-421.

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 gene into canola breeding lines using a marker-assisted selection approach. Acta Hortic, 1005, 599-605.

Chu, Q., Bai, P., Zhu, X., Zhang, X., Mao, L., Zhu, Q. H., ... & Ye, C. Y. (2020). Characteristics of plant circular RNAs. Briefings in Bioinformatics, 21(1), 135-143.

Church, J. A., & White, N. J. (2011). Sea-level rise from the late 19th to the early 21st century. Surveys in Geophysics, 32(4-5), 585-602.

Ciaghi, S., Schwelm, A., & Neuhauser, S. (2019). Transcriptomic response in symptomless roots of clubroot infected kohlrabi (*Brassica oleracea var. gongylodes*) mirrors resistant plants. BMC Plant Biology, 19(1), 288.

Clevenger, J., Chu, Y., Chavarro, C., Botton, S., Culbreath, A., Isleib, T. G., ... & Ozias-Akins, P. (2018). Mapping late leaf spot resistance in peanut (*Arachis hypogaea*) using QTL-seq reveals markers for marker-assisted selection. Frontiers in Plant Science, 9, 83.

Collard, B. C., Jahufer, M. Z. Z., Brouwer, J. B., & Pang, E. C. K. (2005). An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. Euphytica, 142(1-2), 169-196.

Collard, B. C., & Mackill, D. J. (2008). Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. Philosophical Transactions of the Royal Society B: Biological Sciences, 363(1491), 557-572.

Compaan, B., Yang, W. C., Bisseling, T., & Franssen, H. (2001). ENOD40 expression in the pericycle precedes cortical cell division in Rhizobium-legume interaction and the highly conserved internal region of the gene does not encode a peptide. Plant and Soil, 230(1), 1-8.

Compaan, B., Ruttink, T., Albrecht, C., Meeley, R., Bisseling, T., & Franssen, H. (2003). Identification and characterization of a *Zea mays* line carrying a transposon-tagged *ENOD40*. Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression, 1629(1-3), 84-91.

Conn, V. M., Hugouvieux, V., Nayak, A., Conos, S. A., Capovilla, G., Cildir, G., ... & Conn, S. J. (2017). A circRNA from *SEPALLATA3* regulates splicing of its cognate mRNA through R-loop formation. Nature Plants, 3(5), 1-5.

Croteau, R., Kutchan, T. M., & Lewis, N. G. (2000). Natural products (secondary metabolites). Biochemistry and molecular biology of plants, *24*, 1250-1319.

Csorba, T., Questa, J. I., Sun, Q., & Dean, C. (2014). Antisense *COOLAIR* mediates the coordinated switching of chromatin states at *FLC* during vernalization. Proceedings of the National Academy of Sciences, 111(45), 16160-16165.

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. The Plant Journal, 89(3), 577-589.

Danson, J., Wasano, K., & Nose, A. (2000). Infection of rice plants with the sheath blight fungus causes an activation of pentose phosphate and glycolytic pathways. European Journal of Plant Pathology, 106(6), 555-561.

Dai, X., Zhuang, Z., & Zhao, P. X. (2018). psRNATarget: a plant small RNA target analysis server (2017 release). Nucleic Acids Research, 46(W1), W49-w54.

Dao, T. T. H., Linthorst, H. J. M., & Verpoorte, R. (2011). Chalcone synthase and its functions in plant resistance. Phytochemistry Reviews, 10(3), 397-412.

Daugherty, M. P., Zeilinger, A. R., & Almeida, R. P. (2017). Conflicting effects of climate and vector behavior on the spread of a plant pathogen. Phytobiomes, 1(1), 46-53.

Daun, J. K. (2011). Origin, distribution, and production. In Canola .AOCS Press, pp. 1-

27.Delorge, I., Janiak, M., Carpentier, S., & Van Dijck, P. (2014). Fine tuning of trehalose biosynthesis and hydrolysis as novel tools for the generation of abiotic stress tolerant plants. Frontiers in Plant Science, *5*, 147.

Delourme, R., Foisset, N., Horvais, R., Barret, P., Champagne, G., Cheung, W. Y., ... & Renard, M. (1998). Characterisation of the radish introgression carrying the *Rfo* restorer gene for the Ogu-INRA cytoplasmic male sterility in rapeseed (*Brassica napus L.*). Theoretical and Applied Genetics, 97(1-2), 129-134.

Deng, P., Liu, S., Nie, X., Weining, S., & Wu, L. (2018). Conservation analysis of long non-coding RNAs in plants. Science China Life Sciences, 61(2), 190-198.

Derksen, H., Rampitsch, C., & Daayf, F. (2013). Signaling cross-talk in plant disease resistance. Plant Science, 207, 79-87.

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

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

Ding, J., Lu, Q., Ouyang, Y., Mao, H., Zhang, P., Yao, J., ... & Zhang, Q. (2012). A long noncoding RNA regulates photoperiod-sensitive male sterility, an essential component of hybrid rice. Proceedings of the National Academy of Sciences, 109(7), 2654-2659.

Ding, J., Shen, J., Mao, H., Xie, W., Li, X., & Zhang, Q. (2012). RNA-directed DNA methylation is involved in regulating photoperiod-sensitive male sterility in rice. Molecular Plant, 5(6), 1210-1216.

Dixon, R. A., Achnine, L., Kota, P., Liu, C. J., Reddy, M. S., & Wang, L. (2002). The phenylpropanoid pathway and plant defence—a genomics perspective. Molecular Plant Pathology, 3(5), 371-390.

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

Dixon, G. R. (2014). Clubroot (*Plasmodiophora brassicae* Woronin)–an agricultural and biological challenge worldwide. Canadian Journal of Plant Pathology, 36(sup1), 5-18.

Dobson, R., Gabrielson, R., Baker, A., & Bennett, L. (1983). Effects of lime particle size and distribution and fertilizer formulation on clubroot disease caused by *Plasmodiophora brassicae*. Plant Disease, 67(1), 50-52.

Dong, H., & Beer, S. V. (2000). Riboflavin induces disease resistance in plants by activating a novel signal transduction pathway. Phytopathology, 90(8), 801-811.

Dowen, R. H., Pelizzola, M., Schmitz, R. J., Lister, R., Dowen, J. M., Nery, J. R., ... & Ecker, J. R. (2012). Widespread dynamic DNA methylation in response to biotic stress. Proceedings of the National Academy of Sciences, 109(32), E2183-E2191.

dos Santos, J. P. R., Pires, L. P. M., de Castro Vasconcellos, R. C., Pereira, G. S., Von Pinho, R. G., & Balestre, M. (2016). Genomic selection to resistance to *Stenocarpella maydis* in maize lines using DArTseq markers. BMC Genetics, 17(1), 86.

Du, W. W., Fang, L., Yang, W., Wu, N., Awan, F. M., Yang, Z., & Yang, B. B. (2017). Induction of tumor apoptosis through a circular RNA enhancing Foxo3 activity. Cell Death and Differentiation, 24(2), 357-370.

Dudareva, N., Klempien, A., Muhlemann, J. K., & Kaplan, I. (2013). Biosynthesis, function and metabolic engineering of plant volatile organic compounds. New Phytologist, *198*(1), 16-32.

Early, R. (2009). Pathogen control in primary production: crop foods. In Foodborne Pathogens (pp. 205-279). Woodhead Publishing.

Ehness, R., Ecker, M., Godt, D. E., & Roitsch, T. (1997). Glucose and stress independently regulate source and sink metabolism and defense mechanisms via signal transduction pathways involving protein phosphorylation. The Plant Cell, 9(10), 1825-1841.

Eissa, H. F., Hassanien, S. E., Ramadan, A. M., El-Shamy, M. M., Saleh, O. M., Shokry, A. M., ... & Hassan, S. M. (2017). Developing transgenic wheat to encounter rusts and powdery mildew by overexpressing barley *chi26* gene for fungal resistance. Plant Methods, 13(1), 41.

Ellur, R. K., Khanna, A., Yadav, A., Pathania, S., Rajashekara, H., Singh, V. K., ... & Prakash, G. (2016). Improvement of Basmati rice varieties for resistance to blast and bacterial blight diseases using marker assisted backcross breeding. Plant Science, 242, 330-341.

Ernst, T. W., Kher, S., Stanton, D., Rennie, D. C., Hwang, S. F., & Strelkov, S. E. (2019). *Plasmodiophora brassicae* resting spore dynamics in clubroot resistant canola (*Brassica napus*) cropping systems. Plant Pathology, 68(2), 399-408.

FAO. 2002. Trade and food security: Conceptualizing the linkages. In presentation at the Conference on Trade, Agricultural Development, and Food Security: The Impact of Recent Economic and Trade Policy Reform, Food and Agricultural Organization, Rome, July (pp. 11-12).

FAO. 2020. Crop Prospects and Food Situation - Quarterly Global Report No. 1, March 2020. Rome.

Falcone Ferreyra, M. L., Rius, S., & Casati, P. (2012). Flavonoids: biosynthesis, biological functions, and biotechnological applications. Frontiers in Plant Science, 3, 222.

Faigón-Soverna, A., Harmon, F. G., Storani, L., Karayekov, E., Staneloni, R. J., Gassmann, W., ... & Yanovsky, M. J. (2006). A constitutive shade-avoidance mutant implicates TIR-NBS-LRR proteins in *Arabidopsis* photomorphogenic development. The Plant Cell, 18(11), 2919-2928.

Fan, G., Cao, Y., & Wang, Z. (2018). Regulation of long noncoding RNAs responsive to phytoplasma infection in *Paulownia tomentosa*. International Journal of Genomics, 2018, 1-16.

Fang, S., Zhang, L., Guo, J., Niu, Y., Wu, Y., Li, H., ... & Sun, L. (2018). NONCODEV5: a comprehensive annotation database for long non-coding RNAs. Nucleic Acids Research, 46(D1), D308-D314.

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(2), 91-101.

Fei, W., Feng, J., Rong, S., Strelkov, S. E., Gao, Z., & Hwang, S. F. (2016). Infection and gene expression of the clubroot pathogen *Plasmodiophora brassicae* in resistant and susceptible canola cultivars. Plant Disease, 100(4), 824-828.

Feng, H., Zhang, Q., Wang, B., Fu, Y., Huang, L., Wang, X., & Kang, Z. (2015). Exploration of microRNAs and their targets engaging in the resistance interaction between wheat and stripe rust. Frontiers in Plant Science, 6, 469.

Finch, S., Samuel, A., & Lane, G. P. (2014). Lockhart and wiseman's crop husbandry including grassland. Elsevier.

Franco-Zorrilla, J. M., Valli, A., Todesco, M., Mateos, I., Puga, M. I., Rubio-Somoza, I., ... & Paz-Ares, J. (2007). Target mimicry provides a new mechanism for regulation of microRNA activity. Nature Genetics, 39(8), 1033-1037.

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

Fredua-Agyeman, R., & Rahman, H. (2016). Mapping of the clubroot disease resistance in spring *Brassica napus* canola introgressed from European winter canola cv.'Mendel'. Euphytica, 211(2), 201-213.

Fu, D., Ma, B., Mason, A. S., Xiao, M., Wei, L., & An, Z. (2013). Micro RNA-based molecular markers: a novel PCR-based genotyping technique in Brassica species. Plant Breeding, 132(4), 375-381.

Fu, P., Piao, Y., Zhan, Z., Zhao, Y., Pang, W., Li, X., & Piao, Z. (2019). Transcriptome profile of *Brassica rapa* L. reveals the involvement of jasmonic acid, ethylene, and brassinosteroid signaling pathways in clubroot resistance. Agronomy, 9(10), 589.

Fujii, H., Chiou, T. J., Lin, S. I., Aung, K., & Zhu, J. K. (2005). A miRNA involved in phosphate-starvation response in *Arabidopsis*. Current Biology, 15(22), 2038-2043.

Fyles, H., & Madramootoo, C. (2016). Key drivers of food insecurity. In Emerging Technologies for Promoting Food Security (pp. 1-19). Woodhead Publishing.

Gallart, A. P., Pulido, A. H., De Lagrán, I. A. M., Sanseverino, W., & Cigliano, R. A. (2016). GREENC: a Wiki-based database of plant lncRNAs. Nucleic Acids Research, 44(Database issue), D1161.

Ganie, S. A., & Mondal, T. K. (2015). Genome-wide development of novel miRNA-based microsatellite markers of rice (*Oryza sativa*) for genotyping applications. Molecular Breeding, 35(1), 51.

Gassmann, W., Hinsch, M. E., & Staskawicz, B. J. (1999). The *Arabidopsis RPS4* bacterial-resistance gene is a member of the TIR-NBS-LRR family of disease-resistance genes. The Plant Journal, 20(3), 265-277.

German, S. E., & Kolmer, J. A. (1992). Effect of gene *Lr34* in the enhancement of resistance to leaf rust of wheat. Theoretical and Applied Genetics, 84(1-2), 97-105.

Gessler, C., Pertot, I., & Perazzolli, M. (2011). *Plasmopara viticola*: a review of knowledge on downy mildew of grapevine and effective disease management. Phytopathologia Mediterranea, 50(1), 3-44.

Gómez-Gómez, L., & Boller, T. (2000). *FLS2*: an LRR receptor–like kinase involved in the perception of the bacterial elicitor flagellin in *Arabidopsis*. Molecular Cell, 5(6), 1003-1011.

Gómez-Ariza, J., Campo, S., Rufat, M., Estopà, M., Messeguer, J., Segundo, B. S., & Coca, M. (2007). Sucrose-mediated priming of plant defense responses and broad-spectrum disease resistance by overexpression of the maize pathogenesis-related PRms protein in rice plants. Molecular Plant-Microbe Interactions, 20(7), 832-842.

Govind, S. R., Jogaiah, S., Abdelrahman, M., Shetty, H. S., & Tran, L. S. P. (2016). Exogenous trehalose treatment enhances the activities of defense-related enzymes and triggers resistance against downy mildew disease of pearl millet. Frontiers in Plant Science, 7, 1593.

Green, M. R., & Sambrook, J. (2018). Optimizing primer and probe concentrations for use in real-time polymerase chain reaction (PCR) assays. Cold Spring Harbor Protocols, 2018(10), pdb-prot095018.

Gruber, A. R., Lorenz, R., Bernhart, S. H., Neuböck, R., & Hofacker, I. L. (2008). The vienna RNA websuite. Nucleic acids research, 36(suppl_2), W70-W74.

Gultyaev, A. P., & Roussis, A. (2007). Identification of conserved secondary structures and expansion segments in *ENOD40* RNAs reveals new enod40 homologues in plants. Nucleic Acids Research, 35(9), 3144-3152.

Guha, P. K., Mazumder, A., Das, A., Pani, D. R., & Mondal, T. K. (2019). In silico identification of long non-coding RNA based simple sequence repeat markers and their application in diversity analysis in rice. Gene Reports, 16, 100418.

Gupta, P., Nutan, K. K., Singla-Pareek, S. L., & Pareek, A. (2017). Abiotic stresses cause differential regulation of alternative splice forms of GATA transcription factor in rice. Frontiers in Plant Science, 8, 1944.

Guttman, M., Amit, I., Garber, M., French, C., Lin, M. F., Feldser, D., ... & Cabili, M. N. (2009). Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature, 458(7235), 223-227.

Guttman, M., Garber, M., Levin, J. Z., Donaghey, J., Robinson, J., Adiconis, X., ... & Rinn, J. L. (2010). Ab initio reconstruction of cell type–specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. Nature Biotechnology, 28(5), 503-510.

Hanson, P., Lu, S. F., Wang, J. F., Chen, W., Kenyon, L., Tan, C. W., ... & Ledesma, D. (2016). Conventional and molecular marker-assisted selection and pyramiding of genes for multiple disease resistance in tomato. Scientia Horticulturae, 201, 346-354.

Harakava, R. (2005). Genes encoding enzymes of the lignin biosynthesis pathway in Eucalyptus. Genetics and Molecular Biology, 28(3), 601-607.

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, 59(10), 805-815.

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

Haverkort, A. J., Boonekamp, P. M., Hutten, R., Jacobsen, E., Lotz, L. A. P., Kessel, G. J. T., ... & Visser, R. G. F. (2016). Durable late blight resistance in potato through dynamic

varieties obtained by cisgenesis: scientific and societal advances in the DuRPh project. Potato Research, 59(1), 35-66.

Hawkes, E. J., Hennelly, S. P., Novikova, I. V., Irwin, J. A., Dean, C., & Sanbonmatsu, K. Y. (2016). *COOLAIR* antisense RNAs form evolutionarily conserved elaborate secondary structures. Cell Reports, 16(12), 3087-3096.

He, R., Chang, Z., Yang, Z., Yuan, Z., Zhan, H., Zhang, X., & Liu, J. (2009). Inheritance and mapping of powdery mildew resistance gene *Pm43* introgressed from *Thinopyrumintermedium* into wheat. Theoretical and Applied Genetics, 118(6), 1173-1180.

Helliwell, C. A., Robertson, M., Finnegan, E. J., Buzas, D. M., & Dennis, E. S. (2011). Vernalization-repression of *Arabidopsis FLC* requires promoter sequences but not antisense transcripts. PloS one, 6(6), e21513.

Henriquez, M. A., Soliman, A., Li, G., Hannoufa, A., Ayele, B. T., & Daayf, F. (2016). Molecular cloning, functional characterization, and expression of potato (*Solanum tuberosum*) 1-deoxy-d-xylulose 5-phosphate synthase 1 (*StDXS1*) in response to *Phytophthora infestans*. Plant Science, 243, 71-83.

Heo, S. H., Jang, S. J., Choi, J. S., Jang, C. S., Song, J. Y., & Kim, H. G. (2009). Chinese cabbage clubroot pathogen, *Plasmodiophora brassicae*, is genetically stable. Mycobiology, *37*(3), 225-229.

Heo, J. B., & Sung, S. (2011). Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. Science, 331(6013), 76-79.

Herz, S., Kis, K., Bacher, A., & Rohdich, F. (2002). A tomato enzyme catalyzing the phosphorylation of 3, 4-dihydroxy-2-butanone. Phytochemistry, 60(1), 3-11.

Hilaire, E., Young, S. A., Willard, L. H., McGee, J. D., Sweat, T., Chittoor, J. M., ... & Leach, J. E. (2001). Vascular defense responses in rice: peroxidase accumulation in xylem parenchyma cells and xylem wall thickening. Molecular Plant-Microbe Interactions, 14(12), 1411-1419.

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(4), 639-643.

Hofacker, I. L. (2009). R NA Secondary Structure Analysis Using the Vienna RNA Package. Current Protocols in Bioinformatics, 26(1), 12.2.1-12.2.16.

Holdt, L. M., Stahringer, A., Sass, K., Pichler, G., Kulak, N. A., Wilfert, W., ... & Gäbel, G. (2016). Circular non-coding RNA *ANRIL* modulates ribosomal RNA maturation and atherosclerosis in humans. Nature Communications, 7(1), 1-14.

Holme, I. B., Wendt, T., & Holm, P. B. (2013). Intragenesis and cisgenesis as alternatives to transgenic crop development. Plant Biotechnology Journal, 11(4), 395-407.

Horiuchi, S., & Hori, M. (1980). A simple greenhouse technique for obtaining high levels of clubroot incidence. Bulletin of the Chugoku National Agricultural Experiment Station, E, (17), 33-55.

Hou, X., Cui, J., Liu, W., Jiang, N., Zhou, X., Qi, H., ... & Luan, Y. (2020). *LncRNA39026* enhances tomato resistance to *Phytophthora infestans* by Decoying miR168a and inducing *PR* gene expression. Phytopathology, 110(4), 873-880.

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

Hren, M., Ravnikar, M., Brzin, J., Ermacora, P., Carraro, L., Bianco, P. A., ... & Gruden, K. (2009). Induced expression of sucrose synthase and alcohol dehydrogenase I genes in phytoplasma-infected grapevine plants grown in the field. Plant Pathology, 58(1), 170-180.

Hu, Q., Min, L., Yang, X., Jin, S., Zhang, L., Li, Y., ... & Lindsey, K. (2018). Laccase *GhLac1* modulates broad-spectrum biotic stress tolerance via manipulating phenylpropanoid pathway and jasmonic acid synthesis. Plant physiology, 176(2), 1808-1823.

Huang, B., & Zhang, R. (2014). Regulatory non-coding RNAs: revolutionizing the RNA world. Molecular Biology Reports, 41(6), 3915-3923.

Huang, J., Yang, M., Lu, L., & Zhang, X. (2016). Diverse functions of small RNAs in different plant–pathogen communications. Frontiers in Microbiology, 7, 1552.

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), 387.

Huang, C. Y., Wang, H., Hu, P., Hamby, R., & Jin, H. (2019). Small RNAs–big players in plant-microbe interactions. Cell Host & Microbe, 26(2), 173-182.

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(9), 131.

Hudzik, C., Hou, Y., Ma, W., & Axtell, M. J. (2020). Exchange of small regulatory RNAs between plants and their pests. Plant Physiology, 182(1), 51-62.

Hwang, S. F., Strelkov, S. E., Feng, J. I. E., Gossen, B. D., & Howard, R. J. (2012). *Plasmodiophora brassicae*: a review of an emerging pathogen of the Canadian canola (*Brassica napus*) crop. Molecular Plant Pathology, 13(2), 105-113.

Hwang, S. F., Strelkov, S. E., Ahmed, H. U., Manolii, V. P., Zhou, Q., Fu, H., ... & Feindel, D. (2017). Virulence and inoculum density-dependent interactions between clubroot resistant

canola (*Brassica napus*) and *Plasmodiophora brassicae*. Plant Pathology, 66(8), 1318-1328.Ietswaart, R., Wu, Z., & Dean, C. (2012). Flowering time control: another window to the connection between antisense RNA and chromatin. Trends in Genetics, 28(9), 445-453.

Irani, S., Todd, C. D., Wei, Y., & Bonham-Smith, P. C. (2019). Changes in phenylpropanoid pathway gene expression in roots and leaves of susceptible and resistant *Brassica napus* lines in response to *Plasmodiophora brassicae* inoculation. Physiological and Molecular Plant Pathology, 106, 196-203.

Irani, S., Todd, C. D., Wei, Y., & Bonham-Smith, P. C. (2019). Changes in phenylpropanoid pathway gene expression in roots and leaves of susceptible and resistant *Brassica napus* lines in response to *Plasmodiophora brassicae* inoculation. Physiological and Molecular Plant Pathology, 106, 196-203.

Jabnoune, M., Secco, D., Lecampion, C., Robaglia, C., Shu, Q., & Poirier, Y. (2013). A rice cis-natural antisense RNA acts as a translational enhancer for its cognate mRNA and contributes to phosphate homeostasis and plant fitness. The Plant Cell, 25(10), 4166-4182.

Jaganathan, D., Bohra, A., Thudi, M., & Varshney, R. K. (2020). Fine mapping and gene cloning in

the post-NGS era: advances and prospects. Theoretical and Applied Genetics, 1-20.

Jahn, L., Mucha, S., Bergmann, S., Horn, C., Staswick, P., Steffens, B., ... & Ludwig-Müller, J.

(2013). The clubroot pathogen (*Plasmodiophora brassicae*) influences auxin signaling to regulate auxin homeostasis in *Arabidopsis*. Plants, 2(4), 726-749.

Jain, P., Sharma, V., Dubey, H., Singh, P. K., Kapoor, R., Kumari, M., ... & Mondal, T. K. (2017). Identification of long non-coding RNA in rice lines resistant to rice blast pathogen *Maganaporthe oryzae*. Bioinformation, 13(8), 249.

Jaiswal, V., Rawoof, A., Dubey, M., Chhapekar, S. S., Sharma, V., & Ramchiary, N. (2020). Development and characterization of non-coding RNA based simple sequence repeat markers in *Capsicum* species. Genomics, 112(2), 1554-1564.

Janga, H., Aznaourova, M., Boldt, F., Damm, K., Grünweller, A., & Schulte, L. N. (2018). Cas9-mediated excision of proximal DNaseI/H3K4me3 signatures confers robust silencing of microRNA and long non-coding RNA genes. PloS one, 13(2), e0193066.

Jefferies, S. P., King, B. J., Barr, A. R., Warner, P., Logue, S. J., & Langridge, P. (2003). Marker-assisted backcross introgression of the *Yd2* gene conferring resistance to barley yellow dwarf virus in barley. Plant Breeding, 122(1), 52-56.

Ji, R., Wang, Y., Wang, X., Liu, Y., Shen, X., & Feng, H. (2018). Proteomic analysis of the Interaction between *Plasmodiophora brassicae* and Chinese cabbage (*Brassica rapa L. ssp. Pekinensis*) at the initial infection stage. Scientia Horticulturae, 233, 386-393.

Jiang, N., Cui, J., Shi, Y., Yang, G., Zhou, X., Hou, X., ... & Luan, Y. (2019). Tomato *lncRNA23468* functions as a competing endogenous RNA to modulate NBS-LRR genes by decoying miR482b in the tomato-*Phytophthora infestans* interaction. Horticulture Research, 6(1), 1-11.

Jia, H., Wei, X., Yang, Y., Yuan, Y., Wei, F., Zhao, Y., ... & 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(1-2), 93-105.

Jiao, F., Pahwa, K., Manning, M., Dochy, N., & Geuten, K. (2019). Cold induced antisense transcription of *FLOWERING LOCUS C* in distant grasses. Frontiers in Plant Science, 10, 72.

Jones, J. D., & Dangl, J. L. (2006). The plant immune system. Nature, 444(7117), 323-329.

Joshi, R. K., Megha, S., Basu, U., Rahman, M. H., & Kav, N. N. (2016). Genome wide identification and functional prediction of long non-coding RNAs responsive to *Sclerotinia sclerotiorum* infection in *Brassica napus*. PLoS One, 11(7), e0158784.

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

Kalia, R. K., Rai, M. K., Kalia, S., Singh, R., & Dhawan, A. K. (2011). Microsatellite markers: an overview of the recent progress in plants. Euphytica, 177(3), 309-334.

Kang, C., & Liu, Z. (2015). Global identification and analysis of long non-coding RNAs in diploid strawberry *Fragaria vesca* during flower and fruit development. BMC genomics, 16(1), 815.

Kapranov, P., Cheng, J., Dike, S., Nix, D. A., Duttagupta, R., Willingham, A. T., ... & Bell, I. (2007). RNA maps reveal new RNA classes and a possible function for pervasive transcription. Science, 316(5830), 1484-1488.

Karakülah, G., Kurtoğlu, K. Y., & Unver, T. (2016). PeTMbase: a database of plant endogenous target mimics (eTMs). PloS one, 11(12), e0167698.

Karlik, E., Ari, S., & Gozukirmizi, N. (2019). LncRNAs: genetic and epigenetic effects in plants. Biotechnology & Biotechnological Equipment, 33(1), 429-439.

Kawasaki, T., Koita, H., Nakatsubo, T., Hasegawa, K., Wakabayashi, K., Takahashi, H., ... & Shimamoto, K. (2006). Cinnamoyl-CoA reductase, a key enzyme in lignin biosynthesis, is an effector of small GTPase Rac in defense signaling in rice. Proceedings of the National Academy of Sciences, 103(1), 230-235.

Kaur, J., Fellers, J., Adholeya, A., Velivelli, S. L., El-Mounadi, K., Nersesian, N., ... & Shah, D. (2017). Expression of apoplast-targeted plant defensin *MtDef4.2* confers resistance to leaf rust pathogen *Puccinia triticina* but does not affect mycorrhizal symbiosis in transgenic wheat. Transgenic research, 26(1), 37-49.

Kerpedjiev, P., Hammer, S., & Hofacker, I. L. (2015). Forna (force-directed RNA): simple and effective online RNA secondary structure diagrams. Bioinformatics, 31(20), 3377-3379.

Kirubakaran, S. I., Begum, S. M., Ulaganathan, K., & Sakthivel, N. (2008). Characterization of a new antifungal lipid transfer protein from wheat. Plant physiology and Biochemistry, 46(10), 918-927.

Khraiwesh, B., Zhu, J. K., & Zhu, J. (2012). Role of miRNAs and siRNAs in biotic and abiotic stress responses of plants. Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms, 1819(2), 137-148.

Kim, Y. G., Cha, J., & Chandrasegaran, S. (1996). Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. Proceedings of the National Academy of Sciences, 93(3), 1156-1160.

Kim, D. H., & Sung, S. (2017). Vernalization-triggered intragenic chromatin loop formation by long noncoding RNAs. Developmental cell, 40(3), 302-312.

Kim, D. H., Xi, Y., & Sung, S. (2017). Modular function of long noncoding RNA, *COLDAIR*, in the vernalization response. PLoS Genetics, 13(7), e1006939.

Koch, A., Schlemmer, T., Hoefle, L., Werner, B., Preusser, C., Hardt, M., ... & Jelonek, L. (2020). Host-induced gene silencing involves transfer of dsRNA-derived siRNA via extracellular vesicles. bioRxiv.

Kondo, K., Uritani, I., & Oba, K. (2003). Induction mechanism of 3-hydroxy-3-methylglutaryl-CoA reductase in potato tuber and sweet potato root tissues. Bioscience, Biotechnology, and Biochemistry, 67(5), 1007-1017.

Kong, L., Zhang, Y., Ye, Z. Q., Liu, X. Q., Zhao, S. Q., Wei, L., & Gao, G. (2007). CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. Nucleic Acids Research, 35(suppl_2), W345-W349.

Kong, L., Zhao, K., Gao, Y., Miao, L., Chen, C., Deng, H., ... & Yu, X. (2018). Comparative analysis of cytokinin response factors in Brassica diploids and amphidiploids and insights into the evolution of Brassica species. BMC Genomics, 19(1), 728.

Kopp, F., & Mendell, J. T. (2018). Functional classification and experimental dissection of long noncoding RNAs. Cell, 172(3), 393-407.

Kornberg, R. D. (2007). The molecular basis of eukaryotic transcription. Proceedings of the National Academy of Sciences, 104(32), 12955-12961.

Kornienko, A. E., Guenzl, P. M., Barlow, D. P., & Pauler, F. M. (2013). Gene regulation by the act of long non-coding RNA transcription. BMC Biology, 11(1), 59.

Kost, T. D., Gessler, C., Jänsch, M., Flachowsky, H., Patocchi, A., & Broggini, G. A. (2015). Development of the first cisgenic apple with increased resistance to fire blight. PLoS One, 10(12), e0143980.

Kouchi, H., Takane, K. I., So, R. B., Ladha, J. K., & Reddy, P. M. (1999). Rice *ENOD40*: isolation and expression analysis in rice and transgenic soybean root nodules. The Plant Journal, 18(2), 121-129.

Kowata-Dresch, L. S., & May-De Mio, L. L. (2012). Clubroot management of highly infested soils. Crop protection, 35, 47-52.

Kumar, D., Singh, D., Kanodia, P., Prabhu, K. V., Kumar, M., & Mukhopadhyay, K. (2014). Discovery of novel leaf rust responsive microRNAs in wheat and prediction of their target genes. Journal of Nucleic Acids, 2014, 1-12

Kumar, R. (2014). Role of microRNAs in biotic and abiotic stress responses in crop plants. Applied Biochemistry and Biotechnology, 174(1), 93-115.

Kumar, V., Khare, T., Shriram, V., & Wani, S. H. (2018). Plant small RNAs: the essential epigenetic regulators of gene expression for salt-stress responses and tolerance. Plant Cell Reports, 37(1), 61-75.

Kunze, G., Zipfel, C., Robatzek, S., Niehaus, K., Boller, T., & Felix, G. (2004). The N terminus of bacterial elongation factor Tu elicits innate immunity in *Arabidopsis* plants. The Plant Cell, 16(12), 3496-3507.

Kutter, C., Schöb, H., Stadler, M., Meins, F., & Si-Ammour, A. (2007). MicroRNA-mediated regulation of stomatal development in *Arabidopsis*. The Plant Cell, 19(8), 2417-2429.

Kwenda, S., Birch, P. R., & Moleleki, L. N. (2016). Genome-wide identification of potato long intergenic noncoding RNAs responsive to *Pectobacterium carotovorum* subspecies *brasiliense* infection. BMC Genomics, 17(1), 614.

Kwon, T. (2016). Cytokinin response factor 2 positively regulates salicylic acid-mediated plant immunity in *Arabidopsis thaliana*. Plant Biotechnology, 33(3), 207-210.

Lahlali, R., Song, T., Chu, M., Yu, F., Kumar, S., Karunakaran, C., & Peng, G. (2017). Evaluating changes in cell-wall components associated with clubroot resistance using fourier transform infrared spectroscopy and RT-PCR. International Journal of Molecular Sciences, 18(10), 2058.

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. Nature Methods, 9(4), 357-359.

Lau, W. C., Rafii, M. Y., Ismail, M. R., Puteh, A., Latif, M. A., & Ramli, A. (2015). Review of functional markers for improving cooking, eating, and the nutritional qualities of rice. Frontiers in Plant Science, 6, 832.

Laule, O., Fürholz, A., Chang, H. S., Zhu, T., Wang, X., Heifetz, P. B., ... & Lange, M. (2003). Crosstalk between cytosolic and plastidial pathways of isoprenoid biosynthesis in *Arabidopsis thaliana*. Proceedings of the National Academy of Sciences, 100(11), 6866-6871.

Lawson, M. J., & Zhang, L. (2006). Distinct patterns of SSR distribution in the *Arabidopsis thaliana* and rice genomes. Genome Biology, 7(2), R14.1- R14.11.

Lemarié, S., Robert-Seilaniantz, A., Lariagon, C., Lemoine, J., Marnet, N., Jubault, M., ... & Gravot, A. (2015). Both the jasmonic acid and the salicylic acid pathways contribute to resistance to the biotrophic clubroot agent *Plasmodiophora brassicae* in *Arabidopsis*. Plant and Cell Physiology, 56(11), 2158-2168.

Le Roy, J., Huss, B., Creach, A., Hawkins, S., & Neutelings, G. (2016). Glycosylation is a major regulator of phenylpropanoid availability and biological activity in plants. Frontiers in Plant Science, 7, 735.

Li, X., Clarke, J. D., Zhang, Y., & Dong, X. (2001). Activation of an *EDS1*-mediated R-gene pathway in the *snc1* mutant leads to constitutive, *NPR1*-independent pathogen resistance. Molecular Plant-Microbe Interactions, 14(10), 1131-1139.

Li, T., Liu, B., Spalding, M. H., Weeks, D. P., & Yang, B. (2012). High-efficiency TALEN-based gene editing produces disease-resistant rice. Nature biotechnology, 30(5), 390-392.

Li, L., Eichten, S. R., Shimizu, R., Petsch, K., Yeh, C. T., Wu, W., ... & Evans, M. M. (2014). Genome-wide discovery and characterization of maize long non-coding RNAs. Genome Biology, 15(2), R40.1-R40.15.

Li, S. H., Dudler, R., Ji, R., Yong, M. L., Wang, Z. Y., & Hu, D. W. (2014). Long non-coding RNAs in wheat are related to its susceptibility to powdery mildew. Biologia Plantarum, 58(2), 296-304.

Li, Z., Huang, C., Bao, C., Chen, L., Lin, M., Wang, X., ... & Zhu, P. (2015). Exon-intron circular RNAs regulate transcription in the nucleus. Nature Structural & Molecular Biology, 22(3), 256.

Li, X., Zhang, S., Bai, J., & He, Y. (2016). Tuning growth cycles of Brassica crops via natural antisense transcripts of *Br FLC*. Plant Biotechnology Journal, 14(3), 905-914.

Li, W. Q., Jia, Y. L., Liu, F. Q., Wang, F. Q., Fan, F. J., Wang, J., ... & Yang, J. (2018). Genome-wide identification and characterization of long non-coding RNAs responsive to *Dickeya zeae* in rice. RSC advances, 8(60), 34408-34417.

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, 207.

Li, R., Fu, D., Zhu, B., Luo, Y., & Zhu, H. (2018). CRISPR/Cas9-mediated mutagenesis of *lncRNA1459* alters tomato fruit ripening. The Plant Journal, 94(3), 513-524.

Li, L., Long, Y., Li, H., & Wu, X. (2020). Comparative transcriptome analysis reveals key pathways and hub genes in rapeseed during the early stage of *Plasmodiophora brassicae* infection. Frontiers in Genetics, *10*, 1275.

Liabeuf, D., Sim, S. C., & Francis, D. M. (2018). Comparison of marker-based genomic estimated breeding values and phenotypic evaluation for selection of bacterial spot resistance in tomato. Phytopathology, 108(3), 392-401.

Liang, Y., Srivastava, S., Rahman, M. H., Strelkov, S. E., & Kav, N. N. (2008). Proteome changes in leaves of *Brassica napus* L. as a result of *Sclerotinia sclerotiorum* challenge. Journal of Agricultural and Food Chemistry, 56(6), 1963-1976.

Liang, Y. S., Ermawati, N., Cha, J. Y., Jung, M. H., Su'udi, M., Kim, M. G., ... & Son, D. (2010). Overexpression of an AP2/ERF-type transcription factor *CRF5* confers pathogen resistance to *Arabidopsis* plants. Journal of the Korean Society for Applied Biological Chemistry, 53(2), 142-148.

Liang, G., Yang, Y., Niu, G., Tang, Z., & Li, K. (2017). Genome-wide profiling of *Sus scrofa* circular RNAs across nine organs and three developmental stages. DNA Research, 24(5), 523-535.

Liao, B., Lee, S. Y., Meng, K., Yin, Q., Huang, C., Fan, Q., ... & Chen, S. (2019). Characterization and novel EST-SSR marker development of an important Chinese medicinal plant, *Morinda officinalis* How (Rubiaceae). Biotechnology & Biotechnological Equipment, 33(1), 1311-1318.

Lin, M. F., Jungreis, I., & Kellis, M. (2011). PhyloCSF: a comparative genomics method to distinguish protein coding and non-coding regions. Bioinformatics, 27(13), i275-i282.

Lin, L., Allemekinders, H., Dansby, A., Campbell, L., Durance-Tod, S., Berger, A., & Jones, P. J. (2013). Evidence of health benefits of canola oil. Nutrition reviews, 71(6), 370-385.

Lindsay, M. A., Griffiths-Jones, S., Wight, M., & Werner, A. (2013). The functions of natural antisense transcripts. Essays in biochemistry, 54, 91-101.

Liu, F., Marquardt, S., Lister, C., Swiezewski, S., & Dean, C. (2010). Targeted 3′ processing of antisense transcripts triggers *Arabidopsis FLC* chromatin silencing. Science, 327(5961), 94-97.

Liu, J., Jung, C., Xu, J., Wang, H., Deng, S., Bernad, L., ... & Chua, N. H. (2012). Genome-wide analysis uncovers regulation of long intergenic noncoding RNAs in *Arabidopsis*. The Plant Cell, 24(11), 4333-4345.

Liu, H., Guo, Z., Gu, F., Ke, S., Sun, D., Dong, S., ... & Liu, Y. (2017). 4-*coumarate*-CoA *ligase*-like gene *OsAAE3* negatively mediates the rice blast resistance, floret development and lignin biosynthesis. Frontiers in Plant Science, 7, 2041.

Liu, J., Ben, Q., Lu, E., He, X., Yang, X., Ma, J., ... & Wang, H. (2018). Long noncoding RNA *PANDAR* blocks *CDKN1A* gene transcription by competitive interaction with p53 protein in gastric cancer. Cell Death & Disease, 9(2), 1-13.

Liu, Q., Luo, L., & Zheng, L. (2018). Lignins: biosynthesis and biological functions in plants. International Journal of Molecular Sciences, 19(2), 335.

Liu, X., Li, D., Zhang, D., Yin, D., Zhao, Y., Ji, C., ... & Hu, S. (2018). A novel antisense long noncoding RNA, *TWISTED LEAF*, maintains leaf blade flattening by regulating its associated sense *R2R3-MYB* gene in rice. New Phytologist, 218(2), 774-788.

Liu, T., Zhang, L., Chen, G., & Shi, T. (2017). Identifying and characterizing the circular RNAs during the lifespan of *Arabidopsis* leaves. Frontiers in Plant Science, 8, 1278.

Liu, X., Hao, L., Li, D., Zhu, L., & Hu, S. (2015). Long non-coding RNAs and their biological roles in plants. Genomics, Proteomics & Bioinformatics, 13(3), 137-147.

Liu, X., Zhao, C., Yang, L., Zhang, Y., Wang, Y., Fang, Z., & Lv, H. (2020). Genome-Wide identification, expression profile of the TIFY gene family in *Brassica oleracea var. capitata,* and their divergent response to various pathogen infections and phytohormone treatments. Genes, 11(2), 127.

López Sánchez, A., Stassen, J. H., Furci, L., Smith, L. M., & Ton, J. (2016). The role of DNA (de) methylation in immune responsiveness of *Arabidopsis*. The Plant Journal, 88(3), 361-374.

Lorenz, R., Bernhart, S. H., Zu Siederdissen, C. H., Tafer, H., Flamm, C., Stadler, P. F., & Hofacker, I. L. (2011). ViennaRNA Package 2.0. Algorithms for Molecular Biology, 6(1), 26.

Lu, T., Cui, L., Zhou, Y., Zhu, C., Fan, D., Gong, H., ... & Luo, J. (2015). Transcriptome-wide investigation of circular RNAs in rice. RNA, 21(12), 2076-2087.

Lu, D., Wu, S., Gao, X., Zhang, Y., Shan, L., & He, P. (2010). A receptor-like cytoplasmic kinase, *BIK1*, associates with a flagellin receptor complex to initiate plant innate immunity. Proceedings of the National Academy of Sciences, 107(1), 496-501.

Luan, M., Xu, M., Lu, Y., Zhang, Q., Zhang, L., Zhang, C., ... & Wang, L. (2014). Family-wide survey of miR169s and NF-YAs and their expression profiles response to abiotic stress in maize roots. PloS one, 9(3), e91369.

Luo, Y., Dong, D., Su, Y., Wang, X., Peng, Y., Peng, J., & Zhou, C. (2018). Transcriptome analysis of *Brassica juncea var*. *tumida Tsen* responses to *Plasmodiophora brassicae* primed by the biocontrol strain *Zhihengliuella aestuarii*. Functional & Integrative Genomics, 18(3), 301-314.

Lyons, R., Rusu, A., Stiller, J., Powell, J., Manners, J. M., & Kazan, K. (2015). Investigating the association between flowering time and defense in the *Arabidopsis thaliana-Fusarium oxysporum* interaction. PloS one, 10(6), e0127699.

Lyu, J. I., Min, S. R., Lee, J. H., Lim, Y. H., Kim, J. K., Bae, C. H., & Liu, J. R. (2013). Overexpression of a trehalose-6-phosphate synthase/phosphatase fusion gene enhances tolerance and photosynthesis during drought and salt stress without growth aberrations in tomato. Plant Cell, Tissue and Organ Culture (PCTOC), 112(2), 257-262.

Mackey, D., Holt III, B. F., Wiig, A., & Dangl, J. L. (2002). *RIN4* interacts with Pseudomonas syringae type III effector molecules and is required for *RPM1*-mediated resistance in *Arabidopsis*. Cell, 108(6), 743-754.

Macfarlane, I. (1970). Germination of resting spores of *Plasmodiophora brassicae*. Transactions of the British Mycological Society, 55(1), 97-112.

Malinowski, R., Novák, O., Borhan, M. H., Spíchal, L., Strnad, M., & Rolfe, S. A. (2016). The role of cytokinins in clubroot disease. European Journal of Plant Pathology, 145(3), 543-557.

Mannur, D. M., Babbar, A., Thudi, M., Sabbavarapu, M. M., Roorkiwal, M., Sharanabasappa, B. Y., ... & Chamarthi, S. K. (2019). Super Annigeri 1 and improved JG 74: two Fusarium wilt-resistant introgression lines developed using marker-assisted backcrossing approach in chickpea (*Cicer arietinum L.*). Molecular Breeding, 39(1), 2.

Manzanares-Dauleux, M. J., Divaret, I., Baron, F., & Thomas, G. (2001). Assessment of biological and molecular variability between and within field isolates of *Plasmodiophora brassicae*. Plant Pathology, 50(2), 165-173.

Mao, X., Cai, T., Olyarchuk, J. G., & Wei, L. (2005). Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. Bioinformatics, 21(19), 3787-3793.

Marino, D., Peeters, N., & Rivas, S. (2012). Ubiquitination during plant immune signaling. Plant physiology, *160*(1), 15-27.

Marone, D., Russo, M. A., Laidò, G., De Leonardis, A. M., & Mastrangelo, A. M. (2013). Plant nucleotide binding site–leucine-rich repeat (NBS-LRR) genes: active guardians in host defense responses. International Journal of Molecular Sciences, 14(4), 7302-7326.

Mattick, J. S., & Rinn, J. L. (2015). Discovery and annotation of long noncoding RNAs. Nature Structural & Molecular Biology, 22(1), 5-7.

Matsui, A., & Seki, M. (2019). The involvement of long noncoding RNAs in response to plant stress. in plant long non-coding RNAs (pp. 151-171). Humana Press, New York, NY.

Matzke, M. A., & Mosher, R. A. (2014). RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. Nature Reviews Genetics, 15(6), 394-408.

Mauch-Mani, B., Baccelli, I., Luna, E., & Flors, V. (2017). Defense priming: an adaptive part of induced resistance. Annual Review of Plant Biology, 68, 485-512.

McDonald, M. R., Sharma, K., Gossen, B. D., Deora, A., Feng, J., & Hwang, S. F. (2014). The role of primary and secondary infection in host response to *Plasmodiophora brassicae*. Phytopathology, 104(10), 1078-1087.

Mei, C., Qi, M., Sheng, G., & Yang, Y. (2006). Inducible overexpression of a rice allene oxide

synthase gene increases the endogenous jasmonic acid level, PR gene expression, and host resistance to fungal infection. Molecular Plant-Microbe Interactions, 19(10), 1127-1137.

Mei, J., Guo, Z., Wang, J., Feng, Y., Ma, G., Zhang, C., ... & Chen, G. (2019). Understanding the resistance mechanism in *Brassica napus* to clubroot caused by *Plasmodiophora brassicae*. Phytopathology, 109(5), 810-818.

Memczak, S., Jens, M., Elefsinioti, A., Torti, F., Krueger, J., Rybak, A., ... & Loewer, A. (2013). Circular RNAs are a large class of animal RNAs with regulatory potency. Nature, 495(7441), 333-338.

Meng, Y., Shao, C., Wang, H., & Jin, Y. (2012). Target mimics: an embedded layer of microRNA-involved gene regulatory networks in plants. BMC Genomics, 13(1), 197.

Meng, X., Li, X., Zhang, P., Wang, J., Zhou, Y., & Chen, M. (2017). Circular RNA: an emerging key player in RNA world. Briefings in Bioinformatics, 18(4), 547-557.

Miller, T. E., Reader, S. M., Ainsworth, C. C., & Summers, R. W. (1988, December). The introduction of a major gene for resistance to powdery mildew of wheat, *Erysiphe graminis f. sp. tritici*, from *Aegilops speltoides* into wheat, *Triticum aestivum*. In Cereal breeding related to integrated cereal production: proceedings of the EUCARPIA conference. Pudoc, Wageningen, The Netherlands (pp. 179-183).

Mizobuchi, R., Sato, H., Fukuoka, S., Tanabata, T., Tsushima, S., Imbe, T., & Yano, M. (2013). Mapping a quantitative trait locus for resistance to bacterial grain rot in rice. Rice, 6(1), 13.

Molla, K. A., Azharudheen, T. M., Ray, S., Sarkar, S., Swain, A., Chakraborti, M., ... & Mukherjee, A. K. (2019). Novel biotic stress responsive candidate gene based SSR (cgSSR) markers from rice. Euphytica, 215(2), 17.

Mondal, T. K., & Ganie, S. A. (2014). Identification and characterization of salt responsive miRNA-SSR markers in rice (*Oryza sativa*). Gene, 535(2), 204-209.

Morris, K. V., & Mattick, J. S. (2014). The rise of regulatory RNA. Nature Reviews Genetics, 15(6), 423-437.

Moose, S. P., & Mumm, R. H. (2008). Molecular plant breeding as the foundation for 21st century crop improvement. Plant Physiology, 147(3), 969-977.

Morais, T. P., Zaini, P. A., Chakraborty, S., Gouran, H., Carvalho, C. P., Almeida-Souza, H. O., ... & Nascimento, R. (2019). The plant-based chimeric antimicrobial protein SlP14a-PPC20 protects tomato against bacterial wilt disease caused by Ralstonia solanacearum. Plant Science, 280, 197-205.

Morkunas, I., Borek, S., Formela, M., & Ratajczak, L. (2012). Plant responses to sugar starvation. Carbohydrates-Comprehensive Studies on Glycobiology and Glycotechnology, IntechOpen, Rijeka, Croatia, pp. 409-438.

Morkunas, I., & Ratajczak, L. (2014). The role of sugar signaling in plant defense responses against fungal pathogens. Acta Physiologiae Plantarum, 36(7), 1607-1619.

Mostofa, M. G., Hossain, M. A., Fujita, M., & Tran, L. S. P. (2015). Physiological and biochemical mechanisms associated with trehalose-induced copper-stress tolerance in rice. Scientific reports, *5*, 11433.

Muthusamy, M., Uma, S., Backiyarani, S., & Saraswathi, M. S. (2015). Genome-wide screening for novel, drought stress-responsive long non-coding RNAs in drought-stressed leaf transcriptome of drought-tolerant and-susceptible banana (*Musa spp*) cultivars using Illumina high-throughput sequencing. Plant Biotechnology Reports, 9(5), 279-286.

Muthusamy, M., Uma, S., Suthanthiram, B., Saraswathi, M. S., & Chandrasekar, A. (2019). Genome-wide identification of novel, long non-coding RNAs responsive to *Mycosphaerella eumusae* and *Pratylenchus coffeae* infections and their differential expression patterns in disease-resistant and sensitive banana cultivars. Plant Biotechnology Reports, 13(1), 73-83.

Mutuku, M., & Nose, A. (2010). *Rhizoctonia solani* infection in two rice lines increases mRNA expression of metabolic enzyme genes in glycolytic, oxidative pentose phosphate pathways and secondary metabolism. Tropical Agriculture and Development, 54(4), 119-131.

Mutuku, J. M., & Nose, A. (2012). Changes in the contents of metabolites and enzyme activities in rice plants responding to *Rhizoctonia solani* Kuhn infection: activation of glycolysis and connection to phenylpropanoid pathway. Plant and Cell Physiology, *53*(6), 1017-1032.

Ncube, B., & Van Staden, J. (2015). Tilting plant metabolism for improved metabolite biosynthesis and enhanced human benefit. Molecules, *20*(7), 12698-12731.

Nejat, N., & Mantri, N. (2018). Emerging roles of long non-coding RNAs in plant response to biotic and abiotic stresses. Critical Reviews in Biotechnology, 38(1), 93-105.

Nguyen, K., Grondin, A., Courtois, B., & Gantet, P. (2019). Next-generation sequencing accelerates crop gene discovery. Trends in Plant Science, 24(3), 263-274.

Ning, Y., Wang, Y., Fang, Z., Zhuang, M., Zhang, Y., Lv, H., ... & Yang, L. (2019). Comparative transcriptome analysis of cabbage (*Brassica oleracea* var. *capitata*) infected by *Plasmodiophora brassicae* reveals drastic defense response at secondary infection stage. Plant and Soil, *443*(1-2), 167-183.

Nekrasov, V., Wang, C., Win, J., Lanz, C., Weigel, D., & Kamoun, S. (2017). Rapid generation of a transgene-free powdery mildew resistant tomato by genome deletion. Scientific Reports, 7(1), 1-6.

Nie, Y., Zhou, L., Wang, H., Chen, N., Jia, L., Wang, C., ... & Li, H. (2019). Profiling the epigenetic interplay of lncRNA *RUNXOR* and oncogenic *RUNX1* in breast cancer cells by gene in situ cis-activation. American Journal of Cancer Research, 9(8), 1635.

Noman, A., Aqeel, M., & Lou, Y. (2019). PRRs and NB-LRRs: from signal perception to activation of plant innate immunity. International Journal of Molecular Sciences, 20(8), 1882.

Oba, K., Kondo, K., Doke, N., & Uritani, I. (1985). Induction of 3-hydroxy-3-methylglutaryl CoA reductase in potato tubers after slicing, fungal infection or chemical treatment, and some properties of the enzyme. Plant and cell physiology, *26*(5), 873-880.

O'donnell, P. J., Calvert, C., Atzorn, R., Wasternack, C. H. M. O., Leyser, H. M. O., & Bowles, D. J. (1996). Ethylene as a signal mediating the wound response of tomato plants. Science, 274(5294), 1914-1917.

Ohtani, M. (2018). Plant snRNP biogenesis: a perspective from the nucleolus and Cajal bodies. Frontiers in Plant Science, 8, 2184.

Okada, K., Abe, H., & Arimura, G. I. (2015). Jasmonates induce both defense responses and communication in monocotyledonous and dicotyledonous plants. Plant and Cell Physiology, *56*(1), 16-27.

Omar, A. A., Murata, M. M., El-Shamy, H. A., Graham, J. H., & Grosser, J. W. (2018). Enhanced resistance to citrus canker in transgenic mandarin expressing *Xa21* from rice. Transgenic Research, 27(2), 179-191.

Onaga, G., and Wydra, K., (2016). Advances in plant tolerance to biotic stresses, Plant Genomics, Ibrokhim Y. Abdurakhmonov, IntechOpen, Rijeka, Croatia, pp. 167–228.

Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P., & Lim, W. A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell, 152(5), 1173-1183.

Quek, X. C., Thomson, D. W., Maag, J. L., Bartonicek, N., Signal, B., Clark, M. B., ... & Dinger, M. E. (2015). lncRNAdb v2. 0: expanding the reference database for functional long noncoding RNAs. Nucleic Acids Research, 43(D1), D168-D173.

Takagi, H., Abe, A., Yoshida, K., Kosugi, S., Natsume, S., Mitsuoka, C., ... & Innan, H. (2013). QTL-seq: rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. The Plant Journal, 74(1), 174-183.

Tamang, T., Park, J., Kakeshpour, T., Valent, B., Jia, Y., Want, G., & Park, S. (2018). Development of selectable marker-free cisgenic rice plants expressing a blast resistance gene *Pi9*. World Con-gress on In vitro Biology, 54, 544.

Torquebiau, E., Tissier, J., & Grosclaude, J. Y. (2016). How climate change reshuffles the cards for agriculture. In Climate Change and Agriculture Worldwide (pp. 1-16). Springer, Dordrecht.

Pachauri, R. K., & Reisinger, A. (2008). Climate change 2007. Synthesis report. Contribution of Working Groups I, II and III to the fourth assessment report. IPCC, Geneva, Switzerland.

Padmanabhan, S. Y. (1973). The great Bengal famine. Annual Review of Phytopathology, 11(1), 11-24.

Parker, J. E., Coleman, M. J., Szabò, V., Frost, L. N., Schmidt, R., van der Biezen, E. A., ... &
Jones, J. D. (1997). The *Arabidopsis* downy mildew resistance gene *RPP5* shares similarity to the toll and interleukin-1 receptors with N and L6. The Plant Cell, *9*(6), 879-894.

Parthasarathy, A., Cross, P. J., Dobson, R. C., Adams, L. E., Savka, M. A., & Hudson, A. O. (2018). A three-ring circus: metabolism of the three proteogenic aromatic amino acids and their role in the health of plants and animals. Frontiers in Molecular Biosciences, 5, 29.

Patil, S. V., & Kumudini, B. S. (2019). Seed priming induced blast disease resistance in finger millet plants through phenylpropanoid metabolic pathway. Physiological and Molecular Plant Pathology, 108, 101428.

Peng, G., Lahlali, R., Hwang, S. F., Pageau, D., Hynes, R. K., McDonald, M. R., ... & Strelkov, S. E. (2014). Crop rotation, cultivar resistance, and fungicides/biofungicides for managing clubroot (*Plasmodiophora brassicae*) on canola. Canadian Journal of Plant Pathology, 36(sup1), 99-112.

Peng, Z., He, S., Gong, W., Sun, J., Pan, Z., Xu, F., ... & Du, X. (2014). Comprehensive analysis of differentially expressed genes and transcriptional regulation induced by salt stress in two contrasting cotton genotypes. BMC genomics, 15(1), 760.

Peng, G., Pageau, D., Strelkov, S. E., Gossen, B. D., Hwang, S. F., & Lahlali, R. (2015). A> 2-year

crop rotation reduces resting spores of *Plasmodiophora brassicae* in soil and the impact of

clubroot on canola. European Journal of Agronomy, 70, 78-84.

Peng, A., Chen, S., Lei, T., Xu, L., He, Y., Wu, L., ... & Zou, X. (2017). Engineering canker-resistant plants through CRISPR/Cas9-targeted editing of the susceptibility gene *CsLOB1* promoter in citrus. Plant biotechnology journal, 15(12), 1509-1519.

Peng, Y., Gossen, B. D., Huang, Y., Al-Daoud, F., & McDonald, M. R. (2019). Development of *Plasmodiophora brassicae* in the root cortex of cabbage over time. European Journal of Plant Pathology, 154(3), 727-737.

Phukan, U. J., Jeena, G. S., & Shukla, R. K. (2016). WRKY transcription factors: molecular regulation and stress responses in plants. Frontiers in plant science, *7*, 760.

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(8), 1458-1465.

Piao, H. Y., Guo, S., Wang, Y., & Zhang, J. (2020). Exosomal long non-coding RNA *CEBPA-AS1* inhibits tumor apoptosis and functions as a non-invasive biomarker for diagnosis of gastric cancer. OncoTargets and Therapy, 13, 1365-1374.

Pickering, R. A., Hill, A. M., Michel, M., & Timmerman-Vaughan, G. M. (1995). The transfer of a powdery mildew resistance gene from *Hordeum bulbosum L* to barley (*H. vulgare L.*) chromosome 2 (2I). Theoretical and Applied Genetics, 91(8), 1288-1292.

Pinson, S. R., Shahjahan, A. K., Rush, M. C., & Groth, D. E. (2010). Bacterial panicle blight resistance QTLs in rice and their association with other disease resistance loci and heading date. Crop Science, 50(4), 1287-1297.

Poczai, P., Varga, I., Laos, M., Cseh, A., Bell, N., Valkonen, J. P., & Hyvönen, J. (2013). Advances in plant gene-targeted and functional markers: a review. Plant Methods, *9*(1), 6.

Prerostova, S., Dobrev, P. I., Konradyova, V., Knirsch, V., Gaudinova, A., Kramna, B., ... & Vankova, R. (2018). Hormonal responses to *Plasmodiophora brassicae* infection in *Brassica napus* cultivars differing in their pathogen resistance. International Journal of Molecular Sciences, 19(12), 4024.

Primard-Brisset, C., Poupard, J. P., Horvais, R., Eber, F., Pelletier, G., Renard, M., & Delourme, R. (2005). A new recombined double low restorer line for the Ogu-INRA cms in rapeseed (*Brassica napus L.*). Theoretical and Applied Genetics, 111(4), 736-746.

Punta, M., Coggill, P. C., Eberhardt, R. Y., Mistry, J., Tate, J., Boursnell, C., ... & Heger, A. (2014). The Pfam protein families database: nucleic acids research. Database, (2012), 40.

Rahman, H., Shakir, A., & Jakir Hasan, M. (2011). Breeding for clubroot resistant spring canola (*Brassica napus* L.) for the Canadian prairies: Can the European winter canola cv. Mendel be used as a source of resistance?. Canadian Journal of Plant Science, 91(3), 447-458.

Rahman, H., Harwood, J., & Weselake, R. (2013). Increasing seed oil content in Brassica species through breeding and biotechnology. Lipid Technology, 25(8), 182-185.

Rahman, H., Peng, G., Yu, F., Falk, K. C., Kulkarni, M., & Selvaraj, G. (2014). Genetics and breeding for clubroot resistance in Canadian spring canola (*Brassica napus L.*). Canadian Journal of Plant Pathology, 36(sup1), 122-134.

Rahman, H. (2016). UA AlfaGold Clearfield herbicide-tolerant spring *Brassica napus* canola developed from winter× spring canola cross. Canadian Journal of Plant Science, 97(1), 144-146.

Rai, M. I., Alam, M., Lightfoot, D. A., Gurha, P., & Afzal, A. J. (2019). Classification and experimental identification of plant long non-coding RNAs. Genomics, 111(5), 997-1005.

Ramirez-Prado, J. S., Abulfaraj, A. A., Rayapuram, N., Benhamed, M., & Hirt, H. (2018). Plant immunity: from signaling to epigenetic control of defense. Trends in Plant Science, 23(9), 833-844.

Ranjan, A., Westrick, N. M., Jain, S., Piotrowski, J. S., Ranjan, M., Kessens, R., ... & Kabbage, M. (2019). Resistance against *Sclerotinia sclerotiorum* in soybean involves a reprogramming of the phenylpropanoid pathway and up-regulation of antifungal activity targeting ergosterol biosynthesis. Plant Biotechnology Journal, 17(8), 1567-1581.

Rashotte, A. M., Mason, M. G., Hutchison, C. E., Ferreira, F. J., Schaller, G. E., & Kieber, J. J. (2006). A subset of *Arabidopsis AP2* transcription factors mediates cytokinin responses in concert with a two-component pathway. Proceedings of the National Academy of Sciences, 103(29), 11081-11085.

Raymer, P. L. (2002). Canola: an emerging oilseed crop. Trends in new crops and new uses, 1, 122-

126.

Reader, S. M., & Miller, T. E. (1991). The introduction into bread wheat of a major gene for resistance to powdery mildew from wild emmer wheat. Euphytica, 53(1), 57-60.

Reignault, P. H., Cogan, A., Muchembled, J., Lounes-Hadj Sahraoui, A., Durand, R., & Sancholle, M. (2001). Trehalose induces resistance to powdery mildew in wheat. New Phytologist, 149(3), 519-529.

Rejeb, I. B., Pastor, V., & Mauch-Mani, B. (2014). Plant responses to simultaneous biotic and abiotic stress: molecular mechanisms. Plants, *3*(4), 458-475.

Rekhter, D., Lüdke, D., Ding, Y., Feussner, K., Zienkiewicz, K., Lipka, V., ... & Feussner, I. (2019). Isochorismate-derived biosynthesis of the plant stress hormone salicylic acid. Science, 365(6452), 498-502.

Rempel, C. B., Hutton, S. N., & Jurke, C. J. (2014). Clubroot and the importance of canola in Canada. Canadian Journal of Plant Pathology, *36*(sup1), 19-26.

Ricroch, A. E., & Hénard-Damave, M. C. (2016). Next biotech plants: new traits, crops, developers and technologies for addressing global challenges. Critical Reviews in Biotechnology, 36(4), 675-690.

"RNAcentral: a hub of information for non-coding RNA sequences." Nucleic acids research 47, no. D1 (2019): D221-D229.

Robin, A. H. K., Hossain, M. R., Kim, H. T., Nou, I. S., & Park, J. I. (2019). Role of cytokinins in clubroot disease development. Plant Breeding and Biotechnology, 7(2), 73-82.

Röhrig, H., Schmidt, J., Miklashevichs, E., Schell, J., & John, M. (2002). Soybean *ENOD40* encodes two peptides that bind to sucrose synthase. Proceedings of the National Academy of Sciences, 99(4), 1915-1920.

Rojas, C. M., Senthil-Kumar, M., Tzin, V., & Mysore, K. (2014). Regulation of primary plant metabolism during plant-pathogen interactions and its contribution to plant defense. Frontiers in Plant Science, 5, 17.

Rolfe, S. A., Strelkov, S. E., Links, M. G., Clarke, W. E., Robinson, S. J., Djavaheri, M., ... & Taheri, A. (2016). The compact genome of the plant pathogen *Plasmodiophora brassicae* is adapted to intracellular interactions with host Brassica spp. BMC genomics, *17*(1), 272.

Roux, M., Schwessinger, B., Albrecht, C., Chinchilla, D., Jones, A., Holton, N., ... & Zipfel, C. (2011). The *Arabidopsis* leucine-rich repeat receptor–like kinases *BAK1*/*SERK3* and *BKK1*/*SERK4* are required for innate immunity to hemibiotrophic and biotrophic pathogens. The Plant Cell, 23(6), 2440-2455.

Ruan, J., Zhou, Y., Zhou, M., Yan, J., Khurshid, M., Weng, W., ... & Zhang, K. (2019). Jasmonic acid signaling pathway in plants. International Journal of Molecular Sciences, 20(10), 2479.

Rutter, B. D., & Innes, R. W. (2017). Extracellular vesicles isolated from the leaf apoplast carry stress-response proteins. Plant Physiology, 173(1), 728-741.

Saeidi, H., Rahiminejad, M. R., & Heslop-Harrison, J. S. (2008). Retroelement insertional polymorphisms, diversity and phylogeography within diploid, D-genome *Aegilops tauschii* (*Triticeae*, *Poaceae*) sub-taxa in Iran. Annals of Botany, 101(6), 855-861.

Santos-Sánchez, N. F., Salas-Coronado, R., Hernández-Carlos, B., & Villanueva-Cañongo, C. (2019). Shikimic acid pathway in biosynthesis of phenolic compounds. In *Plant* Physiological Aspects of Phenolic Compounds. IntechOpen, London, UK, 2019.

Saini, K., AbdElgawad, H., Markakis, M. N., Schoenaers, S., Asard, H., Prinsen, E., ... & Vissenberg, K. (2017). Perturbation of auxin homeostasis and signaling by *PINOID* overexpression induces stress responses in *Arabidopsis*. Frontiers in Plant Science, 8, 1308.

Sahu, P. P., Pandey, G., Sharma, N., Puranik, S., Muthamilarasan, M., & Prasad, M. (2013). Epigenetic mechanisms of plant stress responses and adaptation. Plant Cell Reports, 32(8), 1151-1159.

Sahu, S., Rao, A. R., Pandey, J., Gaikwad, K., Ghoshal, S., & Mohapatra, T. (2018). Genome-wide identification and characterization of lncRNAs and miRNAs in cluster bean (*Cyamopsis tetragonoloba*). Gene, 667, 112-121.

Samal, K. C., & Rout, G. R. (2018). Genetic improvement of vegetables using transgenic technology. In Genetic Engineering of Horticultural Crops (pp. 193-224). Academic Press.

Sanchez, A. C., Brar, D. S., Huang, N., Li, Z., & Khush, G. S. (2000). Sequence tagged site marker-assisted selection for three bacterial blight resistance genes in rice. Crop Science, 40(3), 792-797.

Schaart, J. G., van de Wiel, C. C., Lotz, L. A., & Smulders, M. J. (2016). Opportunities for products of new plant breeding techniques. Trends in Plant Science, 21(5), 438-449.

Schluepmann, H., Van Dijken, A., Aghdasi, M., Wobbes, B., Paul, M., & Smeekens, S. (2004). Trehalose mediated growth inhibition of *Arabidopsis* seedlings is due to trehalose-6-phosphate accumulation. Plant physiology, *135*(2), 879-890.

Schmittgen, T. D., & Livak, K. J. (2008). Analyzing real-time PCR data by the comparative C T method. Nature Protocols, 3(6), 1101-1108.

Schouten, H. J., Krens, F. A., & Jacobsen, E. (2006). Cisgenic plants are similar to traditionally bred plants. EMBO Reports, 7(8), 750-753.

Schweighofer, A., Hirt, H., & Meskiene, I. (2004). Plant PP2C phosphatases: emerging functions in stress signaling. Trends in Plant Science, 9(5), 236-243.

Seo, J. S., Sun, H. X., Park, B. S., Huang, C. H., Yeh, S. D., Jung, C., & Chua, N. H. (2017). *ELF18-INDUCED LONG-NONCODING RNA* associates with mediator to enhance expression of innate immune response genes in *Arabidopsis*. The Plant Cell, 29(5), 1024-1038.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., ... & Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Research, 13(11), 2498-2504.

Sharma, N., Rahman, M. H., Strelkov, S., Thiagarajah, M., Bansal, V. K., & Kav, N. N. (2007). Proteome-level changes in two *Brassica napus* lines exhibiting differential responses to the fungal pathogen *Alternaria brassicae*. Plant Science, 172(1), 95-110.

Sharma, S., Taneja, M., Tyagi, S., Singh, K., & Upadhyay, S. K. (2017). Survey of high throughput RNA-Seq data reveals potential roles for lncRNAs during development and stress response in bread wheat. Frontiers in Plant Science, 8, 1019.

Shen, D., Suhrkamp, I., Wang, Y., Liu, S., Menkhaus, J., Verreet, J. A., ... & Cai, D. (2014). Identification and characterization of micro RNAs in oilseed rape (*Brassica napus*) responsive to infection with the pathogenic fungus *Verticillium longisporum* using Brassica AA (*Brassica rapa*) and CC (*Brassica oleracea*) as reference genomes. New Phytologist, 204(3), 577-594.

Shen, E., Zhu, X., Hua, S., Chen, H., Ye, C., Zhou, L., ... & Chen, X. (2018). Genome-wide identification of oil biosynthesis-related long non-coding RNAs in allopolyploid *Brassica napus*. BMC Genomics, 19(1), 1-13.

Shin, D., Moon, S. J., Han, S., Kim, B. G., Park, S. R., Lee, S. K., ... & Yi, B. Y. (2011). Expression of *StMYB1R*-1, a novel potato single MYB-like domain transcription factor, increases drought tolerance. Plant Physiology, 155(1), 421-432.

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

Singh, S., Sidhu, J. S., Huang, N., Vikal, Y., Li, Z., Brar, D. S., ... & Khush, G. S. (2001). Pyramiding three bacterial blight resistance genes (*xa5, xa13 and Xa21*) using marker-assisted selection into indica rice cultivar PR106. Theoretical and Applied Genetics, 102(6-7), 1011-1015.

Singh, V., Louis, J., Ayre, B. G., Reese, J. C., & Shah, J. (2011). Trehalose phosphate synthase11-dependent trehalose metabolism promotes *Arabidopsis thaliana* defense against the phloem-feeding insect *Myzus persicae*. The Plant Journal, 67(1), 94-104.

Singh, A. K., Chaurasia, S., Kumar, S., Singh, R., Kumari, J., Yadav, M. C., ... & Jacob, S. R.

(2018). Identification, analysis, and development of salt responsive candidate gene based SSR markers in wheat. BMC Plant Biology, 18(1), 249.

Somssich, I. E., & Hahlbrock, K. (1998). Pathogen defence in plants—a paradigm of biological complexity. Trends in Plant Science, 3(3), 86-90.

Song, T., Chu, M., Lahlali, R., Yu, F., & Peng, G. (2016). Shotgun label-free proteomic analysis of clubroot (*Plasmodiophora brassicae*) resistance conferred by the gene *Rcr1* in *Brassica rapa*. Frontiers in Plant Science, 7, 1013.

Slaughter, A., Daniel, X., Flors, V., Luna, E., Hohn, B., & Mauch-Mani, B. (2012). Descendants of primed *Arabidopsis* plants exhibit resistance to biotic stress. Plant Physiology, 158(2), 835-843.

Spanu, P. D., & Panstruga, R. (2017). Biotrophic plant-microbe interactions. Frontiers in Plant Science, 8, 192.

Stasolla, C., Katahira, R., Thorpe, T. A., & Ashihara, H. (2003). Purine and pyrimidine nucleotide metabolism in higher plants. Journal of Plant Physiology, 160(11), 1271-1295.

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

Strelkov, S. E., & Hwang, S. F. (2014). Clubroot in the Canadian canola crop: 10 years into the outbreak. Canadian Journal of Plant Pathology, 36(sup1), 27-36.

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(3), 517-529.

Strelkov, S. E., Hwang, S. F., Manolii, V. P., Cao, T., Fredua-Agyeman, R., Harding, M. W., ... & 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(2), 284-298.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(10), 1360-1377.

Summanwar, A., Basu, U., Rahman, H., & Kav, N. N. (2020). Non-coding RNAs as emerging targets for crop improvement. Plant Science, 110521.

Sun, L., Luo, H., Bu, D., Zhao, G., Yu, K., Zhang, C., ... & Zhao, Y. (2013). Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. Nucleic acids Research, 41(17), e166-e166.

Sun, Q., Csorba, T., Skourti-Stathaki, K., Proudfoot, N. J., & Dean, C. (2013). R-loop stabilization represses antisense transcription at the *Arabidopsis FLC* locus. Science, 340(6132), 619-621.

Sundin, G. W., Castiblanco, L. F., Yuan, X., Zeng, Q., & Yang, C. H. (2016). Bacterial disease management: challenges, experience, innovation and future prospects: Challenges in bacterial molecular plant pathology. Molecular Plant Pathology, 17(9), 1506-1518.

Suwabe, K., Tsukazaki, H., Iketani, H., Hatakeyama, K., Fujimura, M., Nunome, T., ... & Hirai, M. (2003). Identification of two loci for resistance to clubroot (*Plasmodiophora brassicae* Woronin) in *Brassica rapa* L. Theoretical and Applied Genetics, 107(6), 997-1002.

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(1), 309-319.

Swarbreck, D., Wilks, C., Lamesch, P., Berardini, T. Z., Garcia-Hernandez, M., Foerster, H., ... & Radenbaugh, A. (2007). The *Arabidopsis* Information Resource (TAIR): gene structure and function annotation. Nucleic Acids Research, 36(suppl_1), D1009-D1014.

Swiezewski, S., Liu, F., Magusin, A., & Dean, C. (2009). Cold-induced silencing by long antisense transcripts of an *Arabidopsis* Polycomb target. Nature, 462(7274), 799-802.

Szaker, H. M., Darkó, É., Medzhiradszky, A. R., Janda, T., Liu, H. C., Charng, Y. Y., & Csorba, T. (2019). miR824/AGAMOUS-LIKE16 module integrates recurring environmental heat stress changes to fine-tune post-stress development. Frontiers in Plant Science, 10, 1454.

Szcześniak, M. W., Rosikiewicz, W., & Makałowska, I. (2016). CANTATAdb: a collection of plant long non-coding RNAs. Plant and Cell Physiology, 57(1), e8-e8.

Takemoto, D., Shibata, Y., Ojika, M., Mizuno, Y., Imano, S., Ohtsu, M., ... & Camagna, M. (2018). Resistance to *Phytophthora infestans*: exploring genes required for disease resistance in Solanaceae plants. Journal of General Plant Pathology, 84(5), 312-320.

Tang, B., Hao, Z., Zhu, Y., Zhang, H., & Li, G. (2018). Genome-wide identification and functional analysis of circRNAs in *Zea mays*. PLoS One, 13(12), e0202375.

Tav, C., Tempel, S., Poligny, L., & Tahi, F. (2016). miRNAFold: a web server for fast miRNA precursor prediction in genomes. Nucleic Acids Research, 44(W1), W181-W184.

Tayeh, C., Randoux, B., Vincent, D., Bourdon, N., & Reignault, P. (2014). Exogenous trehalose induces defenses in wheat before and during a biotic stress caused by powdery mildew. Phytopathology, 104(3), 293-305.

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(1), 143-144.

Thimm, O., Bläsing, O., Gibon, Y., Nagel, A., Meyer, S., Krüger, P., ... & Stitt, M. (2004). MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. The Plant Journal, 37(6), 914-939.

Todd, C. D., & Polacco, J. C. (2006). *AtAAH* encodes a protein with allantoate amidohydrolase activity from *Arabidopsis thaliana*. Planta, 223(5), 1108-1113.

Tohge, T., Watanabe, M., Hoefgen, R., & Fernie, A. R. (2013). Shikimate and phenylalanine biosynthesis in the green lineage. Frontiers in Plant Science, 4, 62.

Tomita, H., Shimizu, M., Doullah, M. A. U., Fujimoto, R., & Okazaki, K. (2013). Accumulation of quantitative trait loci conferring broad-spectrum clubroot resistance in *Brassica oleracea*. Molecular Breeding, 32(4), 889-900.

Torrens-Spence, M. P., Bobokalonova, A., Carballo, V., Glinkerman, C. M., Pluskal, T., Shen, A., & Weng, J. K. (2019). *PBS3* and *EPS1* complete salicylic acid biosynthesis from isochorismate in *Arabidopsis*. Molecular Plant, 12(12), 1577-1586.

Toojinda, T., Baird, E., Booth, A., Broers, L., Hayes, P., Powell, W., ... & Young, G. (1998). Introgression of quantitative trait loci (QTLs) determining stripe rust resistance in barley: an example of marker-assisted line development. Theoretical and Applied Genetics, 96(1), 123-131.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., ... & Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nature Protocols, 7(3), 562-578.

Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., ... & Pachter, L. (2010). Transcript assembly and abundance estimation from RNA-Seq reveals thousands of new transcripts and switching among isoforms. Nature Biotechnology, 28(5), 511-515.

Treutter, D. (2006). Significance of flavonoids in plant resistance: a review. Environmental Chemistry Letters, 4(3), 147-157.

Trouvelot, S., Héloir, M. C., Poinssot, B., Gauthier, A., Paris, F., Guillier, C., ... & Adrian, M. (2014). Carbohydrates in plant immunity and plant protection: roles and potential application as foliar sprays. Frontiers in Plant Science, 5, 592.

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

"Growing at a Slower Pace, World Population Is Expected to Reach 9.7 Billion in 2050 and Could

Peak at Nearly 11 Billion around 2100 | UN DESA Department of Economic and Social Affairs." United Nations, United Nations, 17 June 2019, Retrieved May 23, 2020. www.un.org/development/desa/en/news/population/world-population-prospects-2019.html.

Van Rhijn, P., Fang, Y., Galili, S., Shaul, O., Atzmon, N., Wininger, S., ... & Fujishige, N. (1997). Expression of early nodulin genes in alfalfa mycorrhizae indicates that signal transduction pathways used in forming arbuscular mycorrhizae and Rhizobium-induced nodules may be conserved. Proceedings of the National Academy of Sciences, 94(10), 5 467-5472.

Van Loon, L. C., & Van Strien, E. A. (1999). The families of pathogenesis-related proteins, their activities, and comparative analysis of *PR-1* type proteins. Physiological and Molecular Plant Pathology, 55, 85-97.

Van Inghelandt, D., Melchinger, A. E., Martinant, J. P., & Stich, B. (2012). Genome-wide association mapping of flowering time and northern corn leaf blight (*Setosphaeria*

turcica) resistance in a vast commercial maize germplasm set. BMC Plant Biology, 12(1), 56.

Vasistha, N. K., Balasubramaniam, A., Mishra, V. K., Srinivasa, J., Chand, R., & Joshi, A. K. (2017). Molecular introgression of leaf rust resistance gene *Lr34* validates enhanced effect on resistance to spot blotch in spring wheat. Euphytica, 213(12), 262.

Veronese, P., Narasimhan, M. L., Stevenson, R. A., Zhu, J. K., Weller, S. C., Subbarao, K. V., & Bressan, R. A. (2003). Identification of a locus controlling *Verticillium* disease symptom response in *Arabidopsis thaliana*. The Plant Journal, 35(5), 574-587.

Verma, S. S., Rahman, M. H., Deyholos, M. K., Basu, U., & Kav, N. N. (2014). Differential expression of miRNAs in *Brassica napus* root following infection with *Plasmodiophora brassicae*. PLoS One, 9(1), e86648.

Vogt, T. (2010). Phenylpropanoid biosynthesis. Molecular Plant, 3(1), 2-20.

Voorrips, R. E., & Visser, D. L. (1993). Examination of resistance to clubroot in accessions of *Brassica oleracea* using a glasshouse seedling test. Netherlands Journal of Plant Pathology, 99(5-6), 269-276.

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

Wagner, G., Laperche, A., Lariagon, C., Marnet, N., Renault, D., Guitton, Y., ... & Gravot, A. (2019). Resolution of quantitative resistance to clubroot into QTL-specific metabolic modules. Journal of Experimental Botany, 70(19), 5375-5390.

Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. Nature Reviews Genetics, 10(1), 57-63.

Wang, K. C., Yang, Y. W., Liu, B., Sanyal, A., Corces-Zimmerman, R., Chen, Y., ... & Wysocka, J. (2011). A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. Nature, 472(7341), 120-124.

Wang, S., & Fu, J. (2011). Insights into auxin signaling in plant–pathogen interactions. Frontiers in Plant Science, 2, 74.

Wang, H., Nagegowda, D. A., Rawat, R., Bouvier-Navé, P., Guo, D., Bach, T. J., & Chye, M. L. (2012). Overexpression of *Brassica juncea* wild-type and mutant HMG-CoA synthase 1

in *Arabidopsis* up-regulates genes in sterol biosynthesis and enhances sterol production and stress tolerance. Plant Biotechnology Journal, 10(1), 31-42.

Wang, H., Chung, P. J., Liu, J., Jang, I. C., Kean, M. J., Xu, J., & Chua, N. H. (2014). Genome-wide identification of long noncoding natural antisense transcripts and their responses to light in *Arabidopsis*. Genome Research, 24(3), 444-453.

Wang, Y., Cheng, X., Shan, Q., Zhang, Y., Liu, J., Gao, C., & Qiu, J. L. (2014). Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. Nature Biotechnology, 32(9), 947-951.

Wang, Y., Fan, X., Lin, F., He, G., Terzaghi, W., Zhu, D., & Deng, X. W. (2014). *Arabidopsis* noncoding RNA mediates control of photomorphogenesis by red light. Proceedings of the National Academy of Sciences, 111(28), 10359-10364.

Wang, J., Yu, W., Yang, Y., Li, X., Chen, T., Liu, T., ... & Zhang, B. (2015). Genome-wide analysis of tomato long non-coding RNAs and identification as endogenous target mimic for microRNA in response to TYLCV infection. Scientific Reports, 5, 16946.

Wang, Y., Bouchabke-Coussa, O., Lebris, P., Antelme, S., Soulhat, C., Gineau, E., ... & Legée, F. (2015). LACCASE5 is required for lignification of the *Brachypodium distachyon Culm*. Plant Physiology, 168(1), 192-204.

Wang, J., Yu, W., Yang, Y., Li, X., Chen, T., Liu, T., ... & Zhang, B. (2015). Genome-wide analysis of tomato long non-coding RNAs and identification as endogenous target mimic for microRNA in response to TYLCV infection. Scientific Reports, 5, 16946.

Wang, F., Wang, C., Liu, P., Lei, C., Hao, W., Gao, Y., ... & Zhao, K. (2016). Enhanced rice blast resistance by CRISPR/Cas9-targeted mutagenesis of the ERF transcription factor gene *OsERF922*. PloS one, 11(4), e0154027.

Wang, C. Y., Liu, S. R., Zhang, X. Y., Ma, Y. J., Hu, C. G., & Zhang, J. Z. (2017). Genome-wide screening and characterization of long non-coding RNAs involved in flowering development of trifoliate orange (*Poncirus trifoliata L. Raf.*). Scientific Reports, 7, 43226.

Wang, N., Cao, P., Xia, W., Fang, L., & Yu, H. (2017). Identification and characterization of long non-coding RNAs in response to early infection by *Melampsora larici-populina* using genome-wide high-throughput RNA sequencing. Tree Genetics & Genomes, 13(2), 34.

Wang, Z., Liu, Y., Li, L., Li, D., Zhang, Q., Guo, Y., ... & Huang, H. (2017). Whole transcriptome sequencing of *Pseudomonas syringae pv. actinidiae*-infected kiwifruit plants reveals species-specific interaction between long non-coding RNA and coding genes. Scientific Reports, 7(1), 1-15.

Wang, Z., Liu, Y., Li, D., Li, L., Zhang, Q., Wang, S., & Huang, H. (2017). Identification of circular RNAs in kiwifruit and their species-specific response to bacterial canker pathogen invasion. Frontiers in Plant Science, 8, 413.

Wang, B., Sumit, R., Sahu, B. B., Ngaki, M. N., Srivastava, S. K., Yang, Y., ... & Bhattacharyya, M. K. (2018). *Arabidopsis* novel glycine-rich plasma membrane PSS1 protein enhances disease resistance in transgenic soybean plants. Plant Physiology, 176(1), 865-878.

Wang, J., Yang, Y., Jin, L., Ling, X., Liu, T., Chen, T., ... & Zhang, B. (2018). Re-analysis of long non-coding RNAs and prediction of circRNAs reveal their novel roles in susceptible tomato following TYLCV infection. BMC Plant Biology, 18(1), 104.

Wang, X., Pei, X., Guo, G., Qian, X., Dou, D., Zhang, Z., ... & Duan, X. (2020). Exosome-mediated transfer of long noncoding RNA *H19* induces doxorubicin resistance in breast cancer. Journal of Cellular Physiology, 235(10), 6896-6904.

Wasternack, C., & Song, S. (2017). Jasmonates: biosynthesis, metabolism, and signaling by proteins activating and repressing transcription. Journal of Experimental Botany, 68(6), 1303-1321.

Wei, S., Gruber, M. Y., Yu, B., Gao, M. J., Khachatourians, G. G., Hegedus, D. D., ... & Hannoufa, A. (2012). *Arabidopsis* mutant sk156 reveals complex regulation of *SPL15* in a miR156-controlled gene network. BMC Plant Biology, 12(1), 169.

Westholm, J. O., Miura, P., Olson, S., Shenker, S., Joseph, B., Sanfilippo, P., ... & Lai, E. C. (2014). Genome-wide analysis of drosophila circular RNAs reveals their structural and sequence properties and age-dependent neural accumulation. Cell Reports, 9(5), 1966-1980.

Widjaja, I., Lassowskat, I., Bethke, G., Eschen-Lippold, L., Long, H. H., Naumann, K., ... & Lee, J. (2010). A protein phosphatase 2C, responsive to the bacterial effector *AvrRpm1* but not to the AvrB effector, regulates defense responses in *Arabidopsis*. The Plant Journal, 61(2), 249-258.

Wildermuth, M. C., Dewdney, J., Wu, G., & Ausubel, F. M. (2001). Isochorismate synthase is

required to synthesize salicylic acid for plant defence. Nature, *414*(6863), 562-565.

Williams, P. H. (1966). A system for the determination of races of *Plasmodiophora brassicae* that infect cabbage and rutabaga. Phytopathology, 56(6), 624-626.

Williams, B. P., & Gehring, M. (2017). Stable transgenerational epigenetic inheritance requires a DNA methylation-sensing circuit. Nature Communications, 8(1), 1-8.

Wilusz, J. E., Sunwoo, H., & Spector, D. L. (2009). Long noncoding RNAs: functional surprises from the RNA world. Genes & Development, 23(13), 1494-1504.

Wingler, A., Fritzius, T., Wiemken, A., Boller, T., & Aeschbacher, R. A. (2000). Trehalose induces the ADP-glucose pyrophosphorylase gene, ApL3, and starch synthesis in *Arabidopsis*. Plant Physiology, 124(1), 105-114.

Wu, H. J., Wang, Z. M., Wang, M., & Wang, X. J. (2013). Widespread long noncoding RNAs as endogenous target mimics for microRNAs in plants. Plant physiology, 161(4), 1875-1884.

Wu, H., Yang, L., & Chen, L. L. (2017). The diversity of long noncoding RNAs and their generation. Trends in Genetics, 33(8), 540-552.

Xia, Y., Suzuki, H., Borevitz, J., Blount, J., Guo, Z., Patel, K., ... & Lamb, C. (2004). An extracellular aspartic protease functions in *Arabidopsis* disease resistance signaling. The EMBO journal, *23*(4), 980-988.

Xiao, W., Sheen, J., & Jang, J. C. (2000). The role of hexokinase in plant sugar signal transduction
and growth and development. Plant Molecular Biology, 44(4), 451-461.

Xie, M., Zhang, J., Tschaplinski, T. J., Tuskan, G. A., Chen, J. G., & Muchero, W. (2018). Regulation of lignin biosynthesis and its role in growth-defense tradeoffs. Frontiers in Plant Science, 9, 1427.

Xin, M., Wang, Y., Yao, Y., Song, N., Hu, Z., Qin, D., ... & 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(1), 61.

Xiang, L., Cai, C., Cheng, J., Wang, L., Wu, C., Shi, Y., ... & Yuan, Y. (2018). Identification of circularRNAs and their targets in *Gossypium* under *Verticillium* wilt stress based on RNA-seq. PeerJ, 6, e4500.

Xiang, C., Zhao, P. X., Miao, Z. Q., Zhang, J., & Liu, Q. Q. (2019). MADS-box factor *AGL16* negatively regulates drought resistance via stomatal density and stomatal movement. bioRxiv, 723106.

Xuan, H., Zhang, L., Liu, X., Han, G., Li, J., Li, X., ... & Zhang, S. (2015). PLNlncRbase: a resource for experimentally identified lncRNAs in plants. Gene, 573(2), 328-332.

Yang, C. L., Liang, S., Wang, H. Y., Han, L. B., Wang, F. X., Cheng, H. Q., ... & Xia, G. X. (2015). Cotton major latex protein 28 functions as a positive regulator of the ethylene responsive factor 6 in defense against *Verticillium dahliae*. Molecular Plant, 8(3), 399-411.

Yang, F., Zhao, D., Fan, H., Zhu, X., Wang, Y., Liu, X., ... & Chen, L. (2020). Functional analysis of long non-coding RNAs reveal their novel roles in biocontrol of bacteria-induced tomato resistance to *Meloidogyne incognita*. International Journal of Molecular Sciences, 21(3), 911.

Yang, J., Hsiang, T., Bhadauria, V., Chen, X. L., & Li, G. (2017). Plant fungal pathogenesis. Biomed Research International, 2017, 1-2.

Ye, C. Y., Chen, L., Liu, C., Zhu, Q. H., & Fan, L. (2015). Widespread noncoding circular RNAs in plants. New Phytologist, 208(1), 88-95.

Yeung, E. C., & Saxena, P. K. (2005). Histological techniques. In Protocol for Somatic Embryogenesis in Woody Plants (pp. 517-537). Springer, Dordrecht.

Yi, H., & Richards, E. J. (2007). A cluster of disease resistance genes in *Arabidopsis* is coordinately regulated by transcriptional activation and RNA silencing. The Plant Cell, *19*(9), 2929-2939.

Yi, X., Zhang, Z., Ling, Y., Xu, W., & Su, Z. (2015). PNRD: a plant non-coding RNA database. Nucleic Acids Research, 43(D1), D982-D989.

Yi, Y., Zhao, Y., Li, C., Zhang, L., Huang, H., Li, Y., ... & Hu, Y. (2017). RAID v2. 0: an updated resource of RNA-associated interactions across organisms. Nucleic Acids Research, 45(D1), D115-D118.

Young, M. D., Wakefield, M. J., Smyth, G. K., & Oshlack, A. (2012). Goseq: Gene Ontology testing for RNA-seq datasets. R Bioconductor, 8, 1-25.

Yoon, J. H., Abdelmohsen, K., & Gorospe, M. (2013). Posttranscriptional gene regulation by long noncoding RNA. Journal of Molecular Biology, 425(19), 3723-3730.

Yin, Y., Yan, P., Lu, J., Song, G., Zhu, Y., Li, Z., ... & Orkin, S. H. (2015). Opposing roles for the lncRNA haunt and its genomic locus in regulating *HOXA* gene activation during embryonic stem cell differentiation. Cell Stem Cell, 16(5), 504-516.

Yu, F., Zhang, X., Huang, Z., Chu, M., Song, T., Falk, K. C., ... & Gossen, B. D. (2016). Identification of genome-wide variants and discovery of variants associated with *Brassica rapa* clubroot resistance gene *Rcr1* through bulked segregant RNA sequencing. PloS one, 11(4), e0153218.

Yu, Y., Zhou, Y. F., Feng, Y. Z., He, H., Lian, J. P., Yang, Y. W., ... & Chen, Y. Q. (2020). Transcriptional landscape of pathogen-responsive lnc RNA s in rice unveils the role of ALEX 1 in jasmonate pathway and disease resistance. Plant Biotechnology Journal, *18*(3), 679-690.

Zare, K., Shademan, M., Seno, M. M. G., & Dehghani, H. (2018). CRISPR/Cas9 knockout strategies to ablate *CCAT1* lncRNA gene in cancer cells. Biological Procedures Online, 20(1), 21.

Zaynab, M., Fatima, M., Abbas, S., Umair, M., Sharif, Y., & Raza, M. A. (2018). Long non-coding RNAs as molecular players in plant defense against pathogens. Microbial Pathogenesis, 121, 277-282.

Zhang, Y., Goritschnig, S., Dong, X., & Li, X. (2003). A gain-of-function mutation in a plant disease resistance gene leads to constitutive activation of downstream signal transduction pathways in suppressor of *NPR1-1*, constitutive 1. The Plant Cell, 15(11), 2636-2646.

Zhang, S., Yang, X., Sun, M., Sun, F., Deng, S., & Dong, H. (2009). Riboflavin-induced priming for pathogen defense in *Arabidopsis thaliana*. Journal of Integrative Plant Biology, 51(2), 167-174.

Zhang, X., Zou, Z., Gong, P., Zhang, J., Ziaf, K., Li, H., ... & Ye, Z. (2011). Over-expression of microRNA169 confers enhanced drought tolerance to tomato. Biotechnology Letters, 33(2), 403-409.

Zhang, H., Chen, X., Wang, C., Xu, Z., Wang, Y., Liu, X., ... & Ji, W. (2013). Long non-coding genes implicated in response to stripe rust pathogen stress in wheat (*Triticum aestivum L.*). Molecular Biology Reports, 40(11), 6245-6253.

Zhang, H., Hu, W., Hao, J., Lv, S., Wang, C., Tong, W., ... & Ji, W. (2016). Genome-wide identification and functional prediction of novel and fungi-responsive lincRNAs in *Triticum aestivum*. BMC Genomics, 17(1), 238.

Zhang, R., Sun, B., Chen, J., Cao, A., Xing, L., Feng, Y., ... & Chen, P. (2016). *Pm55*, a developmental-stage and tissue-specific powdery mildew resistance gene introgressed from *Dasypyrum villosum* into common wheat. Theoretical and Applied Genetics, 129(10), 1975-1984.

Zhang, X., Liu, Y., Fang, Z., Li, Z., Yang, L., Zhuang, M., ... & 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, 1929.

Zhang, J., Chen, L., Fu, C., Wang, L., Liu, H., Cheng, Y., ... & Liang, Y. (2017). Comparative transcriptome analyses of gene expression changes triggered by *Rhizoctonia solani* AG1 IA infection in resistant and susceptible rice varieties. Frontiers in Plant Science, 8, 1422.

Zhang, L., Wang, M., Li, N., Wang, H., Qiu, P., Pei, L., ... & Liu, S. (2018). Long noncoding RNA s involve in resistance to *Verticillium dahliae*, a fungal disease in cotton. Plant Biotechnology Journal, 16(6), 1172-1185.

Zhang, M., Zhao, K., Xu, X., Yang, Y., Yan, S., Wei, P., ... & Yang, X. (2018). A peptide encoded by circular form of *LINC-PINT* suppresses oncogenic transcriptional elongation in glioblastoma. Nature Communications, 9(1), 4475.

Zhang, J., Wei, L., Jiang, J., Mason, A. S., Li, H., Cui, C., ... & Jiang, L. (2018). Genome-wide identification, putative functionality and interactions between lncRNAs and miRNAs in Brassica species. Scientific Reports, 8(1), 1-11.

Zhang, J., Yang, Z., Feng, P., Zhong, X., Ma, Q., Su, Q., ... & Yang, Y. (2019). Identification and the potential roles of long non-coding RNAs in cotton leaves damaged by *Aphis gossypii*. Plant Growth Regulation, 88(3), 215-225.

Zhang, Z., Zheng, Y., Ham, B. K., Zhang, S., Fei, Z., & Lucas, W. J. (2019). Plant lncRNAs are enriched in and move systemically through the phloem in response to phosphate deficiency. Journal of Integrative Plant Biology, 61(4), 492-508.

Zhao, J. P., Jiang, X. L., Zhang, B. Y., & Su, X. H. (2012). Involvement of microRNA-mediated gene expression regulation in the pathological development of stem canker disease in *Populus trichocarpa*. PloS one, 7(9), e44968.

Zhao, X., Tan, G., Xing, Y., Wei, L., Chao, Q., Zuo, W., ... & Xu, M. (2012). Marker-assisted introgression of *qHSR1* to improve maize resistance to head smut. Molecular Breeding, 30(2), 1077-1088.

Zhao, Y., Li, H., Fang, S., Kang, Y., Wu, W., Hao, Y., ... & Chen, R. (2016). NONCODE 2016: an informative and valuable data source of long non-coding RNAs. Nucleic acids research, 44(D1), D203-D208.

Zhao, W., Cheng, Y., Zhang, C., You, Q., Shen, X., Guo, W., & Jiao, Y. (2017). Genome-wide identification and characterization of circular RNAs by high throughput sequencing in soybean. Scientific Reports, 7(1), 1-11.

Zhao, Y., Bi, K., Gao, Z., Chen, T., Liu, H., Xie, J., ... & Jiang, D. (2017). Transcriptome analysis of *Arabidopsis thaliana* in response to *Plasmodiophora brassicae* during early infection. Frontiers in Microbiology, 8, 673.

Zhao, T., Tao, X., Feng, S., Wang, L., Hong, H., Ma, W., ... & Guan, X. (2018). LncRNAs in polyploid cotton interspecific hybrids are derived from transposon neofunctionalization. Genome Biology, 19(1), 1-17.

Zhao, W., Chu, S., & Jiao, Y. (2019). Present scenario of circular RNAs (circRNAs) in plants. Frontiers in Plant Science, 10.

Zhen, Q., Fang, T., Peng, Q., Liao, L., Zhao, L., Owiti, A., & Han, Y. (2018). Developing gene-tagged molecular markers for evaluation of genetic association of apple *SWEET* genes with fruit sugar accumulation. Horticulture Research, 5(1), 1-12.

Zheng, B., Wang, Z., Li, S., Yu, B., Liu, J. Y., & Chen, X. (2009). Intergenic transcription by RNA polymerase II coordinates Pol IV and Pol V in siRNA-directed transcriptional gene silencing in *Arabidopsis*. Genes & Development, 23(24), 2850-2860.

Zhou, J., Peng, Z., Long, J., Sosso, D., Liu, B., Eom, J. S., ... & White, F. F. (2015). Gene targeting by the TAL effector PthXo2 reveals cryptic resistance gene for bacterial blight of rice. The Plant Journal, 82(4), 632-643.

Zhou, B., Zhao, H., Yu, J., Guo, C., Dou, X., Song, F., ... & Zhou, Y. (2018). EVLncRNAs: a manually curated database for long non-coding RNAs validated by low-throughput experiments. Nucleic Acids Research, 46(D1), D100-D105.

Zhou, R., Zhu, Y., Zhao, J., Fang, Z., Wang, S., Yin, J., ... & Ma, D. (2018). Transcriptome-wide identification and characterization of potato circular RNAs in response to *Pectobacterium carotovorum* subspecies *brasiliense* infection. International Journal of Molecular Sciences, 19(1), 71.

Zhou, Y., Cho, W. K., Byun, H. S., Chavan, V., Kil, E. J., Lee, S., & Hong, S. W. (2019). Genome-wide identification of long non-coding RNAs in tomato plants irradiated by neutrons followed by infection with Tomato yellow leaf curl virus. PeerJ, 7, e6286.

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), 1-9.

Zhu, X., Li, Z., Xu, H., Zhou, M., Du, L., & Zhang, Z. (2012). Overexpression of wheat lipid transfer protein gene *TaLTP5* increases resistances to *Cochliobolus sativus* and *Fusarium graminearum* in transgenic wheat. Functional & Integrative Genomics, 12(3), 481-488.

Zhu, Q. H., & Wang, M. B. (2012). Molecular functions of long non-coding RNAs in plants. Genes, 3(1), 176-190.

Zhu, Q. H., Stephen, S., Taylor, J., Helliwell, C. A., & Wang, M. B. (2014). Long noncoding RNA s responsive to *Fusarium oxysporum* infection in *Arabidopsis thaliana*. New Phytologist, 201(2), 574-584.

Zhu, Y. X., Jia, J. H., Yang, L., Xia, Y. C., Zhang, H. L., Jia, J. B., ... & Liu, L. C. (2019). Identification of cucumber circular RNAs responsive to salt stress. BMC plant biology, 19(1), 164.

Zou, C., Wang, P., & Xu, Y. (2016). Bulked sample analysis in genetics, genomics and crop improvement. Plant Biotechnology Journal, 14(10), 1941-1955.

Zrenner, R., Stitt, M., Sonnewald, U., & Boldt, R. (2006). Pyrimidine and purine biosynthesis and degradation in plants. Annual Reviews of Plant Biology, *57*, 805-836.

Zubko, E., & Meyer, P. (2007). A natural antisense transcript of the Petunia hybrida Sho gene suggests a role for an antisense mechanism in cytokinin regulation. The Plant Journal, 52(6), 1131-1139.