

School of Molecular and Life Sciences

**The Role of Small RNAs During the Interaction Between
Sclerotinia Sclerotiorum and Brassica Napus**

Roshan Regmi
0000-0002-9717-0609


**This thesis is presented for the Degree of
Doctor of Philosophy
of
Curtin University**

January 2022

I. DECLARATION

I declare that this thesis contains no material previously published by any other person except where due acknowledgement has been made.

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.


21/01/2022

Roshan Regmi

II. AUTHORSHIP DECLARATION

This thesis contains one paper published in a peer-reviewed scientific journal (BMC Plant Biology) and additional manuscripts prepared for publication in due course. I have the permission from the journal to reproduce the article as a thesis chapter.

Details of work:

Regmi R, Newman TE, Kamphuis LG, Derbyshire MC (2021) Identification of *Brassica napus* small RNAs responsive to infection by a necrotrophic pathogen. *BMC Plant Biology*, **21**, Article number 399.

Location in thesis:

This research paper is included as **Chapter 5**.

Student contributions:

- Conceived the design of the study with MD, MCD, and LGK
- Designing the experiment with the help of MD
- Conducted all experiment and collecting data
- Data analysis with assistance from MD
- Lead writer of the manuscript, with comments and suggestions from co-authors.

Co-author signatures:

Signed by LGK, the coordinating lead supervisor on behalf of all co-authors.



Date: 21/01/2022

III. ACKNOWLEDGEMENT

A dream does not come true through magic in a single day, it needs hard work, determination and strong support from others. I wish to show my appreciation to a group of people and helping hands that I came across with throughout my PhD journey.

At first, I would like to express my deep gratitude and respect to my supervisors' team at Curtin University Dr. Lars Kamphuis, Dr. Mark Derbyshire and Dr. Toby Newman for their continuous support, guidance on all aspects of the generation of this thesis and scientific papers. Working under their guidance motivated me in creating a thesis, which not only fulfilled my degree requirement, but also enhanced my knowledge and skills in the research field.

I would also like to thank Dr. Yuphin Khentry, research technician at Curtin University for providing me her advice and expertise with the experimental work related to growth and care of plants and conduction of infection assays.

I would also like to thank the Australian Government Research Training Program and Commonwealth Scientific Research Organization for the financial support throughout my PhD. I thank Curtin University and Grains Research and Development Corporation for resources and facilities.

My PhD journey would be more difficult without the immense support of my special friends and family members. Especially my wife Ranjita who always provides me the freedom and chance to continue my academic journey despite many obstacles and difficult times we have been through. I would also like to thank her to raise our beautiful little girl Roszyn who is as old as my PhD. I will always be grateful for every opportunity I had at Centre for Crop and Disease Management.

IV. TABLE OF CONTENTS

I. DECLARATION.....	1
II. AUTHORSHIP DECLARATION	2
III. ACKNOWLEDGEMENT	3
IV. TABLE OF CONTENTS	4
V. ABSTRACT.....	10
VI. LIST OF FIGURES AND TABLES	14
1.1 Introduction.....	20
1.2 Molecular mechanisms of host and pathogen interactions	21
1.3 Small RNAs (sRNAs) and their roles in plant disease	25
1.4 Small RNA-mediated cross-kingdom silencing.....	27
1.4.1 <i>Fusarium oxysporum</i> and banana	30
1.4.2 <i>Blumeria</i> spp. and barley	30
1.4.3 <i>Fusarium graminearum</i> and wheat.....	31
1.4.4 <i>Botrytis cinerea</i> and Arabidopsis.....	32
1.4.5 <i>Zymoseptoria tritici</i> and wheat	33
1.4.6 <i>Sclerotinia sclerotiorum</i> and its host plants.....	33
1.4.7 <i>Cuscuta</i> spp. and Arabidopsis.....	34
1.5 Role of plant sRNAs in plant immunity response	35
1.6 The function of small RNAs in <i>Brassica napus</i>	37
1.6.1 Oil content.....	37
1.6.2 Cadmium stress.....	38
1.6.3 Salt and drought stress	39
1.6.4 <i>Sclerotinia</i> stem rot	40
1.7 Importance of bioinformatics in small RNA analysis.....	42

1.7.1 Pre-processing and quality control of sRNA sequencing dataset	43
1.7.2 sRNA alignment and annotation	44
1.7.3 Differential expression analysis	45
1.7.4 Target prediction	46
1.7.5 sRNA/mRNA target validation	47
1.8 Knowledge gap in the literature	49
1.9 Conclusion and thesis objectives	50
1. 10 References	51
 Chapter 2: Transformation of <i>Sclerotinia sclerotiorum</i> with a gene encoding green fluorescent protein	 63
Abstract	63
2.1 Introduction	64
2.2 Methods	67
2.2.1 Transformation vector	67
2.2.2 Plasmid amplification for <i>S. sclerotiorum</i> transformation	68
2.2.3 Maintenance and culture of <i>Sclerotinia sclerotiorum</i>	70
2.2.4 PEG-mediated protoplast transformation of <i>Sclerotinia sclerotiorum</i> with GFP ...	70
2.2.5 Transformant selection	72
2.2.6 DNA extraction	73
2.2.7 Confirmation of insertion of <i>GFP</i> in transformants by PCR	74
2.3 Results	76
2.3.1 Putative transgenic strains were confirmed for the insertion of <i>GFP</i> by PCR	76
2.3.2 Fluorescence microscopy identifies three stable GFP expressing strains of <i>Sclerotinia sclerotiorum</i> isolate CU8.24	77
2.3.3 Transgenic and wild type strains maintain the same growth rate <i>in vitro</i>	79
2.3.4 Infection of plants with GFP-expressing strains of <i>Sclerotinia sclerotiorum</i> shows no difference to wild type pathogenicity	81

2.4 Discussion.....	83
Chapter 3: Genome wide identification of <i>Sclerotinia sclerotiorum</i> small RNAs and their endogenous targets.....	90
Abstract.....	90
3.1 Introduction.....	91
3.2 Methods.....	93
3.2.1 Fungal and plant sources.....	93
3.2.2 RNA isolation and sequencing.....	94
3.2.3 Analysis of sequencing data.....	95
3.2.3 Differential expression analysis.....	96
3.2.4 Small RNA target prediction.....	96
3.2.5 Detection of PHAS loci.....	97
3.3 Results.....	98
3.3.1 Sequencing data analysis.....	98
3.3.2 Assessment of sRNA population and characteristics.....	100
3.3.3 Annotation of sRNA generating loci in <i>Sclerotinia sclerotiorum</i>	104
3.3.4 Identification of <i>Sclerotinia sclerotiorum</i> miRNAs.....	104
3.3.5 Small RNAs are differentially expressed in <i>Sclerotinia sclerotiorum</i> mycelium after host inoculation.....	106
3.3.4 Prediction of phased loci in <i>Sclerotinia sclerotiorum</i>	112
3.3.6 Endogenous targets of fungal sRNAs were mostly repetitive elements.....	114
3.4. Discussion.....	115
Chapter 4: Small RNA profiling in <i>Sclerotinia sclerotiorum</i> and target identification in <i>Brassica napus</i>	130
Abstract.....	130
4.2 Materials and Methods.....	133
4.2.1 Plant material and inoculation.....	133

4.2.2 RNA extraction and construction of sRNA and degradome sequencing libraries	134
4.2.3 Filtering of reads and assignment for mapping.....	135
4.2.4 Prediction of fungal sRNAs	135
4.2.5 Normalization and differential expression analysis	137
4.2.6 Prediction of fungal sRNA target genes in plant hosts using the psRNA target server.....	137
4.2.7 Quantification of target genes using quantitative polymerase chain reaction	138
4.2.8 5' RACE validation of target genes	138
4.2.9 Degradome sequencing and analysis	139
4.2.10 Gene ontology enrichment analysis of overrepresented target gene domains	139
4.3 Results.....	139
4.3.1 Overview of small RNA sequencing from the fungal pathogen <i>Sclerotinia sclerotiorum</i>	139
4.3.2 Prediction of fungal small RNAs.....	140
4.3.3 Identification of miRNA-like loci in <i>Sclerotinia sclerotiorum</i>	142
4.3.4 Small RNAs are differentially expressed in <i>Sclerotinia sclerotiorum</i> mycelium during host infection	144
4.3.5 Small RNA target genes in the host are enriched with plant immune response domains	145
4.3.6 qPCR shows that fungal sRNA target genes are repressed during infection	146
4.3.7 5' RACE reveals cleavage of putative target genes.....	147
4.3.8 High throughput fungal sRNA target identification in <i>Brassica napus</i> using degradome sequencing	149
4.4 Discussion	154
4.5 References.....	158
Chapter 5: Identification of <i>Brassica napus</i> small RNAs responsive to infection by a necrotrophic pathogen.....	166
Abstract.....	166

5.1 Introduction.....	167
5.2 Methods.....	170
5.2.1 Biological materials	171
5.2.2 Total RNA extraction and sequencing.....	171
5.2.3 Analysis of small RNA sequencing data	172
5.2.4 Analysis of sRNA targeting using <i>in silico</i> target prediction and degradome sequencing data.....	173
5.2.5 Gene ontology enrichment analysis	174
5.2.6 Five prime rapid amplification of cDNA ends of a cleaved target.....	174
5.2.7 Quantitative polymerase chain reaction.....	175
5.3. Results.....	176
5.3.1 Overview of sequencing results	176
5.3.2 Characteristic features of the <i>Brassica napus</i> small RNA population.....	178
5.3.3 A total of 730 unique <i>Brassica napus</i> small RNAs are upregulated in response to <i>Sclerotinia sclerotiorum</i> infection	180
5.3.5 Identification of conserved miRNAs in <i>Brassica napus</i>	186
5.3.8 RNA structure-aided prediction algorithms identify 135 novel <i>Brassica napus</i> micro RNA loci.....	190
5.3.9 Nine <i>Brassica napus</i> PHAS loci are differentially expressed in response to <i>Sclerotinia sclerotiorum</i> infection	191
5.4 Discussion	198
5.6 References.....	205
Chapter 6: Generation of disruption mutants of RNAi machinery genes in <i>Sclerotinia sclerotiorum</i>	215
Abstract.....	215
6.1 Introduction.....	215
6.2.1 Identification of RNAi machinery genes in the <i>S. sclerotiorum</i> CU8.24 genome.....	219
6.2.1 Amplification of RNAi gene flanking sequences	219

6.2.2 Assembly of knockout plasmid constructs	220
6.2.3 Confirmation of correctly assembled constructs.....	221
6.2.3 PEG-mediated protoplast transformation	222
6.2.4 Screening of transformants	222
6.2.5 Phenotypic analysis of transformants	222
6.3 Results.....	223
6.3.1 Presence of RNAi machinery genes in the <i>Sclerotinia sclerotiorum</i> genome.....	223
6.3.2 Confirmation of knockout plasmid constructs	225
6.3.3. Disruption mutants of <i>Dicer</i> and <i>RdRp</i> genes.....	227
6.3.4 RNAi gene mutants of <i>S. sclerotiorum</i> do not exhibit altered mycelium growth but do have reduced sclerotia formation	228
6.3.5 <i>Dicer1</i> disruption mutants do not exhibit reduced pathogenicity.....	230
6.4 Discussion.....	231
6.4 References.....	234
Chapter 7: General discussion and conclusions.....	240
7.1 General discussion and conclusions.....	240
7.2 Final comments and future direction	246
7.3 References.....	248
APPENDIX.....	254

V. ABSTRACT

Canola (*Brassica napus*) is an important oilseed crop that is used for human consumption, as fodder for livestock, and as a biodiesel. Every year, millions of dollars of economic losses occur in canola due to various diseases. Among them is Sclerotinia stem rot (SSR) caused by *Sclerotinia sclerotiorum*, a cosmopolitan fungal pathogen, which causes significant economic losses to growers. So far, there are commercial canola cultivars that exhibit only limited resistance to SSR, although other management strategies have proven to be successful to some extent, like fungicide usage and cultural control methods. Given that fungicide application is expensive, it may be important to combine molecular breeding techniques with other management strategies to manage SSR to the best possible outcome. To this end, it is important to understand the molecular interactions between plant and pathogen.

Small RNAs (sRNAs) are short non-coding RNAs that regulate various biological processes in eukaryotes. Recently, sRNAs were reported to have a role during host and pathogen interactions. However, there are limited studies on the involvement of sRNAs in the interaction between *S. sclerotiorum* and *B. napus*. Therefore, the main objective of this thesis was to investigate how sRNAs are involved in the molecular interaction between *S. sclerotiorum* and *B. napus*. In addition to this, an Australian *S. sclerotiorum* strain expressing the green fluorescent protein (ssGFP) was developed and checked to see that it maintained virulence. This strain could be useful for research in Australia where an import permit and strict biosecurity measures are required for experimentation with foreign isolates.

To understand the roles of fungal sRNAs in the interaction, sRNA profiling of *S. sclerotiorum* was done with Illumina-based sRNA sequencing. Target prediction using these

sRNAs in *B. napus* was done with an *in-silico* method and targets were further investigated with degradome sequencing which provides experimental evidence of sRNA-mediated mRNA cleavage. Based on the analysis of the sRNA and degradome sequencing, three different sRNA interactions were further investigated using wet lab techniques. Using the same sRNA dataset, endogenous targets were also predicted in *S. sclerotiorum* and analysed bioinformatically. To try and understand the broad role of RNAi components in *S. sclerotiorum* pathogenesis, knockout strains missing key components of the RNAi machinery were also generated. To understand the roles of *B. napus* sRNAs in immunity, sRNA and degradome sequencing were used in conjunction with wet lab techniques. Together, this thesis provides insight into the molecular mechanism of *S. sclerotiorum* and *B. napus* interactions at the sRNA level.

Expression of the green fluorescent protein (GFP) is an important molecular tool that facilitates the investigation of interactions between pathogens and their hosts. The first research chapter describes how an Australian *S. sclerotiorum* strain was transformed with GFP without compromising virulence. Fungi were transformed with GFP following the protoplast-mediated transformation method. The confirmation of insertion of the *GFP* gene was accomplished with PCR and confocal microscopy confirmed expression. In addition, *in vitro* and *in planta* growth assays were conducted to check whether the transformed strains maintained wild-type virulence. The results demonstrated the successful insertion of the *GFP* gene and no significant difference in morphology and virulence of transformed strains, suggesting these strains can be used to study the interactions between *S. sclerotiorum* and its host species. This is important for Australian research as experiments involving foreign strains, such as the reference strain ‘1980’, require an import permit and strict biosecurity measures.

Chapter three analyses endogenous fungal sRNA-mRNA interactions. This chapter provides information on how fungal sRNAs change their expression at 12 and 24 HPI in comparison to *in vitro* and what types of genes they regulate. The results show that the fungal sRNAs mostly silence transposable elements.

The fourth chapter describes sRNA profiling of *S. sclerotiorum* genome and *B. napus* target identification and validation. This chapter demonstrated cross-kingdom silencing where widespread fungal sRNA-mediated cleavage of host mRNAs occurs. This chapter also describes development of rapid amplification cDNA ends (RACE) experiments to further study *B. napus* mRNA cleavage products. This chapter identified that *B. napus* defence-associated genes AP2/ERF, a protein kinase, and a zinc finger domain-containing gene are likely silenced via fungal sRNA-directed cleavage.

In chapter 5, detailed analyses of sRNA-mediated plant immune responses during the *S. sclerotiorum* and *B. napus* interaction were conducted. This chapter demonstrates that under pathogen attack, the plant recruits different sRNA biogenesis pathways and differentially expresses sRNAs including miRNAs that might play key roles in regulating endogenous plant immune response genes. An important ethylene response factor gene was shown to be regulated by a novel sRNA in the plant during pathogen attack and this was validated through 5'RACE and degradome sequencing. This chapter also extensively identifies different classes of sRNAs such as miRNAs, tasiRNAs, and siRNAs in the *B. napus* genome. The findings of this chapter provide insight into the endogenous regulation of plant sRNAs during pathogen attack.

Chapter 6 describes the presence of the RNAi machinery in *S. sclerotiorum* and the generation of gene knockout constructs, the homologous recombination method of gene knockout via protoplast transformation, a method of validating the transformation and also explains the challenges of obtaining pure transformed homokaryotic strains.

VI. LIST OF FIGURES AND TABLES

Fig. 1. 1. A schematic diagram showing cross-kingdom silencing mediated by small RNAs.	28
Fig. 1. 2. Bioinformatics analysis of small RNA sequencing analysis	42
Fig. 1. 3. A model showing possible small RNA interaction during pathogen and host interaction of <i>Sclerotinia sclerotiorum</i> and <i>Brassica napus</i>	50
Fig. 2. 1. “puc57.Ss1186-GFP” plasmid flanked by <i>Sclerotinia sclerotiorum</i> sequences (red colour) for target insertion of the GFP gene.	67
Fig. 2.2. Colony PCR showing amplification of green fluorescent protein (GFP) gene and two <i>Sclerotinia sclerotiorum</i> flanking sequences SSA (left) and SSB (right) in the pUC57 plasmid. SSA spans outside the promoter region while SSB presents outside the terminator region. Lane 1 is 1Kb Bioline hyperladder; lane 2 amplification of left flanking sequence; lane 3 amplification of GFP and lane 4 amplification of right flanking sequence.....	69
Fig. 2.3. The confirmation of GFP insertion in <i>Sclerotinia sclerotiorum</i> genome with the use of different set of primers. (A) The schematic of primer position for the amplification of left and right flanks. (B). The electrophoresis gel images of the amplicon produce for wild type (WT) and three putative transgenic strains (T1, T2 and T3) with GFP, Hyg, F1 and F2 primers. NTC is a no template control.....	77
Fig. 2. 4. Confocal microscopy images after six hyphal tip regenerations of wild type and green fluorescent protein (GFP) transformed strains of <i>S. sclerotiorum</i> isolate CU 8.24. WT, T1, T2, and T3 are wild type, transgenic strain 1, transgenic strain 2, and transgenic strain 3, respectively. For each individual the left image (GFP) shows the GFP expression following excitation wavelength of 488.6 nm and an emission wavelength of 525.0 nm and the right image the bright field (BF).	79
Fig. 2. 5. Comparison of mycelium growth measured at four different time points between wild type and <i>Sclerotinia sclerotiorum</i> -GFP strains. No significant differences were observed between the wild type (WT) and the three GFP-transformed stains (T1, T2 and T3) by Tukey HSD-test ($p < 0.05$).	80
Fig. 2. 6. Lesion development on stems of three GFP-transformed <i>S. sclerotiorum</i> strains (T1, T2 and T3) and one wild type (WT) strain after two weeks placed in a growth room at 16 H day and 8 H night at 24°C.	82
Fig. 2. 7. Comparison of lesion growth on <i>Brassica napus</i> plant measured at three different time points between wild type and <i>Sclerotinia sclerotiorum</i> -GFP strains. No significant differences were observed between the wild type (WT) and the three GFP-transformed stains (T1, T2 and T3) by Tukey HSD-test ($p < 0.05$).	82

Fig. 3.1. Analysis of small RNA sequencing dataset of <i>Sclerotinia sclerotiorum</i> . The x-axis shows the sRNA libraries from where the sRNA reads were generated, y-axis shows the number of total reads identified for different reads; insert shows the representative figure of fungal mycelial mat experiment at 12 HPI and 24 HPI for sampling, R1, R2 and, R3 represent the data for three biological replicates.	99
Fig. 3.2. Percentage of reads mapping to the <i>Sclerotinia sclerotiorum</i> and <i>Brassica napus</i> genomes. The x-axis shows the replicates of treatment and the y-axis shows the percentage of reads mapping to either <i>S. sclerotiorum</i> (orange) or <i>B. napus</i> (turquoise).	100
Fig. 3.3. length distribution of <i>Sclerotinia sclerotiorum</i> sRNAs. The percentage of reads (y-axis) according to nucleotide (nt) sequence length (x-axis) obtained in 0 HPI (A), 12 HPI, (B) and 24 HPI (C) pooled samples. The red bars indicate percentage of library for each length of total reads while the blue bars indicate percentage of library for each length of unique reads.	102
Fig. 3. 4. 5' nucleotide bias of <i>Sclerotinia sclerotiorum</i> small RNAs. The percentage of adenine (orange), cytosine (green), guanine (torquiose) and uridine (purple) in the 5' nucleotide (nt) position according to read length for 0 HPI total reads (A) and unique reads (B), 12 HPI total reads (C) and unique reads (D), 24 HPI total reads (E) and unique reads (F).	103
Fig. 3.5. Assessment of sRNA origin of 1,073 Dicer-derived loci in the <i>Sclerotinia sclerotiorum</i> genome.	104
Fig. 3.6. The hairpin loop structure of the miRNA locus predicted by ShortStack. The different colours show the intensity of minimum free energy associated with the secondary structure.	106
Fig. 3.7. Heatmap of sRNAs identified in the <i>Sclerotinia sclerotiorum</i> genome. A Heat map of normalized expression data from the 25 ShortStack loci commonly up-regulated across both time points, 12 and 24 hours post inoculation (HPI). B The expression profile of 16 mature miRNAs common to all three libraries normalized to reads per million.	107
Fig. 3. 8. Venn diagram showing the number of milRNAs identified from <i>Sclerotinia sclerotiorum</i> from different libraries.	108
Fig .4. 1. Characteristics of <i>Sclerotinia sclerotiorum</i> small RNAs. Histogram of read size of 288 highly expressed sRNAs from pooled mock and infected samples.	141
Fig .4. 2. A pie chart showing the origin of 288 sRNA loci in the <i>Sclerotinia sclerotiorum</i> isolate CU8.24 genome.	142
Fig .4.3. Prediction of <i>Sclerotinia sclerotiorum</i> miRNA-like loci with at least 100 reads per million (RPM). (A) Bar graph showing milRNA families that had close homologues in	

- the miRBase database with three or fewer mismatches. (B) Novel miRNA loci with no homologues in miRbase..... 143
- Fig .4. 4. Heatmap showing 63 upregulated sRNA loci of *Sclerotinia sclerotiorum* isolate CU8.24 during infection of *Brassica napus* leaves (24 HPI) relative to *in vitro* controls. The three biological replicates for *in vitro* and infected samples are plotted with normalized counts from DESeq2. 144
- Fig .4.5. GO term enrichment of 156 highly expressed *Sclerotinia sclerotiorum* small RNA targets in *Brassica napus* obtained from the psRNA target server. The y-axis shows the biological process, while x-axis shows the $-\log_{10}(p \text{ adjusted})$ values. 146
- Fig. 4.6. Expression analysis of fungal sRNA target genes in *Brassica napus*. The x-axis displays gene ID and domains found in the genes of interest, *BnaA01g35070D* (AP2/ERF), *BnaC06g14210D* (protein kinase) and *BnaC08g20040D* (MORN motif). The y-axis represents the relative expression levels mean $\log(2^{-\Delta Ct})$ normalized to the *B. napus* actin gene. The error bar shows the standard error values from three biological replicates. 147
- Fig .4. 7. Mapping of the cleavage sites for four *Sclerotinia sclerotiorum* small RNAs targets using 5'RACE. *BnaC06g14210D* encodes a protein kinase, *BnaC01g35070D* encodes an AP2/ERF transcription factor, *BnaC08g20040D* encodes MORN motif and *BnaA02g11730D* encodes plastocyanin-like gene. The red coloured letters indicate the target cleavage site..... 148
- Fig .4.8. Validation of sRNA-mediated cleavage of *Brassica napus* zinc finger encoding transcript. (A) The 5'RACE clones in TOPO-TA vector confirmed by restriction digest with the upper band showing the backbone of the vector and the lower band shows the 5'RACE product size released after the digestion with *EcoRI*. (B) Mapping of the target cleavage site (indicated in with the red letter) in the *BnaC05g04050D* gene, which contains a zinc finger domain. 150
- Fig. 4.9. Experimental evidence of sRNA-mediated cleavage of *Brassica napus* immune related genes by degradome sequencing. T-plots showing transcript position on the x-axis and degradome coverage on the y-axis, CS shows the cleavage site in the gene... 151
- Fig. 5. 1. Changes in size class and 5' nucleotide of *Brassica napus* sRNAs in response to *Sclerotinia sclerotiorum* infection. (A) A principal component analysis based on normalized read counts from DESeq2. The x-axis shows principal component 1, which explained 99 % of the variance, and the y-axis shows principal component 2, which explained 1 % of the variance. The infected samples are depicted with red circles and the mock samples depicted in turquoise. (B) Histogram of read sizes from pooled mock and infected samples. Inset: a representative picture of *Brassica napus* leaves under both treatments (24 hours post-inoculation (HPI)). Left: for the pooled replicates of the mock sample, y axis depicts the percentage of reads across all three replicates and the x axis read length in nucleotides. Right: for the pooled replicates of the infected sample, same information as for the mock sample. (C) Results from pooled replicates for the mock

(left) and infected (right). Showing percentage of reads (y-axis) of each size class (x-axis) that had each of the four nucleotides (AGCU) in their 5' position. 179

Fig. 5. 2. Small RNA population identified from *Brassica napus* genome in response to *Sclerotinia sclerotiorum* infection. (A) A flow chart showing total, Dicer-derived and highly expressed loci with a major RNA reads of ≥ 100 reads, as predicted by ShortStack, and differentially expressed loci identified from DESeq2. (B) Histogram of read sizes from upregulated and downregulated sRNA loci, y-axis depicts the percentage of reads in each category (upregulated or down-regulated) and the x axis read length in nucleotides. (C) Histogram of read sizes showing percentage of reads (y-axis) of each size class (x-axis) that had each of the four nucleotides (AGCU) in their 5' position for upregulated loci (left) and downregulated loci (right). 181

Fig. 5. 3. Representative Target plots (T-plots) for infection-specific targets of upregulated sRNAs. The x-axis shows the transcript position in the target genes and the y-axis shows the 5' read coverage at different positions; cleavage sites predicted by PARESnip2 are labeled in red. The category and p value given by PARESnip are shown in the top left-hand corners of graphs and the gene IDs and their putative functions are shown above. 183

Fig. 5. 4. Prediction of infection-responsive microRNAs from the *Brassica napus* genome. (A) Histogram of the 73 conserved miRNA families. The y-axis shows the identified conserved miRNA family name and the x-axis shows the number of sequences (isomiRs) identified for each miRNA family; red bars are the significantly differentially expressed miRNA families. An isomiR is one of a family of highly similar miRNA sequences derived from either the guide or passenger strand. Green miRNA families were not differentially expressed. Those in other colours contained differentially expressed miRNAs and the colours correspond with D. (B) Histogram of read sizes from 529 conserved miRNAs. The y axis depicts the percentage of reads and the x-axis read length in nucleotides. (C) Histogram of the 529 miRNAs showing percentage of reads (y-axis) of each size class (x-axis) that had each of the four nucleotides (AGCU) in their 5' position. (D) 20 differentially expressed miRNAs with $\log_2(\text{fold change})$ on the y-axis. Colours correspond with miRNA families in A. 189

Fig. 5. 5. . Degradome validation of a TAS gene triggered by miRNA1885. (A) Target plot (T-plot) of the cleavage site on a TAS gene putatively targeted by miRNA1885. The transcript position in the target gene is on the x-axis and the y-axis shows the 5' read coverage at each position; the cleavage site predicted by PARESnip2 is labelled in red. The category and p value of the PARESnip2 test are given in the top left-hand corner of the graph. (B) T-plot of cleavage site on a galactose oxidase gene putatively targeted by one of the ta-siRNAs produced by the miRNA 1885-triggered TAS gene. The x-axis shows the transcript position in the target gene and the y-axis shows read 5' coverage at each position; the cleavage site predicted by PARESnip2 is labelled in red. The category and p value from PARESnip2 are given in the top right-hand corner of the graph. (C) A heat map of 9 differentially expressed PHAS loci plotted with normalized counts from DESeq2. 193

Fig. 5. 6. 5'Rapid amplification of cDNA ends (RACE), degradome result, and qPCR for an ethylene response factor gene putatively cleaved by a novel sRNA. (A) Gel electrophoresis of the 5'RACE result showing a band of the correct size; the second lane is a no template control. (B) Sequence complementarity of the sRNA and its target. The blue arrow shows the cleavage site identified from sequencing the 5'RACE product and the red arrow shows the cleavage site identified with degradome sequencing. (C) Target-plot (T-plot) of the degradome result of the 5'RACE validated gene showing the transcript position in the target gene (x-axis) and read 5' coverage at each position (y-axis); the cleavage site predicted by PARESnip2 is labelled in red. (D) RT-qPCR result of target gene in mock and infected sample (x-axis) and relative expression of gene to the house keeping actin gene. 197

Fig. 6. 1. Plasmid used for the assembly of transformation constructs showing the restriction site, hygromycin cassette and background. The yellow coloured arrows represent hygromycin cassette for fungal selection, neomycin phosphotransferase III coding sequence, and trfA (trans-acting product essential for vegetative plasmid replication), the green coloured arrows represent PtrpC, is the *Aspergillus nidulans* tryptophan biosynthesis protein constitutive promoter to drive expression of the hph gene, the red coloured arrows represent TtrpC, a tryptophan terminator from *Aspergillus nidulans*, and blue the coloured arrows represent origin of replication. 221

Fig. 6. 2. Neighbor-joining consensus tree based on the alignment of RNAi protein sequences of 5 filamentous fungi. The protein sequences were retrieved from funRNA database. 225

Fig. 6. 3. Confirmation of true construct with Sanger sequencing. Consensus is the original construct and coverage is 5' left and 3' right flank mapped with the tool map to reference in Geneious prime. The aligned file is sequenced file obtained from the sanger sequencing..... 226

Fig. 6. 4 Comparison of *Sclerotinia sclerotiorum* mycelium morphology in wild type (WT) and gene disruption mutants. A. Photographs of mycelium length taken after 24 hours of incubation. B. Box plot showing mycelium measurement after 24 hours of incubation. Each coloured dot represents data recorded for each of the replicates. C. Photographs of sclerotia formation after two weeks of incubation. D. Box plot showing sclerotia number. Each coloured dot represents the data recorded for each of the replicates and upper, middle and lower quartiles..... 229

Fig. 6. 5. Effects of Dicer1 gene disruption on *Sclerotinia sclerotiorum* CU8.24 pathogenicity in *Brassica napus* plants. Lesion length at 24 HPI (A) and 48 HPI (B). 231

Table 1.1. Different effectors from *Sclerotinia sclerotiorum* causing *Sclerotinia* stem rot. ... 24

Table 1.2 List of pathogenic and non-pathogenic fungi for which small RNAs studies have been conducted..... 28

Table 1.3 Identified conserved miRNAs in <i>Brassica napus</i> with a role in oil content, seed development, response to <i>Sclerotinia sclerotiorum</i> infection or response to cadmium, salt or drought stress.	41
Table. 3. 1. Description of 24 miRNA loci with mature miRNA and miRNA* sequences with homology to known miRNA sequences in miRBase identified in the <i>Sclerotinia sclerotiorum</i> CU8.24 genome, using the sRNA sequencing data and the miRCat software.....	109
Table. 3. 2. Identification of phasiRNAs producing loci in <i>Sclerotinia sclerotiorum</i>	112
Table. 3.3. Summary of endogenous targets predicted from <i>Sclerotinia sclerotiorum</i> sRNAs. The number of targets that overlapped genes or repeat elements in the <i>S. sclerotiorum</i> CU8.24 genome using second approach with the input of both gene and repeat element files.....	115
Table. 4.1. The list of <i>Sclerotinia sclerotiorum</i> smallRNAs and their <i>Brassica napus</i> gene targets with their predicted cleavage site position in base pairs using psRNAtarget server and their validated cleavage site by degradome sequencing.	152
Table 5. 1. Summary of the sequencing data	177
Table 5.2 Predicted degraded <i>Brassica napus</i> targets of upregulated <i>B. napus</i> small RNAs based on degradome sequencing data analysed using PARESnip2.....	184
Table 5.3. An overview of the characteristics of PHAS loci identified using PhaseTank. ...	195

Chapter 1: Literature Review

1.1 Introduction

Canola is a form of *Brassica napus* (L.) ascribed after 1970 following the development of cultivars with low erucic acid (<2%) and low glucosinolate (30 $\mu\text{mol g}^{-1}$) content. Canola oil is now the third most important source of vegetable oil after palm and soybean [1]. Canola is affected underground and above ground by multiple pathogens including bacteria, viruses, phytoplasmas, nematodes and fungi [2]. The focus of this project is Sclerotinia stem rot (SSR), which is an important fungal disease that causes large economic losses for canola growers. Losses of 0.5% of the potential yield (equivalent to 12.75 kg/ha) have been estimated for every unit percentage of SSR incidence [3]. In Western Australia, annual losses to the canola industry are estimated to be \$AUD 23 million [4]. SSR is caused by *Sclerotinia sclerotiorum* (L.) de Bary, a fungus that has a broad host range of more than 400 species. These include legume crops like bean, peas and soybean, along with oilseed crops like sunflower and canola [5]. Hosts can be infected in two ways, via myceliogenic germination (basal infection) and carpogenic germination [6]. Myceliogenic germination leads to the infection of ground level tissues via hyphae, while carpogenic germination affects above ground tissues via infection of petals by ascospores [7]. Infection via ascospores occurs when ascospores infect senescent blossoms which serve as an initial source of nutrition for the pathogen. Once *S. sclerotiorum* colonizes the senescing petals, it forms microscopic aggregates of appressoria, commonly known as infection cushions [8]. Infection cushions penetrate the healthy stem tissue of the host. As *S. sclerotiorum* grows through the stem tissue it causes characteristic necrotic lesions that give rise to its common name, Sclerotinia stem rot (SSR). The pathogen can spread throughout the stem and complete its lifecycle by producing sclerotia in the stem. The sclerotia

are retained in the stubble or released into the soil from the stems, becoming an inoculum source for the subsequent crop. Stem infection is the main cause of a loss in yield.

Various methods have been implemented to prevent and manage this disease worldwide. The primary control methods include both cultural methods and fungicide application. Development of durable resistant varieties, together with the current management practices (cultural and chemical methods), is important for maintaining a profitable canola oilseed industry [9].

1.2 Molecular mechanisms of host and pathogen interactions

Pathogens secrete various molecules to suppress the plant immune system. These molecules are commonly called effectors. Similarly, plants produce a large arsenal of molecules to combat the pathogen attack. Therefore, there is warfare between plants and pathogens during host and pathogen interactions where complex molecular networks are at play. Typically, there are two defence mechanisms that plants exhibit under pathogen attack. The first layer is called pathogen-associated molecular pattern (PAMP)-triggered immunity (PTI) and the second layer is effector-triggered immunity (ETI) [10]. Once a pathogen infects the host, PTI is activated through recognition of conserved microbial features termed PAMPs, such as fungal chitin or bacterial flagellin [11]. This is effective at resisting the majority of pathogens that are encountered. However, successful pathogens produce effectors that can suppress PTI resulting in effector-triggered susceptibility. In turn, plants have evolved resistance proteins that can recognise the presence of effectors and activate ETI, a stronger defence response that often culminates in localised cell death termed the hypersensitive response (HR). In resistant plants, PTI and/or ETI prevent pathogen spread and disease symptoms, whereas in susceptible plants,

the pathogen can overcome defence responses and cause disease symptoms, which can lead to the complete destruction of plants and yield loss in crop species.

Pathogens can be classified broadly into three categories based upon their infection modes in the host: biotrophic, hemi-biotrophic and necrotrophic. Biotrophs require a living host to derive nutrients from, whereas necrotrophs kill host cells and derive nutrients from the resulting dead host tissue. Hemi-biotrophs have a short biotrophic lifestyle and then switch to a necrotrophic lifestyle.

Pathogens of any lifestyle produce molecules termed effectors that generally aid the successful infection of a host. Biotrophic pathogens often secrete virulence effectors into the host which generally function to suppress host immunity and promote disease. These effectors can be recognised by corresponding plant disease resistance genes which trigger ETI. In this case, the effectors become known as avirulence factors. In contrast, necrotrophs often secrete necrosis-inducing effectors which are recognised by the products of corresponding sensitivity genes in the host [12]. The molecular mechanisms of plant responses to biotrophs are well understood, where resistance genes recognise effectors and trigger the plant immune response. In contrast, resistance to necrotrophic pathogens appears more complex and is often conferred by multiple genes with small effects and is commonly referred to as quantitative resistance.

To date, characterized effectors have mostly been proteins and secondary metabolites. Biotrophic pathogens secrete effectors that, in general, suppress host immune responses. Effectors of hemibiotrophs and necrotrophs both suppress host immune responses and elicit host cell necrosis.

In *S. sclerotiorum* several protein and secondary metabolite effectors have been identified. Oxalic acid (OA) is a potent secondary metabolite effector for SSR infection in *B. napus*, and has been used to screen plants for resistance to *S. sclerotiorum* [13]. In addition, there are various other effectors that have been identified. Wang et al. (2009) identified the SsPG1D pathogenicity factor in *S. sclerotiorum* by employing a yeast two hybrid assay, where SsPG1D was reported to bind with the C2 domain signaling protein from plants [14]. Zhu et al. (2013) identified SsITL, a fungal integrin-like protein that is involved during early-stage infection of SSR via suppression of the jasmonic acid/ethylene signalling pathway in plants [15]. Liang et al. (2013) confirmed the SsV263 secreted protein as a virulence factor for *S. sclerotiorum* by demonstrating that mutant strains containing a disrupted *SsV263* gene were less virulent on susceptible *B. napus* hosts [16]. Similarly, Xiao et al. (2014) adopted knockout experiments to confirm Ss-CAF1 as a crucial pathogenicity factor of *S. sclerotiorum*. Absence of a functional Ss-CAF1 protein leads to failure of the pathogen to induce lesions and appressorium formation even in the presence of oxalic acid, illustrating its role in infection [17]. Lyu et al. (2016) found SsSSVP1, a vital virulence factor that triggers host plant death by binding with the plant QCR8 protein, a subunit of the cytochrome complex in the mitochondrial respiratory chain in plants [18]. Recently, Seifbarghi et al [19] identified six necrotrophic effectors in *S. sclerotiorum*. The secretion of five of these effectors has been demonstrated to require plant receptor like kinases. Details about these *S. sclerotiorum* effectors are summarized in Table 1.1.

Table 1.1. Different effectors from *Sclerotinia sclerotiorum* causing Sclerotinia stem rot.

Effector	Class	Domains	Probable Signaling pathway	Targeted host	References
Oxalic acid	Metabolite	Ca ²⁺ chelation	Oxidative burst	<i>Nicotiana</i>	[20]
				<i>benthamiana</i>	
				<i>Glycine max</i>	
SsPG1D	Protein	Endopolygalacturonase	Disruption of plasma membrane and nucleus	<i>Allium cepa</i>	[14]
SsITL	Protein	Integrin protein	Suppression of jasmonic/ethylene,	<i>Arabidopsis</i>	[15, 21]
			Salicylic acid	<i>thaliana</i>	
SsV263	Protein	Hypothetical secreted protein	NA	<i>Brassica napus</i>	[16]
SS-CAF1	Protein	Ca ²⁺ binding motif	Not known	<i>Nicotiana benthamiana</i>	[17]
SsSSVP1	Protein	Cysteine rich Secreted protein	Formation of dimer with a subunit of cytochrome complex in mitochondrial respiratory chain.	<i>Nicotiana</i>	[18]
				<i>benthamiana</i>	

In addition to proteinaceous effectors, several fungi are known to secrete small RNA effectors, which target plant genes to manipulate and colonize their hosts. Small RNAs (sRNAs) are non-coding RNAs with a size range of 18-30 nt and they regulate gene expression by silencing their target genes by sequence complementarity. The role of sRNAs in plant development, metabolism, and responses to biotic and abiotic stress are extensively investigated in various studies [22-26]. However, the involvement of pathogen sRNAs during invasion of the host is quite a new area of research.

The subsequent sections of this chapter cover the importance of sRNAs as pathogenicity factors in different host-pathogen interactions and discuss the importance of host sRNAs in regulating plant immune response genes under biotic and abiotic stress. Furthermore, a systematic review of different bioinformatic methods used to study small RNAs is conducted.

1.3 Small RNAs (sRNAs) and their roles in plant disease

Small RNAs are double stranded RNAs between 18 to 30 nucleotides in length produced by endonucleases in the Dicer family (DCLs). They are loaded into a complex with ARGONAUTE proteins (AGOs), termed the RNA-induced silencing complex (RISC), to target either coding RNAs or non-coding RNAs by sequence complementarity. Depending on the nature of the target transcript and the AGO types involved, this process might lead to target cleavage and degradation, translational repression or recruitment of additional cofactors [27]. Therefore, sRNAs have major roles in silencing other RNAs. Among the sRNAs are two major groups, the microRNAs (miRNAs) and the short interfering RNAs (siRNAs). miRNAs and siRNAs differ in terms of their precursors, with the former derived from hairpin-shaped single

stranded RNAs whereas the latter are produced from double stranded RNAs [28]. siRNAs can be further classified into secondary siRNAs which are produced in a distinctive phased configuration called phasiRNAs. The loci that generate phasiRNAs are called PHAS loci, which are cleaved at first by AGO1 RISC loaded with the 21-22 nt miRNAs [29]. Trans-acting siRNAs (ta-siRNAs) are a special class of siRNAs that are produced from *TAS* genes and act in *trans* to regulate target transcripts [30].

While the endogenous roles of sRNAs have been well characterized in plants, a picture is emerging for the role of sRNAs derived from microbial pathogens during host infection [23, 31-33]. Small RNAs produced by some fungi have been demonstrated to suppress the plant immune system by silencing host immune system genes through cross-kingdom interference [19]. Small RNAs released by pathogens can translocate into plant cells and aid infection; these have been termed sRNA effectors. In *Botrytis cinerea*, a necrotrophic fungal pathogen closely related to *S. sclerotiorum*, several sRNAs were identified that are able to manipulate the host defense response [34, 35]. Weiberg et al. [34] characterized three promising Bc-sRNA putative effectors for validation out of 73 Bc-sRNAs that were predicted to target host genes in *Arabidopsis* and tomato. Wang et al. [35] validated the function of the small RNA Bc-SiR37. Overexpression of Bc-SiR37 resulted in down-regulation of eight predicted target genes in *Arabidopsis*. Knockout mutant lines of Bc-SiR37 targets enhanced disease susceptibility. Target genes in *A. thaliana* included transcription factors, kinases and cell wall modifying enzymes. Thus, this finding encourages researchers to look for sRNA effectors in other pathogens. The schematic diagram of cross-kingdom RNAi is shown in Figure 1.1.

1.4 Small RNA-mediated cross-kingdom silencing

During cross-kingdom silencing, sRNAs are generated by the pathogen and translocated into the host where they hijack the host Argonautes and silence the host mRNAs that are related to plant immune response genes or vice-versa (Fig. 1.1) [35]. Recently, it has been reported that extracellular vesicles play a key role in transporting sRNAs across two kingdoms [36]. Apart from pathogen sRNAs, plant sRNAs can also move to the pathogen and silence genes [37]. Therefore, sRNAs are potentially a new paradigm that can have a potent role in plant disease during host-pathogen interactions. For example, *Arabidopsis* small RNAs are transported to *B. cinerea* and may play a role in silencing pathogen genes [37]. Due to the importance of sRNAs in the regulation of biological processes, *in silico* and wet lab studies have been conducted to investigate the role of sRNAs in different pathogenic and non-pathogenic fungi (Table 1.2). The following paragraphs will focus on the pathogens for which sRNA profiling following host infection has been conducted.

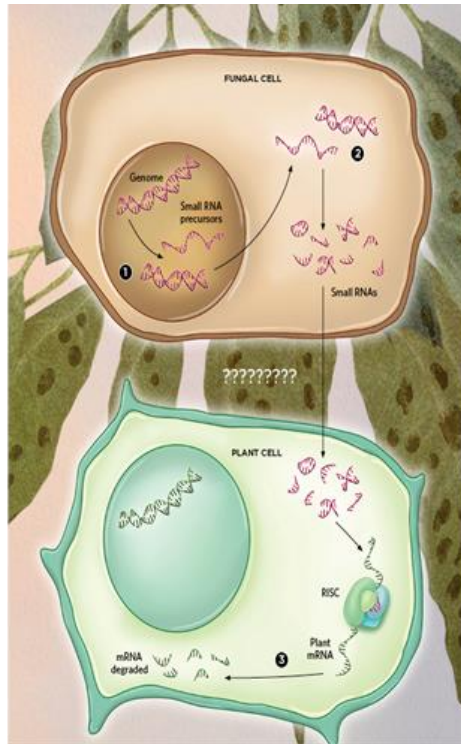


Fig. 1. 1. A schematic diagram showing cross-kingdom silencing mediated by small RNAs.

Source Infographic Cross-Kingdom RNAi | The Scientist Magazine® (the-scientist.com)

Table 1.2 List of pathogenic and non-pathogenic fungi for which small RNAs studies have been conducted.

Species	Pathogenicity on plants
<i>Mucor circinelloides</i> [4]	Non-pathogenic
<i>Neurospora crassa</i> [38]	Non-pathogenic
<i>Cryptococcus neoformans</i> [39]	Non-pathogenic
<i>Trichoderma reesei</i> [40]	Non-pathogenic

<i>Penicillium marneffei</i> [41]	Non-pathogenic
<i>Pennicillium chrysogenum</i> [42]	Non-pathogenic
<i>Aspergillus flavus</i> [43]	Non-pathogenic
<i>Metarhizium anisopliae</i> [44]	Non-pathogenic
<i>Rhizophagus irregularis</i> [45]	Non-pathogenic
<i>Zymoseptoria tritici</i> [46]	Pathogenic
<i>Sclerotinia sclerotiorum</i> [47, 48]	Pathogenic
<i>Botrytis cinerea</i> [49]	Pathogenic
<i>Fusarium oxysporum</i> [25]	Pathogenic
<i>Puccinia striiformis</i> [50]	Pathogenic
<i>Puccinia tritici</i> [41]	Pathogenic
<i>Fusarium graminearum</i> [51, 52]	Pathogenic

1.4.1 *Fusarium oxysporum* and banana

Fusarium oxysporum f. sp. cubense is an important pathogen of banana causing significant economic loss worldwide. sRNA profiling in the Cavendish variety during infection with an aggressive and a non-aggressive strain of *F. oxysporum* identified 122 miRNAs across all infected libraries [25]. A total of 22 miRNA families were upregulated in the aggressive pathotype during infection while no significant differences were observed between the non-aggressive pathotype and mock samples, suggesting the enrichment of miRNAs during banana root infection by aggressive *Fusarium* pathotypes. The expression of three upregulated miRNAs (miR166, miR159 and miR156) were further validated with qRT-PCR. These miRNAs were reported to regulate plant defense pathways in the host such as the salicylic acid signalling pathway. Target genes of all known and novel miRNAs were enriched for seven biological processes, three molecular functions and six cellular components. These included GO terms related to signalling, response to stimulus, localization, regulation of biological processes, biological regulation, metabolic processes and cellular processes. Regulation of transcription factors such as cellular macromolecule biosynthetic process, response to stimulus, regulation of primary metabolic process, gene expression, and RNA metabolic process were more pronounced set of genes in biological process suggesting these genes were involved in regulating gene expression during the pathogen invasion.

1.4.2 *Blumeria* spp. and barley

Blumeria graminis causal agent of powdery mildew is a threatening pathogen of cereal crops. *B. graminis* is found in different forms that have distinct host ranges. Kusch et al. (2018) characterized *B. graminis* sRNAs in barley and wheat powdery mildew and predicted their targets in the host plants [53]. In total 1,963 and 1,250 sRNAs were identified in two strains of

B. graminis, namely *Bgt* and *Bgh*, which targeted 1,180 and 524 genes in wheat and barley, respectively. Altogether, 113 and 138 GO terms were enriched in wheat and barley target genes, respectively. The most represented process in the wheat genes were related to fatty acid catabolism, fatty-acyl-coA transport, ubiquinone biosynthesis, and seed germination while macromolecule-catabolic process was the most enriched in barley. Altogether, 591 and 506 endogenous targets were identified from *Bgt* and *Bgh*, respectively. While several endogenous target sites in the pathogen were reported, unique target sites present only in the host were identified. This suggests that cross-kingdom silencing may be occurring during powdery mildew infection of cereal crops like barley and wheat.

In another study, sRNAs were profiled during *B. graminis* infection of barley plants and their target genes were reported. A total of 1,425 and 1,741 miRNA candidates were identified from the barley and *Bgh* genomes. In addition to sRNA sequencing, degradome sequencing was performed to find the likely cleavage sites of identified sRNAs. The fungal sRNAs were predicted to regulate effectors, metabolic genes, and translation-related genes.

1.4.3 *Fusarium graminearum* and wheat

Fusarium head blight caused by *Fusarium graminearum* is a devastating disease of wheat that causes significant economic and agronomic losses to growers worldwide [54]. Jian et al. [51] identified a total of 4,139 *F. graminearum* sRNAs with 264 sRNAs potentially targeting wheat genes. A wheat gene encoding a chitin elicitor binding protein (*TaCEBip*) gene, which has a likely function in wheat disease resistance signalling was chosen for further validation. The interaction between a fungal sRNA and *TaCEBip* was analysed with the expression analysis of

target genes, western blotting of encoded protein of target genes during infection, and YFP fluorescence analysis of YFP-tagged target genes and co-expressed with sRNAs in host tissues. Altogether, these results suggested the presence of cross-kingdom silencing in the *F. graminearum*-wheat pathosystem.

1.4.4 *Botrytis cinerea* and Arabidopsis

Weiberg et al [34] provided the first experimental evidence of cross-kingdom silencing in the fungal pathogen *B. cinerea*. *B. cinerea* is a necrotrophic fungus that commonly causes gray mold disease in more than 200 species. A total of 832 sRNAs were predicted from *B. cinerea*-infected Arabidopsis and tomato plants. Among them, 73 *B. cinerea*-derived sRNAs were predicted to target genes in both host species including mitogen-activated protein kinases (MAPKs), which have a potential role in plant immunity. Three fungal sRNAs were further tested for their possible role as virulence factors. Repression of MAPK genes in both hosts was investigated with RT-qPCR in the presence of fungal sRNAs. In addition, a co-infiltration assay was conducted to demonstrate the fungal sRNA-mediated suppression of plant target gene expression. The observation was further expanded with another sRNA which was predicted to regulate 15 genes in Arabidopsis including those related to immunity-like WRKY transcription factors, receptor-like kinases, and cell wall-modifying enzymes [35]. These genes exhibited lower expression upon *B. cinerea* infection. The transient expression assay demonstrated reduced expression of host mRNA when co-expressed with the corresponding fungal sRNA, indicating specific sRNA-induced cleavage of the host transcript. Furthermore, an *AGO1* deletion Arabidopsis line did not exhibit the symptoms manifesting the role of AGO1 in the loading of fungal sRNAs which further demonstrated the role of fungal sRNA effectors in hijacking the host RNAi machinery. The observation was further expanded with another *B.*

cinerea sRNA using similar techniques which suppressed the genes related to WRKY transcription factors, receptor-like kinases, and cell wall modifying genes [35].

1.4.5 *Zymoseptoria tritici* and wheat

Xian et al [46] demonstrated no evidence that *Zymoseptoria* sRNA effectors function in silencing plant immune-related genes in wheat. A total of 33 highly expressed *Zymoseptoria* sRNAs during wheat infection were identified. Altogether, 139 genes, which were likely targeted by these sRNAs were downregulated as evident from an RNA sequencing study. However, there is no evidence of cleavage of these targets from degradome datasets, which suggests the lack of evidence for cross-kingdom RNAi. These downregulated target genes were overrepresented for chlorophyll-related process and pigment biosynthesis. Also, wheat target genes related to “microtubule-based movement” were also significantly enriched. This biological process was reported to affect the plant response to pathogen infection [55]. Furthermore, a receptor-like kinase was also present in the set of sRNA target genes. The expression of fungal sRNAs was reported to be lower *in planta* than in *in vitro* samples most probably due to the expression of only one fungal *Dicer* gene during infection. Perhaps the identified sRNAs are important for endogenous regulation of genes rather than silencing of host genes.

1.4.6 *Sclerotinia sclerotiorum* and its host plants

Sclerotinia stem rot is caused by the fungal pathogen *S. sclerotiorum*, which infects more than 400 plant species. With its broad host range and necrotrophic nature, it is a difficult disease to manage. Derbyshire et al [48] predicted 374 highly abundant sRNAs expressed during *S. sclerotiorum* invasion of *Arabidopsis* and *Phaseolus vulgaris*. Most of these sRNA loci

overlapped with transposable elements in *S. sclerotiorum*. Among these highly expressed sRNAs, only 52 sRNAs were significantly upregulated during host infection relative to *in vitro* samples. The top 1 % of sRNA target genes were enriched for signalling with GO terms associated with GO:007165 (signal transduction), GO:0019199 (enzyme binding), GO:0019199 (transmembrane receptor protein kinase). Infection assays with Arabidopsis knockout lines of sRNA target genes (*SERK2* and *SNAK2*) revealed these genes are important for resistance to *S. sclerotiorum*; this aligns with the hypothesis that *S. sclerotiorum* silences these genes during infection to render the plant more susceptible. Altogether, this information provides clues that *S. sclerotiorum* may be producing sRNAs to silence *A. thaliana* genes during infection. However, detailed experimental evidence is needed to prove this phenomenon. This thesis aims to unravel this phenomenon with experimental evidence in the *S. sclerotiorum*-*B. napus* pathosystem. Apart from this, there were two studies that characterized miRNA-like structures in *S. sclerotiorum*. Zhou et al [47] identified 44 miRNA candidates from four different development stages namely, vegetative growth, sclerotial development, myceliogenic germination, and pathogenesis. Furthermore, Xia et al [56] characterized miRNAs involved in sclerotial development and identified a histone acetyltransferase gene as a potential endogenous gene related to sclerotial development. Both of these studies only investigated the presence of miRNAs in *S. sclerotiorum* and did not investigate their potential role in cross-kingdom silencing. However, these studies showed the presence of an RNAi mechanism in *S. sclerotiorum* and opened a new door for further investigation on the role of sRNAs in this pathogen.

1.4.7 *Cuscuta* spp. and Arabidopsis

Apart from sRNA studies in pathogenic and non-pathogenic fungi, their role in obligate parasitic plant interactions has also been investigated. Dodders (*Cuscuta* spp.) are obligate

parasitic plants. Like fungal pathogens, *Cuscuta* species draw nutrients and water from host stems through feeding structures called haustoria. During the interaction of *C. campestris* and *A. thaliana*, 22-nt miRNAs were reported to accumulate [57]. Those miRNAs were able to cleave mRNAs and decrease their expression suggesting the potential role of sRNAs as pathogenicity factors during parasitism. At the point of infection, 76 dodder sRNA loci were significantly upregulated out of which 43 loci were predicted to be of miRNA origin given the formation of hairpin loop structures in these loci. Uncapped mRNAs were more pronounced in parasitized samples when compared to the uninfected samples, further suggesting cleavage of target transcript by miRNAs.

1.5 Role of plant sRNAs in plant immunity response

One of the key functions of RNA silencing in plants is to suppress mRNA transcripts important for pathogenicity of the invading pathogen. The host recruits its RNAi machinery to suppress the RNA of viruses. However, plants have evolved complex defence mechanisms, including PTI and ETI against bacteria and fungi as they do not expose their genomes in host cells during any stage of infection [58]. Therefore, plant sRNAs contribute to PTI and ETI by fine-tuning plant hormone accumulation and/or silencing genes involved in pathogen virulence [59]. While a plant develops an immune system to survive pathogen infection, it also deprives the limited resources available for plant development; therefore, there should be fine tuning of plant hormones. One of the strategies to optimize the regulation of plant hormones during infection is secretion and accumulation of endogenous sRNAs [60]. Different miRNA families and their isomers have been reported to play a role in plant biotic stress responses [23]. Secondary sRNAs called pha-siRNAs also play a key role in regulating endogenous gene expression of plants during pathogen attack [61-63]. Furthermore, miRNAs have been reported to regulate

plant disease resistance genes that contain nucleotide binding site, leucine rich repeat domains (also referred to as NBS-LRR, NB-ARC or NLRs) through the production of phasiRNA [64].

The role of sRNAs in disease responses has been shown in various studies [65-67]. Barley sRNAs were reported to regulate transcription factors encoding auxin response factors, NAC transcription factors, homeodomain transcription factors and several splicing factors [68]. Furthermore, 1,274 pha-siRNA loci with 88.9 % sRNA size length of 24 nt were also identified from the plant genome. Among these pha-siRNA loci, eight and 24 loci were encoded from transcripts related to nucleotide-binding leucine-rich repeat (NBS-LRR) and receptor-like kinase genes. NBS-LRR genes are one of the major sources of pha-siRNAs origin upon binding of specific miRNAs [69]. *Arabidopsis* miR393 was the first miRNA identified to have a role in the plant immune response upon *Pseudomonas syringae* attack [59]. miR393 negatively regulates mRNAs of auxin receptors, transport inhibitor response (TIR1), AFB2 and AFB3 to divert energy to defence signalling over plant growth. In rice, miR393 was also reported to regulate auxin signalling pathways in crown root rot infection [70]. Rice specific miR7695 was upregulated during blast infection [67, 71]. In addition, miR169a, miR172a, and miR398b were also reported to have a role in the basal response of rice against *M. oryzae* [72]. In wheat, *B. graminis* triggered the production of miR167, miR171, miR444, miR408 and miR1338 [65]. Furthermore, in soybean, miR403 was significantly downregulated during *Phytophthora sojae* infection [73]. Altogether, these studies suggest a tight regulation of miRNAs and their target genes during pathogen invasion in host plants.

A detailed summary of the function of sRNAs in response to biotic stress has been discussed in a previous review paper [74]. The next paragraphs discuss the function of sRNAs in *B. napus* in various biological processes including response to biotic stress in *B. napus*.

1.6 The function of small RNAs in *Brassica napus*

Small RNAs have roles in the regulation of various biological processes [28]. In addition, they play a significant role in the plant's response to abiotic and biotic stress [32, 75]. To date, only 92 mature *B. napus* miRNAs have been deposited in miRBase [66] and this section reviews four processes in *B. napus* for which the roles of sRNAs have been investigated.

1.6.1 Oil content

As one of the major oil seed crops, canola plays a key role in the supply of vegetable oil globally [76]. Therefore, the seed oil content of canola is of great importance. Seed oil production depends on the seed yield and seed oil content, which are two of the main breeding targets [77]. Wei et al. [77] investigated the involvement of sRNAs during seed development in canola with sRNA and degradome sequencing data. The comparison of sRNA profiles in two different canola lines varying in seed weights and oil content revealed the sRNA/mRNA regulatory networks. Altogether, 1,276 miRNAs, including 1,248 novel and 28 known miRNAs were predicted from both high seed weight with low oil content and low seed weight and high oil content samples. A total of 57 genes were cleaved from these miRNAs as evident from degradome analysis. Among these degradome validated targets are genes with a potential role in regulation of seed development and nine genes involved with seed oil synthesis. However, only three putative target genes involved in seed development (1) and oil synthesis (2) were regulated by differentially expressed miRNAs. The conserved miRNA bna-miR6029 was significantly downregulated while bna-miR169m was upregulated in the low oil content canola lines. Bna-miR169m regulates genes encoding Ubiquitin ligase E1, Pyruvate dehydrogenase complex E1 and Pyruvate dehydrogenase complex E1 suggesting the role of this miRNA in regulating genes that controls the oil content. Another study also reported the overexpression

of miR169 in low oil content lines [78]. A total of 50 conserved and nine novel miRNAs were identified with 346 putative targets from all conserved miRNAs. Five miRNA families, miR169, miR390, miR394, miR6028, and miR6029 exhibited higher expression level in the low oil content cultivar while miR408* and miR211 were abundant in the high oil content cultivar as confirmed by quantitative reverse transcriptase- polymerase chain reaction (qRT-PCR) and northern blotting. The expression level of miR156 from sequencing did not corroborate the qRT-PCR result. Sequencing showed higher expression of miR156 in higher oil content cultivar; however, qRT-PCR showed similar expression in both lines. Targets of miR156 and miR6029 were reported to relate to *B. napus* oil production. The cleavage of the corresponding targets of these two miRNAs was further validated by a 5' RACE experiment. The comparison of miRNA expression between high oil and low oil content cultivars unveiled the involvement of some of miRNAs in the regulation of *B. napus* seed oil production.

1.6.2 Cadmium stress

Cadmium (Cd) uptake by plants poses a real threat to plant health [79]. Excess Cd accumulation has a negative impact on plant growth and development, including membrane distortion, stunted growth, and inhibition of photosynthesis thereby decreasing the overall yield [80]. miRNAs are reported to control many abiotic stress responses at the post-transcriptional level. In order to identify miRNAs involved in the response to Cd stress, some computational and lab-based studies have been conducted [81-85]. Zhou et al. [82] identified a total of 84 miRNAs from Cd-treated root and shoot samples of *B. napus*. A total of 802 putative targets of these miRNAs were identified by degradome sequencing. Huang et al. [83] identified 12 conserved miRNAs representing nine families. The target genes of miR395, which might have roles in regulating plant abiotic stress responses to Cd, were validated with 5' RACE analysis. In another study, Fu et al. [84] identified 171 miRNAs from 12 sRNA libraries of root control,

root Cd treatment, shoot control, and shoot Cd treatment. Among these, 22 and 29 miRNAs were differentially expressed in the root and shoot, respectively. Expression analysis showed the anti-regulation of 16 miRNA-mRNA interaction pairs in both tissues upon Cd treatment. Jian et al. [85] identified 44 known miRNAs belonging to 27 families and 103 novel miRNAs from Cd-treated *B. napus* samples across 0, 1 and 3 days of treatment. A total of 39 miRNAs were differentially expressed from sequencing data and 13 miRNAs were further validated with qRT-PCR. Eleven target genes regulated from these miRNAs were further tested for their expression changes but only seven genes exhibited reverse expression changes with miRNAs, suggesting that these genes were targeted by the respective miRNAs. The functional annotation of these target candidates was related to transcription factor regulation, biotic stress hormone, ion homeostasis, and secondary metabolism.

1.6.3 Salt and drought stress

Salt and drought stress greatly impact the yield of crops. Drought stress reduces the nutrient diffusion and salinity negatively impact on seed germination, plant growth and development, causing significant losses in crop yield [86]. Jian et al. [87] identified 85 known and 992 novel miRNAs involved in early stage seed germination subject to salt and drought stress in canola. Six miRNA families, namely miR156, miR169, miR860, miR399, miR171, and miR395 were significantly downregulated under drought stress while only miR172 was upregulated. miRNA156 showed the highest abundance in the sequencing dataset which was involved in the downregulation of SQUAMOSA promoter binding protein-like (SPL) transcription factor genes. Only 2 miRNAs, miRNA393 and miR399, were significantly downregulated under salt stress.

1.6.4 Sclerotinia stem rot

There have been three studies conducted so far on the role of *B. napus* sRNAs in response to *S. sclerotiorum* infection [66, 88, 89]. One of the works is an outcome of this thesis which is included in Chapter 4. These studies covered the time-points of 3, 12, 24- and 48-hours post-inoculation.

Altogether 20, 68, and 10 miRNAs were differentially expressed in these studies according to sequencing data [66, 88, 89]. Only bna-miR166 was found to be commonly differentially expressed in Cao et al [88] and Jian et al [89]; however, this miRNA was not significantly expressed in Regmi et al. [66] study. Cao et al. [88] identified 280 *B. napus* miRNA candidates which includes 53 novel and 227 conserved miRNA. miRNA microarray analysis identified 68 differentially expressed miRNAs between *S. sclerotiorum*-inoculated and uninoculated leaves. Expression analysis of six miRNA-mRNA target pairs related to the plant immune response was further investigated with qRT-PCR. Furthermore, the targets of two miRNA miR168a (*BnaC08G46720D*) and bnamir403 (*BnaA05g14760D*) were confirmed with 5' RACE experiments. These miRNAs regulated biological processes related to plant defence to pathogens such as NBS-LRR genes and reactive oxygen species. Moreover, three miRNAs targeted key RNAi component genes *AGO1* and *AGO2*. Jian et al [89] identified 77 known and 176 novel miRNAs with 10 known and 41 novel being differentially expressed. Altogether, 80 cleavage sites were identified from degradome sequencing from these miRNAs. Thirteen miRNAs were further validated with qRT-PCR experiments. From these differentially expressed miRNAs, 15 target genes were differentially expressed in an RNA sequencing dataset.

Table 1.3 Identified conserved miRNAs in *Brassica napus* with a role in oil content, seed development, response to *Sclerotinia sclerotiorum* infection or response to cadmium, salt or drought stress.

miRNAs	Target genes	Annotated information	Biological process	Reference
Novel_bna_miR	<i>GSBRNA2T00056839001</i>	Lipid transport and metabolism	Oil-content*	[77]
Novel_bna_miR	<i>GSBRNA2T00044758001</i>	Cell division, ARF16, ARF16-2	Oil-content*	[77]
bna-miR6029	<i>GSBRNA2T00124698001</i>	Ubiquitin ligase E1, Pyruvate dehydrogenase complex E1, Pyruvate dehydrogenase complex E1	Oil-content*	[77]
bna-miR156	TC93255, TC82011	SPL	Oil-content**	[78]
bna-miR395	<i>BnaAPSI</i> , <i>BnaAPS3</i> , and <i>BnaAPS4</i>	ATP sulphurylases	Cd-stress**	[83]
bna-miR395a	<i>BnaA03g04440D</i>	ABA-binding protein (ABAR)	Cd-stress*	[84]
bna-miR398a-3p	<i>BnaA04g16310D</i>	COPPER/ZINC SUPEROXIDE DISMUTASE 2, CSD2, (CSD2)	Cd-stress*	[84]
bna-miR169_2p	<i>BnaC03g08060D</i>	NAD(P)H dehydrogenase B4 protein (NDB4)	Cd-stress*	[84]
bna-miR6029	<i>BnaA03g11200D</i>	Inner centromere protein	Cd-stress*	[84]
bna-miR398b-3p	<i>BnaA04g16310D</i>	CSD2	Cd-stress*	[84]
bna-miR398a-3p	<i>BnaC04g39580D</i>	CSD2	Cd-stress*	[84]
bna-miR398b-3p	<i>BnaC04g39580D</i>	CSD2	Cd-stress*	[84]
bna-miR398b-3p	<i>BnaC08g42970D</i>	CSD1	Cd-stress*	[84]
bna-miR398b-3p	<i>BnaCmng33420D</i>	CCS	Cd-stress*	[84]
bna-miR2111c-5p	<i>BnaA09g44060D</i>	ACO1	Cd-stress*	[84]
bna-miR395a	<i>BnaC09g46440D</i>	SULTR2;1	Cd-stress*	[84]
bna-miR395b	<i>BnaC09g46440D</i>	SULTR2;1	Cd-stress*	[84]
bna-miR397	<i>BnaC04g07220D</i>	LAC4	Cd-stress*	[84]
bna-miR397	<i>BnaA07g26900D</i>	CDM1	Cd-stress*	[84]
bna-miR408	<i>BnaA04g11970D</i>	PRP2	Cd-stress*	[84]
bna-miR398b-3p	<i>BnaA02g18700D</i>	FLZ6	Cd-stress*	[84]
ath-miR168a	<i>BnaC08g46720D</i>	Argonaut 1	SSR*	[88]
bna-miR403	<i>BnaA05g14760D</i>	Argonaut 2	SSR*	[88]
ath-miR168a	<i>BnaC04g07220D</i>	SGT1a	SSR*	[88]
ath-miR397a	<i>BnaC04g07220D</i>	Nitrite reductase	SSR*	[88]
cme-miR166e	<i>BnaCmng33420D</i>	Copper chaperone for superoxide dismutase	SSR*	[88]
PC-14-5p	<i>BnaA03g53300D</i>	Thioredoxin reductase (NADPH)	SSR*	[88]
bna-miR156a-g	<i>BnaA04g27550D</i> , <i>BnaA06g36780D</i> , <i>BnaA09g27960D</i> <i>BnaC04g02520D</i> <i>BnaC07g17030D</i>	squamosa promoter binding protein-like 15 (SPL15)	SSR*	[89]
bna-miR156d/e/f	<i>BnaAnng03450D</i> , <i>BnaC01g23990D</i>	Phototropic-responsive NPH3 family protein	SSR*	[89]
bna-miR156d/e/f	<i>BnaC06g32630D</i>	Protein of unknown function (DUF579)	SSR*	[89]
bna-miR166f	<i>BnaA02g06170D</i> ,	<i>REVOLUTA (REV)</i>	SSR*	[89]
bna-miR166f	<i>BnaA08g11980D</i>	homeobox gene 8 (HB-8)	SSR*	[89]
	<i>BnaA10g13520D</i>		SSR*	[89]
bna-miR166f	<i>BnaAnng30670D</i>	<i>Homeobox gene 8</i>	SSR*	[89]
Novel miR_87	<i>BnaC03g36760D</i>	<i>Alpha-L-arabinofuranosidase 1</i>	SSR*	[89]
Novel miR_129	<i>BnaA08g11090D</i>	<i>CAN OF WORMS</i>	SSR*	[89]
Novel miR_158	<i>BnaC02g06060D</i>	<i>Casein lytic proteinase B3</i>	SSR*	[89]
miR159_27	<i>BnaA03g22590D</i> , <i>BnaA05g27620D</i> , <i>BnaCmng31260D</i> , <i>BnaCmng49390D</i>	SANT/myb domain	SSR***	[66]
miR5139_2	<i>BnaA02g31560D</i>	Zinc Finger domain	SSR***	[90]
Novel miR_1	<i>BnaA01g27570D</i>	Ethylene response factor	SSR*,***	[90]

* qPCR validation; ** 5' RACE validation; *** Degradome validation.

1.7 Importance of bioinformatics in small RNA analysis

While next-generation sequencing has become a routine method of choice for the investigation of transcriptomes including sRNAs, data analysis remains a challenge. A wide spectrum of computational tools is required to interpret relevant biological information from sRNA sequencing experiments. Therefore, various software packages have been developed to identify sRNAs and their targets and these packages can each give slightly different results. This section reviews the methods and software available for sRNA studies along with their advantages and limitations. The typical bioinformatic pipelines involve pre-processing, quality control, identification of known and novel miRNAs, prediction of differentially expressed sRNAs, and their target prediction. A schematic of the bioinformatic pipeline is shown in Fig. 1.2.

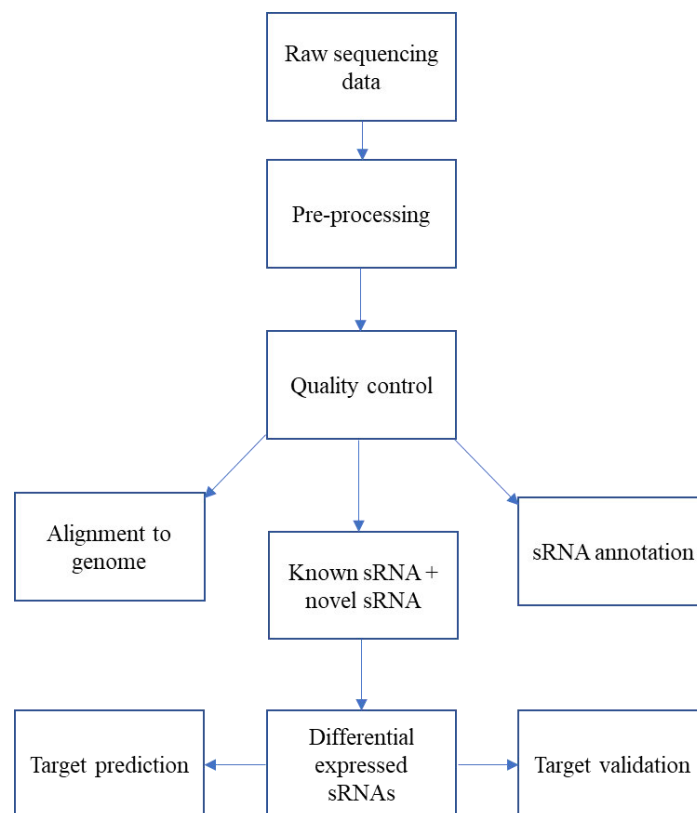


Fig. 1. 2. Bioinformatics analysis of small RNA sequencing analysis

1.7.1 Pre-processing and quality control of sRNA sequencing dataset

Like any other sequencing data, sRNA reads need to be pre-processed before their alignment to a genome. The pre-processing comprises of the removal of adaptor and barcode, size selection, removal of ambiguous reads and generation of unique reads. Cutadapt [91] and trimmomatic [92] are the two most commonly used standalone software packages for adaptor trimming and size selection of sRNA reads. Cutadapt is written in Python and can be operated on Ubuntu Linux, but also on Windows and Mac. Unlike other standalone software packages, it supports FASTQ, FASTA and SOLiD. csfasta/. qual input files. Trimmomatic is Java-based software and can be run on the command line. The impact of these two tools on the quality of data pre-processing has been assessed where no significant changes were observed [93]. In addition to this, programs like ShortStack [94] and the UEA small RNA workbench [95] also have adapter removal packages built in.

While working on reads generated from infected samples, it is important to know whether the sRNA reads are derived from the fungal pathogen or from the host plant. Therefore, some studies [96, 97] split the reads and assign the best matched reads to either pathogen or host using the program called bbsplit [98]. For the annotation of novel and conserved miRNA loci, sRNA reads matching to structural RNAs (ribosomal RNA, transfer RNA or small nuclear RNA) are removed using the Rfam database, as the reads matching to these loci are potentially generated as non-specific degradation products. Rfam is an open-access database which provides information about structural RNAs from a large set of organisms ranging from prokaryotic to eukaryotic [99]. Rfam can either be operated as a web-based platform or through the command line. The Infernal program was developed for RNA homology research and annotation of structural RNAs in sRNA sequencing datasets and can be run on the command line [100]. Fastqc is a java-based program which is mostly used to check the overall quality of

the sequencing dataset [101]. The Dicer-derived origin of sRNA reads results in a unique peak for sRNA length [53] while degradation products give uniform sRNA length; therefore, researchers working on sRNAs investigate the size for their sequencing data and report the typical peak for their data [24, 46, 53, 68, 77].

1.7.2 sRNA alignment and annotation

Different software packages have been developed for the alignment of sRNA reads to a reference genome [94, 95, 102]. SOAP (short oligonucleotide alignment program) is a command-driven program and was developed for efficient gapped and un-gapped alignment of short reads to a genome [102]. SOAP can be operated with a certain number of mismatches or one continuous gap for alignment. The output alignments consist of the best hit with minimal mismatches. SOAP is simply an alignment program and does not give the quantification and annotation of sRNA-producing loci. ShortStack, however, was developed for sRNA sequencing analysis, which investigates overall genome-wide discovery of sRNA-producing loci, and their quantification and categorizes them as Dicer-derived and non-Dicer-derived [94]. Dicer-derived and dicerCall loci are reported based on the distribution of sRNA size and might not reflect the true biogenesis of sRNAs. It also predicts the miRNA-producing loci when run in hairpin mode. The annotation of sRNA loci is based on the sRNA size composition, strandedness, and assignment of multimapping reads to the most likely position. It gives the output in tabular format with the features of all identified loci and counts of each locus which can be further used for differential expression analysis.

The UEA sRNA toolkit is a combination of web-based tools that annotate miRNAs, identify sRNA loci, predict sRNA targets and embeds visualization tools for secondary structures of

miRNAs. However, it does not report sRNA size distribution and non-miRNA hairpin loci. Each of these programs are run separately rather than on the command line [95]. In addition to this, separate programs for the prediction of conserved and novel miRNAs were developed.

A computational method for miRNA discovery relies on the secondary structure of RNA and the ability to form a hairpin loop structure. A detailed review of computational tools in miRNA discovery was previously reported [103]. The commonly used miRNA prediction tools from NGS data include miRDeep [104], miRanalyzer [105], and SSCprofiler [106]. The former two tools have mostly been used for miRNA discovery in plants and fungi while the latter one was originally developed to find cancer-associated miRNAs. The miRDeep software discovers both novel and conserved miRNAs employing miRNA biogenesis properties based on the compatibility of the position and frequency of sRNA reads with the secondary structure of the precursor miRNAs. This program assumes that a true high confident locus must have mature miRNA with higher sequence read matches to stem loop and less frequent reads to of the stem loop called the miRNA star sequence which is normally degraded during the RNAi process. The miRanalyzer package detects known miRNAs from miRBase, finds matches to transcripts and other non-coding RNAs and predicts new miRNAs from the remaining sequences. SSCprofiler uses a similar approach which uses sequence conservation to identify novel miRNAs.

1.7.3 Differential expression analysis

DESeq2 [107] and EdgeR [108] are two common tools used for analysis of differential expression of sRNAs. DESeq2 employs a “geometric” normalization strategy, EdgeR uses a weighted mean of log ratios-based method. A DESeq2 scaling factor is computed as the median of the ratio, for each gene, of its read count over its geometric mean. The main assumption is

that non-DE genes should have similar read counts over all samples. An EdgeR calculates scaling factor to represent sample-specific biases called Trimmed Mean of M-values. This factor is multiplied by each library size to give the significant library size.

EdgeR uses exact test to find differential expressed genes which is based on qCML methods. The exact p-values are calculated by summing all sums of counts that have a probability less than the probability under the null hypothesis of the observed sum of counts.

1.7.4 Target prediction

Various sRNA target prediction tools have been developed based on the concept of seed match, conservation, free energy and site accessibility [109]. miRanda [110], Targetscan [111], TargetMiner [112], miRU [113], and psRNAtarget [114] are some of the most frequently used target prediction tools. The earliest miRNA target predictor miRanda was initially developed to find targets in *Drosophila* and later it was used to find targets in humans [115].

miRanda was written in C and provided as source code. The use of the command line for operating this program poses a technical barrier for novice users, which is one of the drawbacks of this program. TargetScan predicts targets based on seed match and conservation and allows users to search miRNA name, gene name, conserved and poorly conserved miRNA families across several species. Unlike miRanda it is user-friendly as it has a GUI and helpful for novice users who are not proficient in using the command line. TargetMiner predicts targets based on potential seed sites (conserved heptametrical sequences situated mostly at positions 2-7 from miRNA 5'-end) between input sRNA and mRNA of choice. Both GUI and executable versions of Target miner are available.

While most of these programs have been developed for animal miRNA target prediction, miRU and psRNATarget server were developed for sRNA target prediction in plants as target recognition in animals is significantly different to plants. For animal miRNA target prediction, miRNA-duplex free energy is important while for plant miRNA targets, perfect sequence similarity is required between sRNAs and their corresponding targets.

miRU is a web-based tool able to find all potential targets with the user provided sequences and mismatches. False positives can be removed by decreasing mismatches and ensuring the target complementarity conservation in other plant species. The major disadvantage of miRU is it only accepts a single input sRNA sequence at a time. Therefore, psRNATarget was developed, which investigates complementarity between numerous input sRNAs and their target transcripts using the scoring method originally applied by miRU. It is very user-friendly[113]. psRNAtarget finds targets of sRNAs by estimating mRNA target accessibility, which is the thermodynamic energy required to open the secondary structure of mRNA around the target region for target site exposure.

1.7.5 sRNA/mRNA target validation

While computational tools help to find potential sRNA/mRNA target interactions, these interactions need to be validated using experimental methods. Degradome sequencing has been widely used for cleavage site mapping on sRNA/mRNA target pairs [116]. It is a high throughput method of sRNA target validation which was developed for transcriptome-wide detection of sRNA-mediated degradation products. If the degradation of mRNA is due to the sRNA cleavage activity then first 10 nt of its 5' end must contain sRNA complementarity site since AGO1 mediated cleavage in plant occurs between the 10th and 11th nucleotide. If the uncapped mRNAs are mapped to the relative transcriptome and 12-13 nt are extended upstream,

the full sRNA complementarity site can be captured. Alignment of this extended site to the set sRNAs give the cleavage site on the transcript.

CleaveLand [117] and PAREsnip [118] are two computational tools developed for the processing and analysing of degradome data. Both programs need sRNAs, transcriptome and degradome data as input files. Both programs utilize a mismatch-based scoring scheme inferred from a set of experimentally validated sRNA-mRNA target interactions. While CleaveLand can perform analysis on a smaller set of input sRNA sequences. CleaveLand has a slow algorithm compared with PAREsnip. Therefore, PAREsnip can perform analysis of entire sRNA datasets within a smaller timeframe on a typical desktop computer. CleaveLand was written in C and can be run on a linux machine which requires the installation of the EMBOSS package [119] and the Oligomap program [120] while PAREsnip is a Java program. PAREsnip can be run through the command line in linux and also as a web-based program in the UEA small RNA workbench tool [95]. Both tools generate sRNA target candidates and categorize them into one of four categories. CleaveLand considers fragments with an abundance of 1 during the average coverage calculation while PAREsnip removes them to distinguish true lower abundance peaks from background degradation. Category-0 are those candidates with reads more than one and only one maximum coverage site. Category-1 peaks are those that have greater than one read and a shared maximum when there is more than one peak. Category-2 are those that have greater than 1 read and are below the maximum but higher than the the average fragment abundance on the transcript, Category-3 abundance at the position was equal to or less than the average for the transcript. Category-4 has only one read at the position. Category 1 and 2 targets are the highest confidence targets among other categories.

1.8 Knowledge gap in the literature

Currently, there is a lack of understanding and experimental evidence of *S. sclerotiorum*-mediated silencing of *B. napus* genes, although evidence demonstrates that *S. sclerotiorum* expresses sRNAs in host plants [48]. The target prediction during cross-kingdom silencing was limited to the use of an *in-silico* method psRNATarget server which is based on sequence complementarity and might be liable to false positives. Moreover, this phenomenon is reported to be pathosystem-specific, as some pathogen sRNAs do not play any role in disease as reported in the wheat–*Z. tritici* interaction [46].

Little research effort has focused on how canola sRNAs respond during infection of *S. sclerotiorum*, whilst one study has identified several *S. sclerotiorum* sRNAs that might act as effectors [48]. Furthermore, the number of miRNA families identified so far in canola is quite low when compared to the model plants *A. thaliana* and *Medicago truncatula* and other crops suggesting that there is still a great number of miRNAs to be discovered in *B. napus*. SSR is a devastating disease with ongoing research efforts to understand the molecular mechanism of the *S. sclerotiorum*-canola interaction. The further discovery of miRNAs will advance our understanding of the interaction with their target genes.

This project therefore aimed to investigate and validate sRNAs as pathogenicity factors. At first, sRNA profiling of an Australian isolate *S. sclerotiorum* CU8.24 was done with Illumina-based sRNA sequencing. Target prediction of these sRNAs in *B. napus* was done with an *in silico* method and further expanded with the development of degradome sequencing which provides useful experimental evidence of fungal sRNA-mediated *B. napus* mRNA cleavage. Taking advantage of sRNA and degradome sequencing, three different sRNA interactions were investigated. The first is fungal sRNA-mediated fungal mRNA silencing. The second

interaction is fungal sRNA-mediated plant mRNA cleavage also called cross-kingdom silencing. The third interaction is the plant sRNA-mediated plant mRNA cleavage during pathogen attack. The schematic of these interactions is shown in Fig. 1.3. This thesis also investigates the presence of RNAi machineries in *S. sclerotiorum* and demonstrates knockout experiments of the genes predicted to be involved in RNAi. Together, this thesis provides an insight into the molecular mechanism of *S. sclerotiorum* and *B. napus* interactions at the sRNA level.

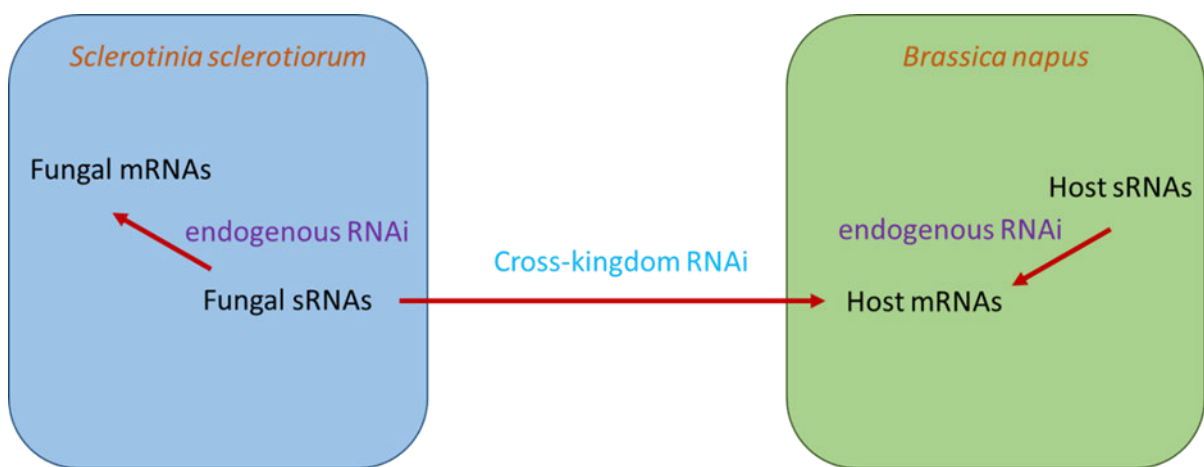


Fig. 1. 3. A model showing possible small RNA interaction during pathogen and host interaction of *Sclerotinia sclerotiorum* and *Brassica napus*.

1.9 Conclusion and thesis objectives

After the identification of the literature gaps described above, this thesis is designed to answer the following objectives:

- Transformation of *S. sclerotiorum* strain with GFP which could facilitate the investigation of how this pathogen interacts with multiple hosts (Chapter 2).

- Provide the experimental evidence of cross-kingdom silencing with the aid of next generation sequencing technology in the *B. napus* – *S. sclerotiorum* pathosystem (sRNA sequencing and degradome sequencing; Chapter 3).
- Validate cross-kingdom sRNA:mRNA interactions experimentally (Chapter 3).
- Elucidate the role of canola sRNAs in manipulating their endogenous genes during *S. sclerotiorum* attack (Chapter 4).
- Investigate the endogenous gene regulation mediated by fungal sRNAs at 12 and 24 HPI (Chapter 5).
- Provide an insight into RNAi machinery in *S. sclerotiorum* through knockout experiments using a homologous recombination technique (Chapter 6).

1. 10 References

1. Carré, P. and A. Pouzet, *Rapeseed market, worldwide and in Europe*. OCL, 2014. **21**(1): p. 1-12.
2. Neik, T.X., M.J. Barbetti, and J. Batley, *Current status and challenges in identifying disease resistance genes in Brassica napus*. *Frontiers in Plant Science*, 2017. **8**.
3. Del Rio, L., et al., *Impact of Sclerotinia stem rot on yield of canola*. *Plant Disease*, 2007. **91**(2): p. 191-194.
4. DAFWA, *DAFWA Annual Report*. 2015.
5. Boland, G.J. and R. Hall, *Index of plant hosts of Sclerotinia sclerotiorum*. *Canadian Journal of Plant Pathology*, 1994. **16**(2): p. 93-108.
6. Derbyshire, M.C. and M. Denton-Giles, *The control of sclerotinia stem rot on oilseed rape (Brassica napus): current practices and future opportunities*. *Plant Pathology*, 2016. **65**(6): p. 859-877.

7. Bardin, S. and H. Huang, *Research on biology and control of Sclerotinia diseases in Canada*. Canadian Journal of Plant Pathology, 2001. **23**(1): p. 88-98.
8. Turkington, T. and R. Morrall, *Use of petal infestation to forecast Sclerotinia stem rot of canola: the influence of inoculum variation over the flowering period and canopy density*. Phytopathology, 1993. **83**(6): p. 682-689.
9. Navabi, Z.K., et al., *Brassica B-genome resistance to stem rot (Sclerotinia sclerotiorum) in a doubled haploid population of brassica napus × brassica carinata*. Canadian Journal of Plant Pathology, 2010. **32**(2): p. 237-246.
10. Dodds, P.N. and J.P. Rathjen, *Plant immunity: towards an integrated view of plant–pathogen interactions*. Nature Reviews Genetics, 2010. **11**(8): p. 539.
11. Navarro, L., et al., *Suppression of the microRNA pathway by bacterial effector proteins*. Science, 2008. **321**(5891): p. 964-967.
12. Wang, X., et al., *The role of effectors and host immunity in plant–necrotrophic fungal interactions*. Virulence, 2014. **5**(7): p. 722-732.
13. Bradley, C., et al., *Response of canola cultivars to Sclerotinia sclerotiorum in controlled and field environments*. Plant Disease, 2006. **90**(2): p. 215-219.
14. Wang, X., et al., *Characterization of a canola C2 domain gene that interacts with PG, an effector of the necrotrophic fungus Sclerotinia sclerotiorum*. Journal of Experimental Botany, 2009. **60**(9): p. 2613-2620.
15. Zhu, W., et al., *A secretory protein of necrotrophic fungus Sclerotinia sclerotiorum that suppresses host resistance*. PLoS One, 2013. **8**(1): p. 1-19.
16. Liang, Y., et al., *Disruption of a gene encoding a hypothetical secreted protein from Sclerotinia sclerotiorum reduces its virulence on canola (Brassica napus)*. Canadian Journal of Plant Pathology, 2013. **35**(1): p. 46-55.

17. Xiao, X., et al., *Novel secretory protein Ss-Caf1 of the plant-pathogenic fungus Sclerotinia sclerotiorum is required for host penetration and normal sclerotial development*. *Molecular Plant-Microbe Interactions*, 2014. **27**(1): p. 40-55.
18. Lyu, X., et al., *A small secreted virulence-related protein is essential for the necrotrophic interactions of Sclerotinia sclerotiorum with its host plants*. *PLoS Pathogens*, 2016. **12**(2): p. e1005435.
19. Seifbarghi, S., et al., *Receptor-like kinases BAK1 and SOBIR1 are required for necrotizing activity of a novel group of Sclerotinia sclerotiorum necrosis-inducing effectors*. *Frontiers in Plant Science*, 2020. **11**: p. 1021.
20. Cessna, S.G., et al., *Oxalic acid, a pathogenicity factor for Sclerotinia sclerotiorum, suppresses the oxidative burst of the host plant*. *The Plant Cell*, 2000. **12**(11): p. 2191-2199.
21. Tang, L., et al., *An effector of a necrotrophic fungal pathogen targets the calcium-sensing receptor in chloroplasts to inhibit host resistance*. *Molecular Plant Pathology*, 2020. **21**(5): p. 686-701.
22. Cabot, C., et al., *A role for zinc in plant defense against pathogens and herbivores*. *Frontiers in Plant Science*, 2019. **10**: p. 1171.
23. Katiyar-Agarwal, S. and H. Jin, *Role of small RNAs in host-microbe interactions*. *Annual Review of Phytopathology*, 2010. **48**: p. 225-246.
24. Wei, W., et al., *Small RNA and degradome profiling involved in seed development and oil synthesis of Brassica napus*. *PloS One*, 2018. **13**(10): p. e0204998.
25. Fei, S., et al., *Small RNA profiling of Cavendish banana roots inoculated with Fusarium oxysporum f. sp. cubense race 1 and tropical race 4*. *Phytopathology Research*, 2019. **1**(1): p. 22.

26. Kettles, G.J., et al., *sRNA profiling combined with gene function analysis reveals a lack of evidence for cross-kingdom RNAi in the wheat–Zymoseptoria tritici pathosystem*. *Frontiers in Plant Science*, 2019. **10**: p. 892.
27. Shanker, A.K. and M. Maheswari, *Small RNA and drought tolerance in crop plants*. *Indian Journal of Plant Physiology*, 2017. **22**(4): p. 422-433.
28. Guleria, P., et al., *Plant small RNAs: biogenesis, mode of action and their roles in abiotic stresses*. *Genomics, Proteomics & Bioinformatics*, 2011. **9**(6): p. 183-199.
29. Liu, Y., et al., *PhasiRNAs in plants: their biogenesis, genic sources, and roles in stress responses, development, and reproduction*. *Plant Cell*, 2020. **32**(10): p. 3059-3080.
30. Felippes, F.F. and D. Weigel, *Triggering the formation of tasiRNAs in Arabidopsis thaliana: the role of microRNA miR173*. *EMBO Reports*, 2009. **10**(3): p. 264-270.
31. Wang, M., A. Weiberg, and H. Jin, *Pathogen small RNAs: a new class of effectors for pathogen attacks*. *Molecular Plant Pathology*, 2015. **16**(3): p. 219-223.
32. Peláez, P. and F. Sanchez, *Small RNAs in plant defense responses during viral and bacterial interactions: similarities and differences*. *Frontiers in Plant Science*, 2013. **4**: p. 343.
33. Baulcombe, D., *Small RNA—The secret of noble rot*. *Science*, 2013. **342**(6154): p. 45-46.
34. Weiberg, A., et al., *Fungal small RNAs suppress plant immunity by hijacking host RNA interference pathways*. *Science*, 2013. **342**(6154): p. 118-123.
35. Wang, M., et al., *Botrytis small RNA Bc-siR37 suppresses plant defense genes by cross-kingdom RNAi*. *RNA Biology*, 2017. **14**(4): p. 421-428.
36. Cai, Q., et al., *Message in a Bubble: Shuttling Small RNAs and Proteins Between Cells and Interacting Organisms Using Extracellular Vesicles*. *Annual Review of Plant Biology*, 2021. **72**: p. 497-524.
37. Wang, M., et al., *Bidirectional cross-kingdom RNAi and fungal uptake of external RNAs confer plant protection*. *Nature Plants*, 2016. **2**(10): p. 1-10.

38. Pickford, A.S., et al., *Quelling in Neurospora crassa*. *Advances in Genetics*, 2002. **46**: p. 277-304.
39. Liu, H., et al., *RNA interference in the pathogenic fungus Cryptococcus neoformans*. *Genetics*, 2002. **160**(2): p. 463-470.
40. Kang, K., et al., *Identification of microRNA-Like RNAs in the filamentous fungus Trichoderma reesei by solexa sequencing*. *PloS One*, 2013. **8**(10): p. e76288.
41. Dubey, H., et al., *Discovery and profiling of small RNAs from Puccinia triticina by deep sequencing and identification of their potential targets in wheat*. *Functional & Integrative Genomics*, 2019. **19**(3): p. 391-407.
42. Dahlmann, T.A. and U. Kück, *Dicer-dependent biogenesis of small RNAs and evidence for microRNA-like RNAs in the penicillin producing fungus Penicillium chrysogenum*. *PloS One*, 2015. **10**(5): p. e0125989.
43. Bai, Y., et al., *sRNA profiling in Aspergillus flavus reveals differentially expressed miRNA-like RNAs response to water activity and temperature*. *Fungal Genetics and Biology*, 2015. **81**: p. 113-119.
44. Zhou, Q., et al., *Genome-wide identification and profiling of microRNA-like RNAs from Metarhizium anisopliae during development*. *Fungal biology*, 2012. **116**(11): p. 1156-1162.
45. Silvestri, A., et al., *In silico analysis of fungal small RNA accumulation reveals putative plant mRNA targets in the symbiosis between an arbuscular mycorrhizal fungus and its host plant*. *BMC Genomics*, 2019. **20**(1): p. 1-18.
46. Ma, X., J. Wiedmer, and J. Palma-Guerrero, *Small RNA bidirectional crosstalk during the interaction between wheat and Zymoseptoria tritici*. *Frontiers in Plant Science*, 2020. **10**: p. 1669.

47. Zhou, J., et al., *Identification of microRNA-like RNAs in a plant pathogenic fungus Sclerotinia sclerotiorum by high-throughput sequencing*. *Molecular Genetics and Genomics*, 2012. **287**(4): p. 275-282.
48. Derbyshire, M., et al., *Small RNAs from the plant pathogenic fungus Sclerotinia sclerotiorum highlight host candidate genes associated with quantitative disease resistance*. *Molecular Plant Pathology*, 2019. **20**(9): p. 1279-1297.
49. Yang, F., *Genome-wide analysis of small RNAs in the wheat pathogenic fungus Zymoseptoria tritici*. *Fungal Biology*, 2015. **119**(7): p. 631-640.
50. Wang, B., et al., *Puccinia striiformis f. sp. tritici mi croRNA-like RNA 1 (Pst-milR1), an important pathogenicity factor of Pst, impairs wheat resistance to Pst by suppressing the wheat pathogenesis-related 2 gene*. *New Phytologist*, 2017. **215**(1): p. 338-350.
51. Jian, J. and X. Liang, *One small RNA of Fusarium graminearum targets and silences CEBiP gene in common wheat*. *Microorganisms*, 2019. **7**(10).
52. Jin, X., et al., *Identification of Fusarium graminearum-responsive miRNAs and their targets in wheat by sRNA sequencing and degradome analysis*. *Functional & Integrative Genomics*, 2019: p. 1-11.
53. Kusch, S., et al., *Small RNAs from cereal powdery mildew pathogens may target host plant genes*. *Fungal Biology*, 2018. **122**(11): p. 1050-1063.
54. Drakopoulos, D., et al., *Control of Fusarium graminearum in wheat with mustard-based botanicals: From in vitro to in planta*. *Frontiers in Microbiology*, 2020. **11**: p. 1595.
55. Lee, A.H.-Y., et al., *A bacterial acetyltransferase destroys plant microtubule networks and blocks secretion*. *PLoS Pathogens*, 2012. **8**(2): p. e1002523.
56. Xia, Z., et al., *Characterization of microRNA-like RNAs associated with sclerotial development in Sclerotinia sclerotiorum*. *Fungal Genetics and Biology*, 2020. **144**: p. 103471.

57. Shahid, S., et al., *MicroRNAs from the parasitic plant Cuscuta campestris target host messenger RNAs*. Nature, 2018. **553**(7686): p. 82-85.
58. Huang, J., et al., *Diverse functions of small RNAs in different plant–pathogen communications*. Frontiers in Microbiology, 2016. **7**: p. 1552.
59. Navarro, L., et al., *A plant miRNA contributes to antibacterial resistance by repressing auxin signaling*. Science, 2006. **312**(5772): p. 436-439.
60. Tian, D., et al., *Fitness costs of R-gene-mediated resistance in Arabidopsis thaliana*. Nature, 2003. **423**(6935): p. 74-77.
61. Cui, C., et al., *A Brassica miRNA regulates plant growth and immunity through distinct modes of action*. Molecular Plant, 2020. **13**(2): p. 231-245.
62. Zhang, C., et al., *Identification of trans-acting siRNAs and their regulatory cascades in grapevine*. Bioinformatics, 2012. **28**(20): p. 2561-2568.
63. Wu, F., et al., *Genome-wide identification and characterization of phased small interfering RNA genes in response to Botrytis cinerea infection in Solanum lycopersicum*. Scientific Reports, 2017. **7**(1): p. 1-10.
64. Zhai, J., et al., *MicroRNAs as master regulators of the plant NB-LRR defense gene family via the production of phased, trans-acting siRNAs*. Genes & Development, 2011. **25**(23): p. 2540-2553.
65. Gupta, O.P., et al., *MicroRNA regulated defense responses in Triticum aestivum L. during Puccinia graminis f. sp. tritici infection*. Molecular Biology Reports, 2012. **39**(2): p. 817-824.
66. Regmi, R., et al., *Identification of B. napus small RNAs responsive to infection by a necrotrophic pathogen*. BMC Plant Biology, 2021. **21**(1): p. 1-21.
67. Sánchez-Sanuy, F., et al., *Osa-miR7695 enhances transcriptional priming in defense responses against the rice blast fungus*. BMC Plant Biology, 2019. **19**(1): p. 1-16.

68. Hunt, M., et al., *Small RNA discovery in the interaction between barley and the powdery mildew pathogen*. BMC Genomics, 2019. **20**(1): p. 610.
69. Deng, P., et al., *Biogenesis and regulatory hierarchy of phased small interfering RNAs in plants*. Plant Biotechnology Journal, 2018. **16**(5): p. 965-975.
70. Bian, H., et al., *Distinctive expression patterns and roles of the miRNA393/TIR1 homolog module in regulating flag leaf inclination and primary and crown root growth in rice (*Oryza sativa*)*. New Phytologist, 2012. **196**(1): p. 149-161.
71. Campo, S., et al., *Identification of a novel micro RNA (mi RNA) from rice that targets an alternatively spliced transcript of the N ramp6 (N atural resistance-associated macrophage protein 6) gene involved in pathogen resistance*. New Phytologist, 2013. **199**(1): p. 212-227.
72. Li, Y., et al., *Identification of microRNAs involved in pathogen-associated molecular pattern-triggered plant innate immunity*. Plant Physiology, 2010. **152**(4): p. 2222-2231.
73. Guo, N., et al., *Microarray profiling reveals microRNAs involving soybean resistance to *Phytophthora sojae**. Genome, 2011. **54**(11): p. 954-958.
74. Huang, J., M. Yang, and X. Zhang, *The function of small RNAs in plant biotic stress response*. Journal of Integrative Plant Biology, 2016. **58**(4): p. 312-327.
75. Zhao, Y.-T., et al., *Small RNA Profiling in Two *Brassica napus* Cultivars Identifies MicroRNAs with Oil Production- and Development-Related Expression and New Small RNA Classes*. Plant Physiology, 2012. **158**(2): p. 813-823.
76. Vaisey Genser, M., *Canola oil : properties and performance*. Publication / Canola Council of Canada, 1987, Winnipeg, Man.: Canola Council. 51 p.
77. Wang, Z., et al., *Small RNA and degradome profiling involved in seed development and oil synthesis of *Brassica napus**. PLoS ONE, 2018. **13**(10):p.1-21.

78. Zhao, Y.-T., et al., *Small RNA profiling in two Brassica napus cultivars identifies microRNAs with oil production-and development-correlated expression and new small RNA classes*. Plant Physiology, 2012. **158**(2): p. 813-823.
79. Gill, S.S. and N. Tuteja, *Cadmium stress tolerance in crop plants: probing the role of sulfur*. Plant Signaling & Behavior, 2011. **6**(2): p. 215-222.
80. El Rasafi, T., et al., *Cadmium stress in plants: A critical review of the effects, mechanisms, and tolerance strategies*. Critical Reviews in Environmental Science and Technology, 2020: p. 1-52.
81. Xie, F.L., et al., *Computational identification of novel microRNAs and targets in Brassica napus*. Febs Letters, 2007. **581**(7): p. 1464-1474.
82. Zhou, Z.S., J.B. Song, and Z.M. Yang, *Genome-wide identification of Brassica napus microRNAs and their targets in response to cadmium*. Journal of Experimental Botany, 2012. **63**(12): p. 4597-4613.
83. Huang, S.Q., et al., *A set of miRNAs from Brassica napus in response to sulphate deficiency and cadmium stress*. Plant Biotechnology Journal, 2010. **8**(8): p. 887-899.
84. Fu, Y., et al., *MicroRNA-mRNA expression profiles and their potential role in cadmium stress response in Brassica napus*. BMC plant biology, 2019. **19**(1): p. 1-20.
85. Jian, H., et al., *Genome-wide identification of microRNAs in response to cadmium stress in oilseed rape (Brassica napus L.) using high-throughput sequencing*. International Journal of Molecular Sciences, 2018. **19**(5): p. 1431.
86. Ma, Y., M.C. Dias, and H. Freitas, *Drought and salinity stress responses and microbe-induced tolerance in plants*. Frontiers in Plant Science, 2020. **11**: p. 1750.
87. Jian, H., et al., *Identification of rapeseed microRNAs involved in early stage seed germination under salt and drought stresses*. Frontiers in Plant Science, 2016. **7**: p. 658.

88. Cao, J.Y., et al., *Tight regulation of the interaction between Brassica napus and Sclerotinia sclerotiorum at the microRNA level*. Plant Molecular Biology, 2016. **92**(1-2): p. 39-55.
89. Jian, H., et al., *Integrated mRNA, sRNA, and degradome sequencing reveal oilseed rape complex responses to Sclerotinia sclerotiorum (Lib.) infection*. Scientific Reports, 2018. **8**(1): p. 1-17.
90. Regmi, R., et al., *Identification of B. napus small RNAs responsive to infection by a necrotrophic pathogen*. BMC Plant Biology, 2021. **21**(1).
91. Martin, M., *Cutadapt removes adapter sequences from high-throughput sequencing reads*. EMBnet. Journal, 2011. **17**(1): p. 10-12.
92. Bolger, A. and F. Giorgi, *Trimmomatic: a flexible trimmer for Illumina sequence data*. Bioinformatics, 2014. **30**(15):p. 2114-2120.
93. He, B., et al., *Assessing the Impact of Data Preprocessing on Analyzing Next Generation Sequencing Data*. Frontiers in Bioengineering and Biotechnology, 2020. **8**: p. 817.
94. Axtell, M.J., *ShortStack: comprehensive annotation and quantification of small RNA genes*. Rna, 2013. **19**(6): p. 740-751.
95. Mohorianu, I., et al., *The UEA small RNA workbench: a suite of computational tools for small RNA analysis*, in *MicroRNA Detection and Target Identification*. 2017, Springer. p. 193-224.
96. Allan, J., et al., *The host generalist phytopathogenic fungus Sclerotinia sclerotiorum differentially expresses multiple metabolic enzymes on two different plant hosts*. Scientific Reports, 2019. **9**(1): p. 1-15.
97. Mwape, V.W., et al., *Analysis of differentially expressed Sclerotinia sclerotiorum genes during the interaction with moderately resistant and highly susceptible chickpea lines*. BMC Genomics, 2021. **22**(1): p. 1-14.

98. Bushnell, B.B.A.F., *Accurate, Splice-Aware Aligner*. United States. , *BBMap: A Fast, Accurate, Splice-Aware Aligner*. United States. 2014.
99. Griffiths-Jones, S., et al., *Rfam: annotating non-coding RNAs in complete genomes*. *Nucleic Acids Research*, 2005. **33**(suppl_1): p. D121-D124.
100. Nawrocki, E.P. and S.R. Eddy, *Infernal 1.1: 100-fold faster RNA homology searches*. *Bioinformatics*, 2013. **29**(22): p. 2933-2935.
101. Andrews, S., *FastQC: a quality control tool for high throughput sequence data*. 2010, Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom.
102. Li, R., et al., *SOAP: short oligonucleotide alignment program*. *Bioinformatics*, 2008. **24**(5): p. 713-714.
103. Gomes, C.P.D.C., et al., *A review of computational tools in microRNA discovery*. *Frontiers in Genetics*, 2013. **4**: p. 81.
104. Friedländer, M.R., et al., *miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades*. *Nucleic Acids Research*, 2012. **40**(1): p. 37-52.
105. Hackenberg, M., et al., *miRanalyzer: a microRNA detection and analysis tool for next-generation sequencing experiments*. *Nucleic acids research*, 2009. **37**(suppl_2): p. W68-W76.
106. Oulas, A., et al., *Finding cancer-associated miRNAs: methods and tools*. *Molecular Biotechnology*, 2011. **49**(1): p. 97-107.
107. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2*. *Genome Biology*, 2014. **15**(12): p. 1-21.
108. Robinson, M.D., D.J. McCarthy, and G.K. Smyth, *edgeR: a Bioconductor package for differential expression analysis of digital gene expression data*. *Bioinformatics*, 2010. **26**(1): p. 139-140.
109. Peterson, S.M., et al., *Common features of microRNA target prediction tools*. *Frontiers in Genetics*, 2014. **5**: p. 23.

110. Enright, A., et al., *MicroRNA targets in Drosophila*. *Genome Biology*, 2003. **4**(11): p. 1-27.
111. Lewis, B.P., et al., *Prediction of mammalian microRNA targets*. *Cell*, 2003. **115**(7): p. 787-798.
112. Bandyopadhyay, S. and R. Mitra, *TargetMiner: microRNA target prediction with systematic identification of tissue-specific negative examples*. *Bioinformatics*, 2009. **25**(20): p. 2625-2631.
113. Zhang, Y., *miRU: an automated plant miRNA target prediction server*. *Nucleic Acids Research*, 2005. **33**(suppl_2): p. W701-W704.
114. Dai, X. and P.X. Zhao, *psRNATarget: a plant small RNA target analysis server*. *Nucleic Acids Research*, 2011. **39**(suppl_2): p. W155-W159.
115. John, B., et al., *Human microRNA targets*. *PLoS Biology*, 2004. **2**(11): p. e363.
116. Addo-Quaye, C., et al., *Endogenous siRNA and miRNA targets identified by sequencing of the Arabidopsis degradome*. *Current Biology*, 2008. **18**(10): p. 758-762.
117. Addo-Quaye, C., W. Miller, and M.J. Axtell, *CleaveLand: a pipeline for using degradome data to find cleaved small RNA targets*. *Bioinformatics*, 2008. **25**(1): p. 130-131.
118. Thody, J., et al., *PAREsnip2: a tool for high-throughput prediction of small RNA targets from degradome sequencing data using configurable targeting rules*. *Nucleic Acids Research*, 2018. **46**(17): p. 8730-8739.
119. Rice, P., I. Longden, and A. Bleasby, *EMBOSS: the European molecular biology open software suite*. *Trends in Genetics*, 2000. **16**(6): p. 276-277.
120. Edris, B., *A comparison of the Oligomap and TargetScan algorithms for miRNA target analysis*. Bmi231. Stanford. Edu, 2011.

Chapter 2: Transformation of *Sclerotinia sclerotiorum* with a gene encoding green fluorescent protein

Abstract

The green fluorescent protein (GFP) is a genetic marker that facilitates the study of host and pathogen interaction when a fungus is stably transformed with this protein. When viewed under a microscope, GFP glows green under the excitation in the blue light spectrum. Thus, GFP transformed pathogen strains can be studied by microscopy for pathogen and host interaction.

S. sclerotiorum strain CU8.24 was transformed with a gene encoding GFP using polyethylene glycol (PEG)-mediated transformation of *S. sclerotiorum* protoplasts. Three stable transformants were identified based on their growth morphology *in vitro*. These putative transformants were validated using polymerase chain reaction (PCR) and microscopy tests. A plant infection assay was conducted in a susceptible *B. napus* variety to compare the pathogenicity of transformants with the wild type of strain. No significant difference was observed in fungal growth among wild type and mutant strains *in vitro* or in disease progression in *B. napus* plants, justifying the use of these strains for downstream microscopy studies. Thus, the developed GFP-expressing Australian *S. sclerotiorum* strain can be used for downstream applications to better understand *S. sclerotiorum* pathogenesis.

2.1 Introduction

Green fluorescent protein (GFP) is a reporter molecule which helps to observe the disease progression of a plant pathogen, when it is tagged with this fluorophore. Therefore, GFP tagged fungal pathogens can be used for the characterization of host and pathogen interaction. Traditionally, host and pathogen interactions were studied using techniques like staining, which are often time-consuming and also do not provide real time images of infection of plant tissue. Also, the use of traditional transgenes like *uidA*, the glucuronidase (*GUS*) gene, requires the destruction of tissues due to the application of substrates used for the leakiness of product and cellular permeabilization of substrates [1]. Different forms of GFP have been used as a reporter molecule in many organisms [2]. A number of fungal pathogens have been transformed with the *GFP* gene allowing researchers to investigate host and pathogen interaction such as in *Ustilago maydis* [3], *Cochliobolus heterostrophus* [4], *Aspergillus nidulans* [5], *Phytophthora parasitica* [6], *Botrytis cinerea* [7], *Leptosphaeria maculans* [2] and *Ascochyta lentis* [8]. Lorang et al. [1] transformed eight ascomycete fungal pathogens including a European *S. sclerotiorum* isolate with the *GFP* gene. Later, researchers used this GFP-transformed isolate to investigate the effect of oxalate production on regulating stomatal guard cells in beans [9]. Moreover, a GFP-transformant of an American isolate of *S.*

sclerotiorum has been developed and its pathogenicity in four crops, including canola, soybean, bean, and sunflower were investigated [10]. Fungal genetic transformation techniques enable scientists to understand the functions of targeted genes, or modify the expression of endogenous genes such as by insertion of an alternative promoter [11]. Since the first transformation of fungi in 1983 [12] hundreds of studies have reported transformation procedures [11]. Protoplast-mediated transformation [13] and the *Agrobacterium*-mediated transformation [14] methods are major methods for fungal transformation adopted by most researchers worldwide. In addition, there are other alternative approaches: electroporation [15] and the biolistic method [16]. Recently, CRISPR-Cas9 has also emerged as a reliable gene knock-out system [17]. In this thesis, we used protoplast-mediated targeted transformation for targeted insertion of DNA into the *S. sclerotiorum* genome. Targeted insertion or homologous recombination helps for precise DNA integration at a desired genetic location. Both circular and linearized plasmid have been used for homologous recombination. It has been shown previously in *S. cerevisiae* [18] and *Alternaria alternata* [19] that the use of linearized plasmid enhanced transformation efficiency.

S. sclerotiorum is a broad host range pathogen which causes an important disease called sclerotinia stem rot (SSR) in important crops like canola, bean, peas, soybean, tomato, and sunflower [20]. Canola is a common name for *Brassica napus* and is the largest oil-seed crop produced in Australia. In Western Australia, almost 1.4 million tonnes of canola were produced contributing 812 million to the state revenue in 2018-2019 (DPIRD). The major control for this disease is the use of fungicides. Therefore, it is important to understand the interactions of *S. sclerotiorum* with its hosts, where an Australian GFP-transformed *S. sclerotiorum* isolate may help in better understanding its interaction with its host crops.

Here, we transformed an Australian isolate of *S. sclerotiorum* with a gene encoding GFP flanked by specific *S. sclerotiorum* genome segment sequences. Three independent *S. sclerotiorum*-GFP transformants were checked for their ability to maintain similar mycelium morphology and infection in comparison to the wild type strain. The GFP-expressing *S. sclerotiorum* strains developed in this chapter will be useful to study disease progression during infection in different *S. sclerotiorum* hosts. The techniques used and optimized in this chapter for the transformation of *S. sclerotiorum* were used for further transformation work in Chapter

6.

2.2 Methods

2.2.1 Transformation vector

Plasmid ‘pUC57.Ss1186-GFP’ was obtained from LIPM, INRA, CNRS, Université de Toulouse, Castanet Tolosan, France. The plasmid contains a *B. cinerea* codon-optimized *GFP* gene under the *niaD* promoter flanked by specific *S. sclerotiorum* genome segments (Ss) (1,086 bp left and 900 bp right) and was used as a transformation vector. The vector contains the ‘spectinomycin (SpecR)’ gene for bacterial selection and ‘Hygromycin (Hyg)’ gene for fungal selection (Fig. 2.1).

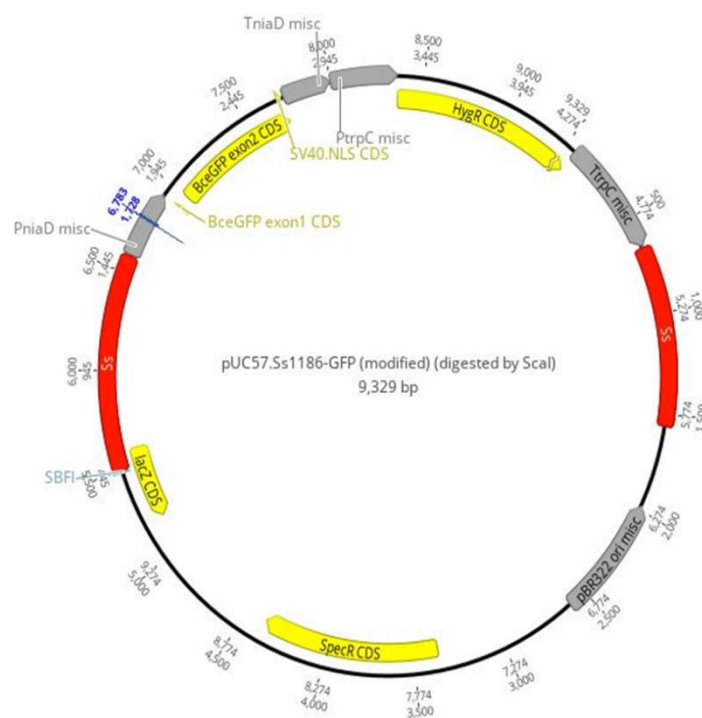


Fig. 2. 1. “puc57.Ss1186-GFP” plasmid flanked by *Sclerotinia sclerotiorum* sequences (red colour) for target insertion of the GFP gene.

2.2.2 Plasmid amplification for *S. sclerotiorum* transformation

Plasmid pUC57 was transformed into *E. coli* strain DH5-alpha using heat shock and incubated at 37 °C overnight for 18 hours. Positive colonies were selected on Lysogeny Broth (LB) agar plates using spectinomycin antibiotic (100µg/mL). Colonies grown in the LB spectinomycin agar plates were subjected to colony PCR to amplify GFP and *S. sclerotiorum* flanking sequences. The *GFP* gene was amplified using forward primer 5'-ACTGGCGTCGTTCCAATCTT-3' and reverse primer 5'-GCAAGACTGGACCATCACCA-3' to yield a 559 bp product. The left flanking sequence of *S. sclerotiorum* (SSA) was amplified using 5'-TGCAGCTGTGATACTCCGTAG -3' and 5'-ACATCAGCCAGTTCAAGCCA-3' forward and reverse primers, respectively. Similarly, the right flanking sequence (SSB) was amplified using forward primer 5'-AAATTGAGCAGACAGGGGCT-3' and reverse primer 5'-AGTCAGTCTGGCCTCAACTC -3'. Both of the flanking sequences gave a 500 bp product (Fig. 2.2). The conditions for the polymerase chain reaction (PCR) were denaturation at 95 °C for 3 min, denaturation at 95 °C for 15 sec, annealing at 55 °C for 20 sec, extension at 72 °C for 1 min 30 sec, and final amplification of 72 °C for 3 min. The PCR program was run for 35 cycles.

After the successful amplification of *GFP* and *S. sclerotiorum* flanking sequences, positive colonies were grown overnight in LB supplemented with spectinomycin antibiotic (100µg/mL) and purified using Promega Miniprep Kit following the manufacturer's protocol. Around 50 micrograms of GFP plasmid was harvested from the multiple miniprep. According to the protocol, 20 micrograms of the plasmid was used for transformation of *S. sclerotiorum*.

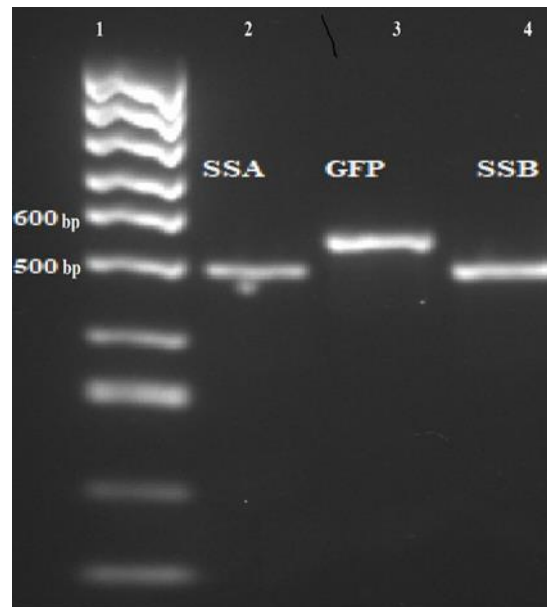


Fig. 2.2. Colony PCR showing amplification of green fluorescent protein (*GFP*) gene and two *Sclerotinia sclerotiorum* flanking sequences *SSA* (left) and *SSB* (right) in the pUC57 plasmid. *SSA* spans outside the promoter region while *SSB* presents outside the terminator region. Lane 1 is 1Kb Bioline hyperladder; lane 2 amplification of left flanking sequence; lane 3 amplification of *GFP* and lane 4 amplification of right flanking sequence.

2.2.3 Maintenance and culture of *Sclerotinia sclerotiorum*

S. sclerotiorum strain CU8.24 originally collected from South Stirling, WA [21] was used to generate a GFP strain. A mature sclerotium was cut into two halves and each half was put into the middle of a 9 cm potato dextrose agar (PDA) Petri dish with the cut end placed down onto the PDA medium [21]. Mycelium that had germinated from the sclerotium was subcultured onto fresh PDA five days post inoculation. After 48 hours, four circular mycelial plugs were taken from the sub-cultured mycelium and added to 100 mL of potato dextrose broth (PDB) and incubated at room temperature shaking at 100 revolutions per minute (RPM) for three days.

2.2.4 PEG-mediated protoplast transformation of *Sclerotinia sclerotiorum* with GFP

A protoplast transformation [22] strategy was adopted for the transformation of the strain CU8.24 with the GFP-containing plasmid “pUC57”. Two mycelial plugs grown in potato dextrose broth (PDB) for three days were recovered to produce protoplasts. Protoplasts were made by digesting mycelium using enzymatic digestion solution containing 0.8 M mannitol, 200 mM citric acid/tri-sodium citrate buffer, and pH 5.5 and 1.5 percent lysing enzyme from *Trichoderma harizanum* (Thermofischer Scientific) with three hours incubation at room

temperature and gentle shaking at 80 RPM. Protoplasts were sequentially filtered through 100 μm and 40 μm cell filtration units (Sigma-Aldrich). The protoplasts were collected by centrifugation at 800 x g for 5 min at 4 °C. Pelleted protoplasts were re-suspended in enzymatic digestion buffer and washed twice with cold STC buffer which constituted 0.8 M sorbitol, 50 mM Tris-HCl, 50 mM CaCl₂. Washed protoplasts were recovered in STC buffer and pelleted again at 800 x g and re-suspended in 200 μL of STC: PTC buffer. (STC: PTC contains 40% PEG in STC buffer. Plasmid DNA (20 μg) and heparin (5 μL of 5 mg/mL stock in STC buffer) were added to protoplast aliquots and incubated on ice for one hour. After incubation, 1 mL of PTC buffer was added to the mixture. This was then mixed gently by inversion and incubated at room temperature for 20 min. The mixture containing protoplasts was poured into liquid regeneration medium (0.1% yeast extract, 0.1% casein hydrolysates, 1 M sucrose, 1.6 % granulated agar). The regeneration medium was then poured into Petri dishes and protoplasts were left to regenerate overnight at 20 °C. After overnight incubation of the protoplasts, selection medium amended with 100 $\mu\text{g}/\text{mL}$ of hygromycin B was poured over the solid regeneration plates and incubated until the growth of fungal colonies on the surface of the regeneration medium was observed.

2.2.5 Transformant selection

Germination of mycelium from putative transformants was observed in the original transformant plates after 10 days. Twelve putative transformants were obtained from the original six transformation plates and subsequently transferred to hygromycin B (100 µg/mL) supplemented PDA plates for testing transgene stability. After two subsequent rounds of selection on PDA plates supplemented with hygromycin B, a mycelial plug from each plate was subcultured multiple times on fresh PDA to enrich GFP-transformed nuclei given the multinucleate nature of *S. sclerotiorum* [10]. Altogether, six subcultures were subcultured in PDA plates supplemented with hygromycin B to enhance the purity of the GFP transformed strains. After the third subculture until the sixth subculture, mycelium was observed in a Nikon confocal A1 plus microscope. Three putative transformants were selected further based upon their intensity of fluorescence under the microscope. These transformants were given the names M1, M2 and M3 and used for *in vitro* and *in planta* validation. Sclerotia were harvested after the sixth hyphal tip regeneration, which were used as the starting material for downstream studies such as *in vitro* and *in planta* assays.

2.2.6 DNA extraction

DNA was extracted from the wild type strain and three putative GFP-transformed *S. sclerotiorum* strains using JW Taylor lysis buffer [23]. Sclerotia of M1, M2, M3 and the wild type strain were cultured on PDA plates. After mycelial germination, a mycelium plug was subcultured on fresh PDA plates. Once mycelium sufficiently grew to cover 9 cm of the PDA plate, a 1 cm mycelial plug was inoculated into 100 mL of potato dextrose broth and incubated for three days with shaking at 80 rpm at room temperature. After three days, one gram of wet mycelium was crushed in liquid nitrogen. Cells were lysed with 750 μ L of lysis buffer and incubated at 65 °C. Thereafter, 700 μ L of chloroform: phenol (1:1) was added and vortexed briefly and centrifuged at 12,000 x g for 10 minutes to separate the organic and aqueous phases. After phase separation, DNA was precipitated with isopropanol. The precipitated DNA was washed with ethanol and eluted into clean molecular grade water. The quantity of extracted DNA was quantified using a NanoDrop spectrophotometer and 100 ng of the DNA template was used for polymerase chain reaction (PCR).

2.2.7 Confirmation of insertion of *GFP* in transformants by PCR

To validate whether the *GFP* sequence was inserted into the *S. sclerotiorum* genome, different sets of primers were used which included primers that bind to GFP, Hyg, and the two flanks (F1 and F2). A schematic of primer positions for the amplification of flanks is shown in Fig

2.3. The left flank (F1) was amplified with the forward primer F1F (5'-AGCAGTGGAAGTTGGAAGGT-3') and reverse primer F1R (5'-GGCGGGATCTGACGAAATGA-3') and the right flank (F2) was amplified with a forward primer F2F (5'-CTCAAGCCTACAGGACACACA-3') and reverse primer F2R (5'-GGGGTAAGAGGAAGAAATGATTGG-3'). Wild type *S. sclerotiorum* genomic DNA was used as a negative control and a no template control was included with water. The PCR reaction was prepared with RedTaq ready mix (Sigma-Aldrich) and run under the conditions: denaturation at 95 °C for 3 min for 1 cycle, 95 °C for 15 sec, annealing at 55 °C for 20 sec, and 72 °C for 1 min 30 sec for 35 cycles, and final amplification of 72 °C for 3 min for 1 cycle.

2.2.8 Measurement of mycelium growth *in vitro*

Three replicates of each transgenic and wild type strain were studied by subculturing them in fresh PDA to determine their growth rate and any morphological differences. The diameter of

mycelial growth was measured at 6,12,24 and 48 H. Three replicates were used at each time point for all the strains. An assessment for significant difference between lesion length measurement at each time point was carried using a one-way ANOVA at a P -value < 0.05 and Tukey's Honest Significant Difference (HSD) post hoc test in R studio version 3.6.1. Following the generation of three independent GFP-transformed strains, their growth was compared to that of the wildtype strain CU8.24.

2.2.9 Pathogenicity comparison of wild type and GFP-transgenic strains

GFP-transformed and wild type sclerotia were cut into halves and cultured on PDA. After 48 hpi, a 0.5 cm mycelial plug was used to inoculate *B. napus* plants with 30-50 percent of flowers open on the main raceme. Seeds of AV-Garnet were acquired from The Australian Grains GeneBank (accession AGG95718BRAS1). The fifth internode from the base was chosen for the point of infection. The mycelial plug was wrapped in parafilm to maintain contact with the stem and to enhance humidity around the infection point. Four replicates were used for each strain. Plants were grown in a growth chamber with a photoperiod of 12 h of light and 12 h of darkness at 22 °C and humidity was maintained at 40 percent using a humidifier. Lesion length was measured for three consecutive weeks at 7, 14 and 21 days post inoculation (DPI) .

Significant differences between lesion length measurements at each timepoint were determined using a one-way ANOVA followed by a Tukey's HSD post hoc test at P -value 0.05.

2.3 Results

2.3.1 Putative transgenic strains were confirmed for the insertion of *GFP* by PCR

To confirm the insertion of *GFP* by PCR, different combinations of primers were used as shown in the schematic in Fig. 2.3A. Fig. 2.3 B shows the amplification of *GFP*, Hyg, F1, and F2 in three transgenic strains and no amplification in the wild type strain, suggesting successful insertion of the *GFP* gene into the genome of isolate CU8.24. The product sizes for *GFP*, Hyg, F1, and F2 were 559 bp, 843 bp, 1,471 bp, and 1,500 bp, respectively.

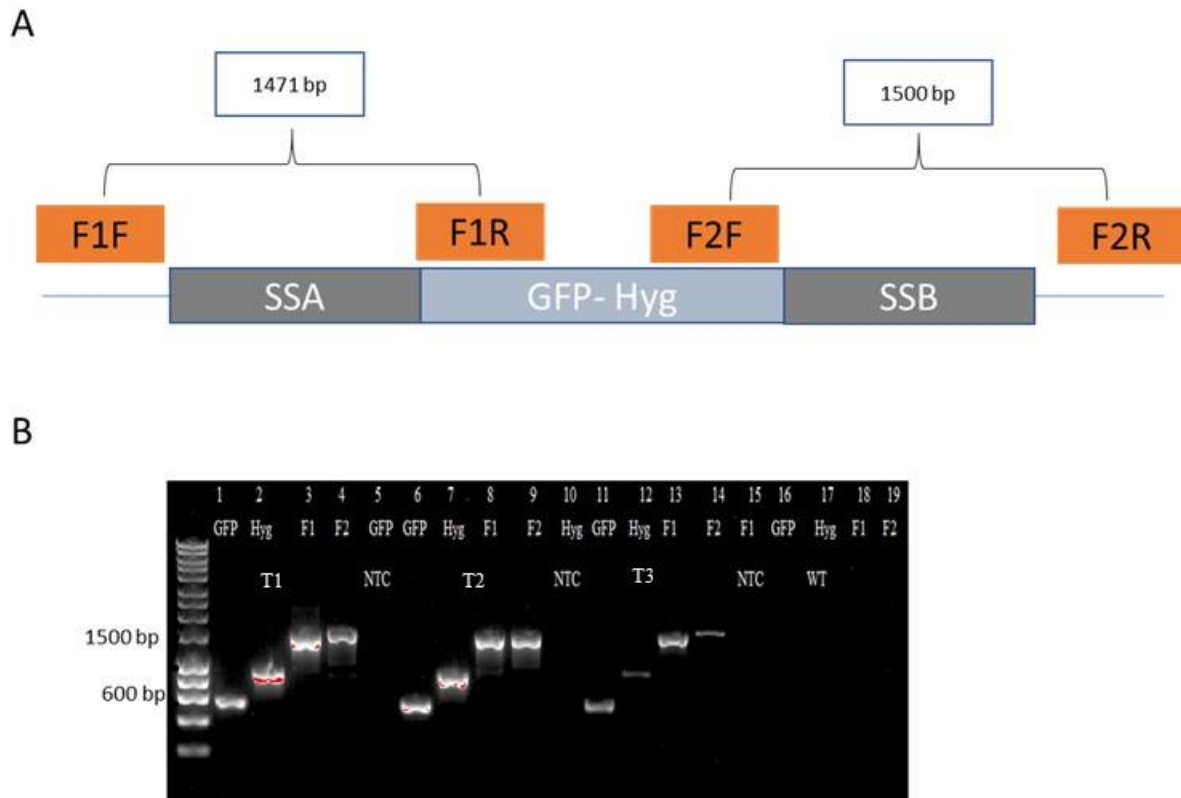


Fig. 2.3. The confirmation of GFP insertion in *Sclerotinia sclerotiorum* genome with the use of different set of primers. (A) The schematic of primer position for the amplification of left and right flanks. (B). The electrophoresis gel images of the amplicon produce for wild type (WT) and three putative transgenic strains (T1, T2 and T3) with GFP, Hyg, F1 and F2 primers. NTC is a no template control.

2.3.2 Fluorescence microscopy identifies three stable GFP expressing strains of *Sclerotinia sclerotiorum* isolate CU8.24

S. sclerotiorum has multiple nuclei per cell. However, stable integration of the GFP construct only has to occur in a single nucleus for the fungus to grow on selectable media. Therefore, hyphal tips of putative transformants were transferred onto selectable medium for six generations. Mycelia from putative transformants and the wild type strain were mounted in 40

% glycerol and observed under a Nikon A1 plus confocal microscope at an excitation wavelength of 488.6 nm and an emission wavelength of 525.0 nm, which is optimized for GFP. No strong fluorescence was observed in the initial three subcultures of the transformants suggesting not all the nuclei contained the *GFP* gene. The intensity of fluorescence exhibited bright green fluorescent colour at the sixth subculture (Fig. 2.4). The reason might be due to the enrichment of the GFP-transformed nuclei after multiple hyphal tip regeneration steps. However, sectoring in fluorescence intensity was observed within the stable transformants. Some hyphae exhibited strong fluorescence while some displayed a weak signal; a similar phenomenon was observed earlier with *S. sclerotiorum* transformed strains [10]. Interestingly, a mycelium sample taken from potato broth culture has enhanced fluorescence compared to that of the PDA (data not shown).

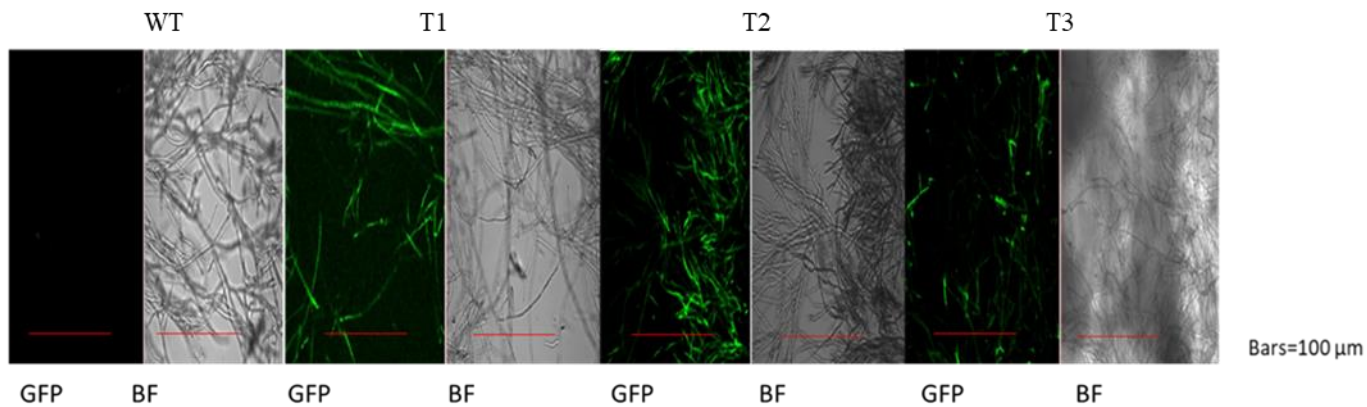


Fig. 2. 4. Confocal microscopy images after six hyphal tip regenerations of wild type and green fluorescent protein (GFP) transformed strains of *S. sclerotiorum* isolate CU 8.24. WT, T1, T2, and T3 are wild type, transgenic strain 1, transgenic strain 2, and transgenic strain 3, respectively. For each individual the left image (GFP) shows the GFP expression following excitation wavelength of 488.6 nm and an emission wavelength of 525.0 nm and the right image the bright field (BF).

2.3.3 Transgenic and wild type strains maintain the same growth rate *in vitro*

To assess whether the transformants maintain the same mycelium growth compared to wild type strains, we measured the mycelial diameter of growth on PDA across four different time points. The mean hyphal diameter for the GFP-transformed strains and wild type at 6 H was 0.18 cm and 0.13 cm respectively. Similarly, mean hyphal growth at 12 H, 24 H, and 48 H for transgenic strains were 0.64 cm, 1.52 cm, and 3.7 cm, respectively. For wild type strains the hyphal diameter recorded at 12 H, 24 H and 48 H were 0.5 cm, 1.5 cm and 3.6 cm, respectively. (Fig 2.5). There were no significant differences (Tukey's HSD post hoc test) at any time points

between the GFP-transformed and wild type strains suggesting *S. sclerotiorum* transformed with GFP maintains the same growth rate *in vitro* as the wild type.

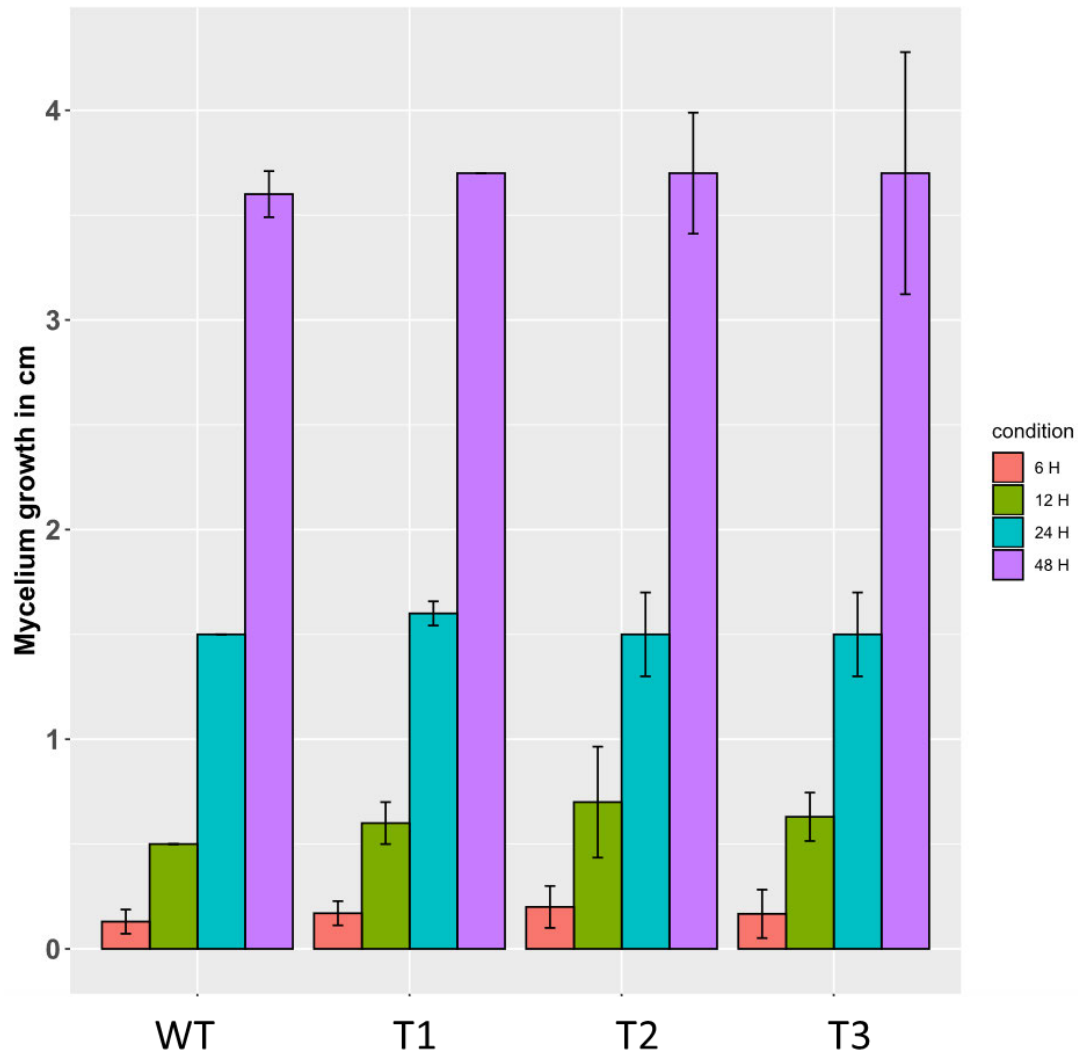


Fig. 2. 5. Comparison of mycelium growth measured at four different time points between wild type and *Sclerotinia sclerotiorum*-GFP strains. No significant differences were observed between the wild type (WT) and the three GFP-transformed strains (T1, T2 and T3) by Tukey HSD-test ($p < 0.05$).

2.3.4 Infection of plants with GFP-expressing strains of *Sclerotinia sclerotiorum* shows no difference to wild type pathogenicity

After comparing the growth rate of the GFP-transformed and wild type strains on plates, susceptible *B. napus* plants (cv AV Garnet) were infected with these strains to investigate their ability to infect canola plants. All transformed strains were pathogenic and initial lesion formation was clearly visible in the stems of canola after a week of infection and increased with time. Fig 2.6 shows the representative figures of disease symptoms on stems of canola by transformed and wild type strains at 14 DPI. Lesion length caused by all GFP strains were similar with no significant difference to those caused by the wild type. T2 exhibited relatively higher lesion size than T1 and T3 strains across all time points. However, no significant difference was observed compared to wild type strains. The mean values of lesion length at 7, 14 and 21 DPI for transformed strains were 4.43, 7.85 and 12.4 cm respectively. For wild type strains the average lesion size at 7, 14, and 21 DPI were 3.98, 9.6, and 15.6 cm respectively (Fig. 2.7). There were no significant differences among the different strains at all-time points, suggesting GFP-transformed strains maintain the same level of pathogenicity as the wild type strain. Therefore, these transformed *S. sclerotiorum* strains can be used for downstream applications like microscopy study of *S. sclerotiorum* growth *in planta*.

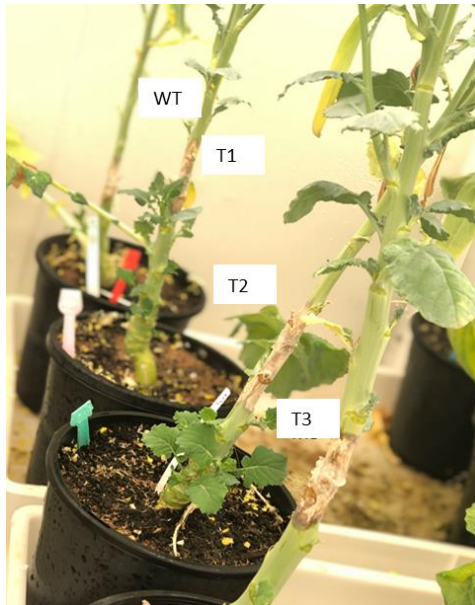


Fig. 2. 6. Lesion development on stems of three GFP-transformed *S. sclerotiorum* strains (T1, T2 and T3) and one wild type (WT) strain after two weeks placed in a growth room at 16 H day and 8 H night at 24°C.

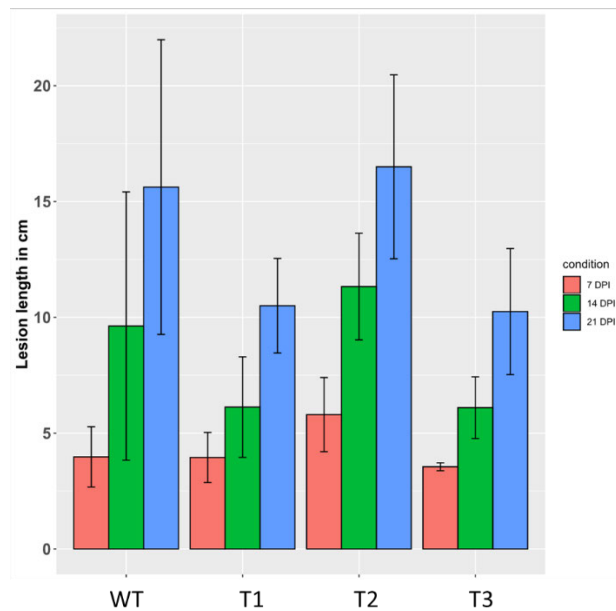


Fig. 2. 7. Comparison of lesion growth on *Brassica napus* plant measured at three different time points between wild type and *Sclerotinia sclerotiorum*-GFP strains. No significant differences were observed between the wild type (WT) and the three GFP-transformed stains (T1, T2 and T3) by Tukey HSD-test ($p < 0.05$).

2.4 Discussion

GFP is a useful genetic marker for tracking fungal disease progression in the host [2]. Various pathogens have been transformed with GFP protein for detailed infection assays [2, 10, 24]. In this study, an aggressive isolate of *S. sclerotiorum* was transformed with GFP using PEG-mediated transformation of protoplasts. The morphology and pathogenicity of transformed strains were assessed in culture plates and by *in planta* infection, respectively. Further validation of the transformants was obtained by PCR and microscopy.

One of the functions of pathogen transformed GFP strains is to track their growth in host tissues [25]. It is important that GFP transformed strains should maintain the same characteristics as wild type *in vitro* and *in planta*. De Silva et al [10] transformed two isolates of *S. sclerotiorum* (originally collected from sunflower and dry bean) with GFP. Subsequently, four transformed strains, two belonging to each isolate, were used to examine the colonization of host tissues of four crops which were canola, soybean, sunflower, and beans. All GFP transformants were pathogenic to their hosts, some of the tested transformed strains of the sunflower collected isolate produced shorter lesion lengths compared to the wild type strain in some hosts while the transformed strains of dry bean collected *S. sclerotiorum* maintained the same lesion size

as the wild type strain [10]. This suggests that the ability of transformed GFP strains to maintain the same pathogenicity as the wild type depends on the isolate. In our study, we tested three GFP-transformed strains from a single isolate CU8.24, which was originally collected from infected canola. The three GFP-strains generated herein produced similar lesion lengths.

Stable transformants varied in the intensity of GFP production. There were two phenomena being observed. At first, the intensity of fluorescence increased with subsequent hyphal tip regeneration and within the stable transformants intensity of GFP fluorescence differed in different hyphae of the same strain. In *B. cinerea*, four subsequent subcultures were done to achieve bright fluorescence while for *S. sclerotiorum* only three were needed [10, 24].

This might be due to the multinucleate nature of the pathogen where subsequent subcultures enriched the GFP transformed nuclei [24]. In a study by De Silva and colleagues in *S. sclerotiorum* [10], some hyphae displayed strong fluorescence while some showed weak fluorescence which is in agreement with this study. During transformation, the GFP construct was most likely integrated in one or a few nuclei of the *S. sclerotiorum* genome. The formation of new hyphal cells due to mitosis is accompanied by the migration of nuclei into the newly

formed hyphal cells; these new cells could receive a higher proportion of transformed nuclei and the fluorescence intensity would be expected to increase.

For effective management of SSR, it is imperative to understand the interactions between *S. sclerotiorum* and *B. napus* under different conditions. Previous research has suggested that the growth of *S. sclerotiorum* is influenced by certain *B. napus* tissues, namely the cortex [26], therefore investigation of tissue-specific resistance would aid in understanding of the *S. sclerotiorum* infection pathway in *B. napus*. Tissue-specific resistance of different lines can be studied at the microscopic and molecular level. Using microscopy, the growth of a pathogen can be traced in the tissues of its host and the rate of growth can be quantified [24]. Uloth et al. [26] provided some insight on tissue-specific resistance while studying various anatomical features associated with the severity of SSR disease using microscopy. It was suggested that canola plants having a greater number of cortical cell layers were more resistant to the disease than those with fewer cortical cell layers. It has been suggested previously that quantitative resistance to *L. maculans* in *B. napus* is also localized to the stem cortex (Huang et al., 2014). Characterization of fungal infection at the macroscopic and microscopic level *in planta* can be facilitated by generating *S. sclerotiorum* strains expressing GFP [27]. Moreover, the nature of

infection between susceptible and resistant varieties can also be unveiled with the help of this strain. Therefore, GFP strains developed in this chapter will be useful for detailed infection assays and microscopic studies for investigating the virulence activities of *S. sclerotiorum* isolate CU8.24 in different host species.

2.5 References

1. Lorang, J.M., et al., *Green Fluorescent Protein Is Lighting Up Fungal Biology*. Applied Environment and Molecular Biology, 2005. **44**. p.1987-1994.
2. Sexton, A. and B. Howlett, *Green fluorescent protein as a reporter in the Brassica–Leptosphaeria maculans interaction*. Physiological and Molecular Plant Pathology, 2001. **58**(1): p. 13-21.
3. Spellig, T., A. Bottin, and R. Kahmann, *Green fluorescent protein (GFP) as a new vital marker in the phytopathogenic fungus Ustilago maydis*. Molecular and General Genetics, 1996. **252**(5): p. 503-509.
4. Maor, R., et al., *Use of green fluorescent protein (GFP) for studying development and fungal-plant interaction in Cochliobolus heterostrophus*. Mycological Research, 1998. **102**(4): p. 491-496.
5. Fernández-Ábalos, J.M., et al., *Plant-adapted green fluorescent protein is a versatile vital reporter for gene expression, protein localization and mitosis in the filamentous fungus, Aspergillus nidulans*. Molecular Microbiology, 1998. **27**(1): p. 121-130.

6. Bottin, A., et al., *Green fluorescent protein (GFP) as gene expression reporter and vital marker for studying development and microbe-plant interaction in the tobacco pathogen Phytophthora parasitica var. nicotianae*. FEMS Microbiology Letters, 1999. **176**(1): p. 51-56.
7. Leroch, M., et al., *Living colors in the gray mold pathogen Botrytis cinerea: Codon-optimized genes encoding green fluorescent protein and mCherry, which exhibit bright fluorescence*. Applied and Environmental Microbiology, 2011. 77(9):p.2887-2897.
8. Henares, B.M., et al., *Agrobacterium tumefaciens-mediated transformation and expression of GFP in Ascochyta lentis to characterize ascochyta blight disease progression in lentil*, PloS One, 2019. **14**(10): p. e0223419.
9. Guimaraes, R.L. and H.U. Stotz, *Oxalate production by Sclerotinia sclerotiorum deregulates guard cells during infection*, Plant Physiology, 2004. **136**(3): p. 3703-3711.
10. De Silva, A., M. Bolton, and B. Nelson, *Transformation of Sclerotinia sclerotiorum with the green fluorescent protein gene and fluorescence of hyphae in four inoculated hosts*, Plant Pathology, 2009. **58**(3): p. 487-496.
11. Martín, J.F., *Fungal transformation: from protoplasts to targeted recombination systems*, in *Genetic Transformation Systems in Fungi*. Springer. **1**: p. 3-18.
12. Buxton, F. and A. Radford, *Cloning of the structural gene for orotidine 5'-phosphate carboxylase of Neurospora crassa by expression in Escherichia coli*. Molecular and General Genetics, 1983. **190**(3): p. 403-405.
13. Turgeon, B.G., et al., *Protoplast transformation of filamentous fungi*, in *Molecular and cell biology methods for fungi*, 2010. Methods in Molecular Biology, 2010. p. 3-19.
14. Michielse, C.B., et al., *Agrobacterium-mediated transformation as a tool for functional genomics in fungi*. Current Genetics, 2005. **48**(1): p. 1-17.

15. Lakrod, K., C. Chairisook, and D. Skinner, *Expression of pigmentation genes following electroporation of albino Monascus purpureus*. Journal of Industrial Microbiology and Biotechnology, 2003. **30**(6): p. 369-374.
16. Armaleo, D., et al., *Biolistic nuclear transformation of Saccharomyces cerevisiae and other fungi*. Current Genetics, 1990. **17**(2): p. 97-103.
17. Doudna, J.A. and E. Charpentier, *The new frontier of genome engineering with CRISPR-Cas9*. Science, 2014. **346**(6213).
18. Orr-Weaver, T.L., J.W. Szostak, and R.J. Rothstein, *Yeast transformation: a model system for the study of recombination*. Proceedings of the National Academy of Sciences, 1981. **78**(10): p. 6354-6358.
19. Shiotani, H. and T. Tsuge, *Efficient gene targeting in the filamentous fungus Alternaria alternata*. Molecular and General Genetics, 1995. **248**(2): p. 142-150.
20. Bolton, M.D., B.P.H.J. Thomma, and B.D. Nelson, *Sclerotinia sclerotiorum (Lib.) de Bary: Biology and molecular traits of a cosmopolitan pathogen*. Molecular Plant Pathology, 2006. **7**(1): p. 1-16.
21. Denton-Giles, M., et al., *Partial stem resistance in Brassica napus to highly aggressive and genetically diverse Sclerotinia sclerotiorum isolates from Australia*. Canadian Journal of Plant Pathology, 2018: 40(4) p. 1-11.
22. Ruiz-Díez, B., *Strategies for the transformation of filamentous fungi*. Journal of applied microbiology, 2002. **92**(2): p. 189-195.
23. Lee, S., M. Milgroom, and J. Taylor, *A rapid, high yield mini-prep method for isolation of total genomic DNA from fungi*. Fungal Genetics Reports, 1988. **35**(1): p. 23.
24. Li, X., T. Zhou, and H. Yu, *Transformation of Botrytis cinerea with a green fluorescent protein (GFP) gene for the study of host-pathogen interactions*. Journal of Plant Pathology, 2006. **6**(2): p. 134-140.

25. Spellig, T., A. Bottin, and R. Kahmann, *Green fluorescent protein (GFP) as a new vital marker in the phytopathogenic fungus Ustilago maydis*. *Molecular and General Genetics*, 1996. 252:p 503-509
26. Uloth, M.B., et al., *Attack modes and defence reactions in pathosystems involving Sclerotinia sclerotiorum, Brassica carinata, B. juncea and B. napus*. *Annals of Botany*, 2016. **117**(1): p. 79-95.
27. Harper, B.K., et al., *Green fluorescent protein as a marker for expression of a second gene in transgenic plants*. *Nature Biotechnology*, 1999. **17**(11): p. 1125-1129.

Chapter 3: Genome wide identification of *Sclerotinia sclerotiorum* small RNAs and their endogenous targets

Abstract

Several phytopathogens produce small non-coding RNAs of approximately 18-30 nt which post-transcriptionally regulate gene expression. Commonly called small RNAs (sRNAs), these small molecules were also reported to be present in the necrotrophic pathogen *Sclerotinia sclerotiorum*. *S. sclerotiorum* causes diseases in more than 400 plant species, including the important oilseed crop *Brassica napus*. sRNAs can further be classified as microRNAs (miRNAs) and short interfering RNAs (siRNAs). Certain miRNAs can activate loci that produce further sRNAs; these secondary sRNA-producing loci are called ‘phased siRNA’ (PHAS) loci, and have only been described in plants. To date, very few studies have characterized sRNAs and their endogenous targets in *S. sclerotiorum*. Here, we used Illumina sequencing to characterize sRNAs from fungal mycelial mats of *S. sclerotiorum* spread over *B. napus* leaves. In total, eight sRNA libraries were prepared from *in vitro*, 12 hours post-inoculation (HPI), and 24 HPI mycelial mat samples. Cluster analysis identified 354 abundant sRNA loci with reads of more than 100 Reads Per Million (RPM). Differential expression analysis revealed upregulation of 34 and 57 loci at 12 and 24 HPI, respectively, in comparison to *in vitro* samples. Among these, 25 loci were commonly upregulated. Altogether, 343 endogenous targets were identified from the major RNAs of 25 loci. Almost 88 % of these targets were annotated as repeat element genes, while the remaining targets were non-repeat element genes. Fungal degradome reads confirmed cleavage of two transposable elements by one upregulated sRNA. Altogether, 23 miRNA loci with ≥ 20 non-redundant sequences were predicted with both mature and miRNA* (star) sequences; these are both criteria associated previously with experimentally verified miRNAs. Degradome sequencing data confirmed the cleavage of 14 targets. These targets were related to repeat element genes, phosphate

acetyltransferases, RNA-binding factor, and exchange factor. A PHAS gene prediction tool identified 26 possible phased interfering loci with 147 phasiRNAs from the *S. sclerotiorum* genome, suggesting this pathogen might produce sRNAs that function similarly to miRNAs in higher eukaryotes. Our results provide new insights into sRNA populations and add a new resource for the study of sRNAs in *S. sclerotiorum*.

3.1 Introduction

There is a complex interaction between pathogens and their hosts during infection [1]. These interactions can be studied with the patterns of gene expression and regulation obtained through RNA sequencing of both host and pathogen [2-5]. Although RNA sequencing helps to identify the key protein-coding genes involved in disease development, the complete picture of gene regulation also demands the investigation of sRNAs, which are 20-30 nucleotide non-coding RNA sequences that regulate gene expression in various biological processes including development and growth, maintenance of genome integrity, and responses to biotic and abiotic stress [6-8]. sRNAs regulate their target transcripts via sequence complementarity mediated by a process called RNA interference (RNAi) [9]. sRNAs are produced through the activity of the enzyme Dicer and are loaded into the RNA induced-silencing complex (RISC). The central protein is Argonaute. sRNAs then direct the complex to complementary nucleotide sequences, which can be mRNAs or DNA. Binding of the sRNA to DNA can lead to silencing of neighbouring genes through methylation while binding to an mRNA can lead to either translational repression or mRNA cleavage [10, 11].

After the discovery of RNAi in the fungus *Neurospora crassa* [12], numerous sRNA studies have been conducted in pathogenic and non-pathogenic fungal species, including *Cryptococcus neoformans* [13], *Trichoderma reesei* [14], *Penicillium marneffeii* [15], *Aspergillus flavus* [16], *Metarhizium anisopliae* [17], *Rhizophagus irregularis* [18], *Botrytis cinerea* [19], *Puccinia*

striiformis [20], *P. tritici* [15], *Fusarium oxysporum* [21] and *F. graminearum* [22, 23]. Some species like *Ustilago maydis* and *Saccharomyces cerevisiae* have lost their RNAi capability [24-26], whilst other species of same genus, such as *S. castellii* and *U. hordei*, have not, suggesting RNAi-related genes are not always essential [27, 28]. The endogenous roles of sRNAs in growth and development of filamentous fungi have been reported before in various studies [29-32].

There are two major classes of small RNA, micro RNAs (miRNAs) and small interfering RNAs (siRNAs). Mature miRNAs are generally 20-24 nt sequences that originate from single-stranded RNA precursors with the ability to form hairpin structures with imperfectly paired arms, whereas siRNAs are formed from perfectly matched dsRNA precursors [33]. Phased siRNAs (phasiRNAs) are a secondary class of siRNAs produced via the miRNA-mediated cleavage of mRNAs or non-coding RNA precursors. They are generated at precise 21-22 nucleotide intervals from the miRNA cleavage site. Although there has been extensive research on phasiRNAs in plants [34-37], whether these are associated with the fungal sRNAs has not yet been considered in detail. A single study conducted before identified phasiRNAs in *S. sclerotiorum* during mycovirus infection [38]. However, no reports are available on whether phasiRNAs are expressed during infection.

In recent years, next generation sequencing has been used for the identification and investigation of fungal sRNAs [29, 39-41]. The fungus *S. sclerotiorum* is a pathogen of hundreds of plant species, including many crops, such as the economically important oilseed *Brassica napus* [42]. A few studies have been conducted on sRNAs in *S. sclerotiorum* [29, 43, 44]. In addition to this, changes in the *S. sclerotiorum* sRNA transcriptome were studied during mycovirus infection [45]. The first study conducted on sRNAs on *S. sclerotiorum* by Zhou et

al [43] identified 44 miRNA-like RNAs. This study was conducted only to identify the presence of miRNA-like structures in *S. sclerotiorum*. Later, Derbyshire et al. [44] identified 374 highly abundant *S. sclerotiorum* sRNAs during infection of two hosts, *Arabidopsis thaliana* and *Phaseolus vulgaris*. This study did not investigate any miRNA-like structures. Recently, Zihao et al. predicted 275 miRNAs associated with sclerotial development [29] and reported endogenous regulation of a *histone acetyltransferase* gene. Mochama et al. demonstrated the presence of antiviral RNA silencing mechanisms in *S. sclerotiorum* during mycoviruses infection [45]. However, there are currently no sRNA transcriptome studies that have investigated the interaction with *B. napus*.

With the goal of identifying endogenous *S. sclerotiorum* sRNA targets regulated *in vitro* and during infection, we sequenced sRNAs from liquid *S. sclerotiorum* cultures and mats of fungal mycelium lifted from infected *B. napus* tissue. We assessed whether any sRNA loci were differentially expressed at 12 and 24 HPI and computationally predicted targets of the sRNAs they produced, verifying 14 with degradome sequencing. Furthermore, we predicted miRNA and PHAS loci; the latter of which have not been studied in detail in fungi to date. Together, the results from this chapter add a new resource for the study of sRNAs in *S. sclerotiorum*.

3.2 Methods

3.2.1 Fungal and plant sources

The *S. sclerotiorum* isolate CU8.24 collected from South Stirling, Western Australia, was used to generate sRNA and degradome datasets from *in vitro* and infected *B. napus* leaves [46]. Fully mature, dry sclerotia were cut in half and placed mycelium-side down on potato dextrose agar (PDA) (Becton Dickinson, USA) and incubated for a week at 20 °C in the dark. The germinated mycelium was subcultured onto the fresh PDA. After 48 hours of incubation, a 1

cm mycelial plug was inoculated onto minimal medium, which consists of 2 g/L NH₄NO₃, 1 g/L KH₂PO₄, 0.1 g/L MgSO₄·7H₂O, 0.5 g/L yeast extract, 4 g/L DL-malic acid and 1 g/L NaOH [47]. The two-day old mycelial mat from the minimal medium was spread over a detached first leaf of one-month-old *B. napus* (cv AV Garnet), placed in a Petri dish containing wet filter paper and incubated at 20 °C. After 12 and 24 HPI, a symptomatic lesion was observed on the surface of the leaves and intact mycelium was carefully separated from the leaves and immediately frozen in liquid nitrogen and stored at -80 °C [48] until RNA extraction. The *in vitro* mycelial mat before inoculation (0 HPI) was used as a reference to investigate sRNA transcriptome changes.

3.2.2 RNA isolation and sequencing

S. sclerotiorum mycelial mats were collected at 0, 12, and 24 HPI of *B. napus* detached leaves. The collected fungal material was ground into a fine powder with liquid nitrogen with a pre-cooled RNase-free mortar and pestle. Total RNA was extracted using the Trizol™ reagent following the manufacturer's protocol (Invitrogen, Carlsbad, CA, USA). Total RNA was quantified using a NanoDrop spectrophotometer and Qubit (Invitrogen) and the integrity was checked by agarose gel electrophoresis. Eight libraries of sRNAs were prepared from total RNA from three replicates each for *in vitro* mycelium and 24 HPI and duplicates for 12 HPI using the TruSeq small RNA sample preparation kit (Illumina, San Diego, CA, USA) according to the manufacturer's protocol. Thereafter, single-end (50 bp) sequencing was performed on an Illumina PE150 by Novogene (Novogene, Beijing, China). These datasets are also analysed in Chapter 4 where they are used to identify mRNA targets in *B. napus*.

3.2.3 Analysis of sequencing data

Total raw reads were trimmed using Cutadapt v1.15 [49], using the command ‘cutadapt -a adaptor.fa -o trimmed.fq raw.fq -m 18 -M30’, where ‘adaptor.fa’ contains the adaptor sequences used for the sRNA library preparation [50]. We retained sequences of a length of 18-30 nt which is the typical size of sRNAs. The quality of filtered reads was assessed with FastQC v0.11.8[51]. Trimmed reads were assigned to the pathogen [52] and the host [53] using bbsplit [54] with the ‘ambig2’ option set to ‘toss’, which discards reads that map equally well to both host and pathogen. Reads that were unique to the fungal genome were kept for prediction of *S. sclerotiorum* sRNAs. *S. sclerotiorum* specific sRNAs potentially derived from structural RNA (ribosomal RNA, snRNA, and snoRNA) were filtered out using the Rfam database [55] using the program Infernal, v1.1.3 with the command ‘cmscan --nohmmonly --rfam --cut_ga --fmt 2 --oclan --oskip -o strRNA.out --tblout strRNA.tblout structuralRNA.cm CU8_24.fasta [56]. First, Rfam hits in the fungal genome were prepared from the Rfam clanin file (cmfetch Rfam.cm). sRNA libraries were then mapped to the Rfam hits in the fungal genome using Bowtie2 [57]. Reads that were not aligned were used for prediction of sRNA loci. BEDtools v2.29.0 [58] was used to find reads overlapping to the genes and repetitive elements.

The clean reads of each library were used as a single entity to find sRNA-producing loci with the program ShortStack v3.8.5 (ShortStack --readfile \$fastq_dir --genomefile \$Genome --outdir \$outdir --bowtie_cores 6 --sort_mem 10G --mismatches 0 --mmap f --bowtie_m all --mincov 100rpm --dicermin 18 --dicermax 30) [33]. ShortStack first identifies significant alignment coverage based on depth of alignment. Significant alignments are then expanded upstream and downstream. Overlapped regions are thereafter merged to form clusters [59]. Clusters are annotated as ‘Dicer-derived’ (Dicer being the main enzyme the generates mature

sRNAs) loci when 80 percent of reads have a length of 18-30 nt. The most abundant sRNAs of Dicer-derived loci are referred to as ‘major RNAs’. Sequences with miRNA-like features were predicted using the miRCat tool implemented in the UEA small RNA workbench v4.5 with default parameters [60]. For prediction of miRNAs, sRNA libraries for each treatment were pooled to make a single file, therefore, three miRCat runs were conducted for 0, 12 and 24 HPI. This program uses PatMan to map sRNA reads to the input genome, using a flank extension of 100 bp, a series of putative miRNA precursors are then excised. Afterward, secondary structures of these putative precursors are predicted and retained below a minimum free energy threshold -25 and *P*-value of 0.05 using RNAfold v.2 [61]. We also used a second approach to predict miRNA loci using ShortStack. We focused on miRNA loci with both miRNA and miRNA* sequences as the presence of the miRNA* sequence in the locus provides strong evidence that these miRNAs are Dicer-derived [43].

3.2.3 Differential expression analysis

We used DESeq2 Bioconductor package in R 3.6.1 to normalize the total raw read counts from Shortstack for differential expression analysis [62]. After normalization of counts, we compared expression of sRNA loci between *in vitro* and 24 HPI, and *in vitro* and 12 HPI using the negative binomial test. Differentially expressed loci were filtered to remove false positive loci using Benjamini-Hochberg (BH) adjustment [63] and further differentially expressed loci were predicted based on a p-value threshold of < 0.01.

3.2.4 Small RNA target prediction

Target prediction of sRNAs was done using the psRNA target online tool (2017 release) [64]. All the parameters were kept default except the expectation score (<3). The psRNATarget

server finds target sequences based on complementarity according to a predefined score scheme and cleavage site recognition by calculating a threshold ratio of unpaired minimum free energy to paired minimum free energy [64]. We used two different approaches for endogenous target detection. The first approach simply comprised the input of sRNA datasets and *S. sclerotiorum*, transcript available in psRNA target server. While using this approach we might miss some of the genuine targets and it does not give enough information of repeat element genes. In the the second approach the sRNA data sets and the *S. sclerotiorum* CU8.24 genome were used as input files with the psRNA target tool. The sRNA/mRNA complementary sites were then used to find their positions in the *S. sclerotiorum* genome using Bedtools v2.29.0. We separately used genes and repeat elements files to characterize the predicted targets as repeat elements and non-repeat element (e.g., genes or introns) with Bedtools. Any targets that overlapped with repeat elements were called repeat element genes and the remaining as non-repeat element genes. For the functional annotation of the targets, 200 bp upstream and downstream were excised and compared with *S. sclerotiorum* InterPro and Pfam domains.

3.2.5 Detection of PHAS loci

The PHAS loci were predicted using ta-si prediction tool from UEA small RNA workbench [60]. This software requires an sRNA dataset and a genome file for input. We separately ran the collapsed file for 0 HPI, 12 HPI and 24 HPI to find the PHAS loci across all time points. At first, the program aligns sRNAs to the genome using PATMAN and any sRNAs that are not matched to the genome are discarded. Using the algorithm developed by Chen et al [65], it calculates the probability of the phasing being significant based on the hypergeometric distribution. Only 21 nt sRNAs were used in the phasing analysis.

3.2.6 Degradome analysis

To find putative cleavage sites, we performed degradome sequencing from mock and infected *in planta* samples at 24 HPI. Samples for degradome libraries were the same as those that were used for *in planta* small RNA sequencing. Three replicates each containing three leaf samples from different plants were pooled together and sequenced using the Illumina platform. Degradome sequencing was done as mentioned in [35]. In brief, construction of the degradome library was started from the degradation site (with a 5' monophosphate group) of the degraded mRNA. The sequencing adaptors were added to both ends of the degraded mRNAs and a library with an insert size of around 200-400 (base pairs) bp was generated. Paired end sequencing (2 x 150bp) was performed on a Hiseq 2500. The read length was shorter than the insert fragment length. Therefore, the sequencing data of the degradome library did not contain adapters, and they were all the same length. The length of the original raw data was 50bp.

3.3 Results

3.3.1 Sequencing data analysis

We generated ~ 62 million high-quality raw reads from eight different *S. sclerotiorum* mycelial mat libraries. Quality and size filtering resulted in a total of ~ 50 million reads. Among these reads, 72 % mapped unambiguously to the *S. sclerotiorum* genome. The structural RNA filtering resulted in ~ 30 million clean fungal sRNA reads, which were used for prediction of sRNA / miRNA loci. Overall, ~ 69 % of total sRNA reads were aligned to the fungal genome. Fig. 3.1 shows a bar diagram summarising the sRNA sequencing dataset.

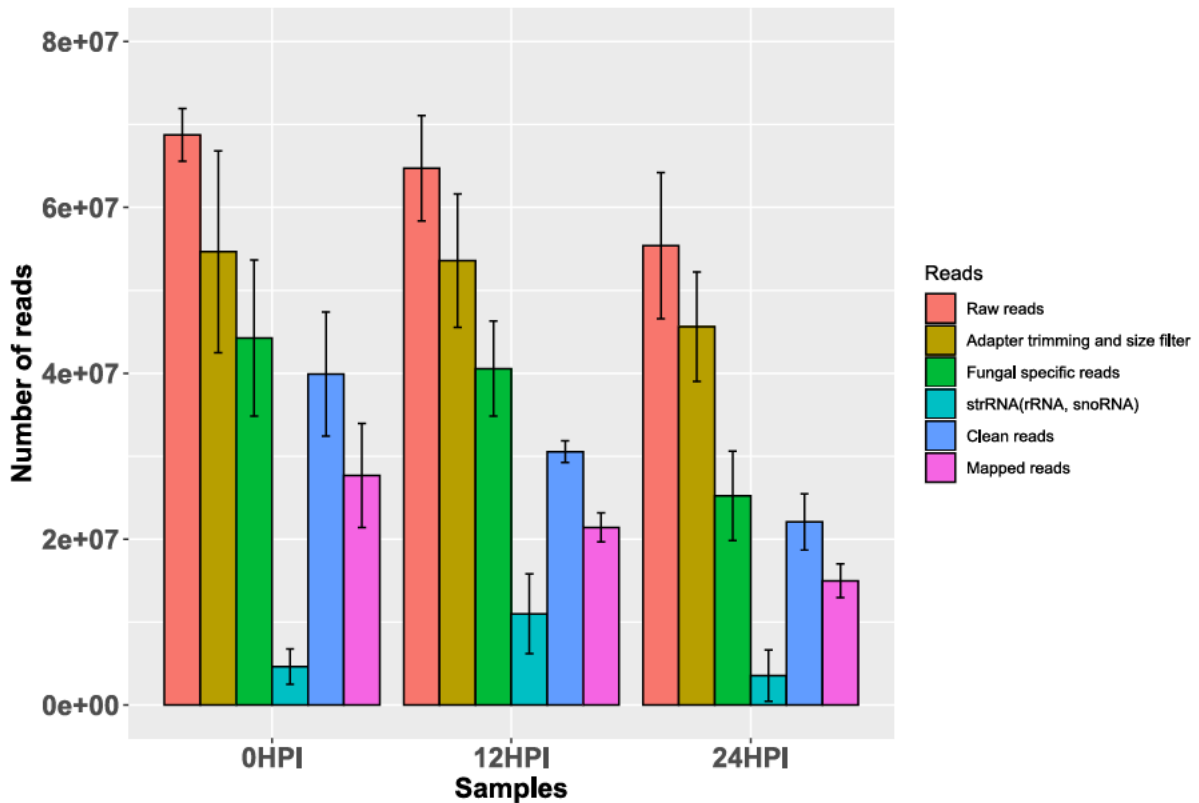


Fig. 3.1. Analysis of small RNA sequencing dataset of *Sclerotinia sclerotiorum*. The x-axis shows the sRNA libraries from where the sRNA reads were generated, y-axis shows the number of total reads identified for different reads; insert shows the representative figure of fungal mycelial mat experiment at 12 HPI and 24 HPI for sampling. R1, R2 and, R3 represent the data for three biological replicates.

Surprisingly, an average of 6.5 % of reads from uninoculated mycelial mats were best matched to the plant genome (Fig. 3.2). Although there was no cross contamination of plant tissue in these samples such mapping of fungal originated sRNAs might be due to transposable element sRNAs that are common in the plant and fungus. Sometimes the same families of repeats are present in unrelated species but the repeats may not have assembled in the *S. sclerotiorum* genome that was used.

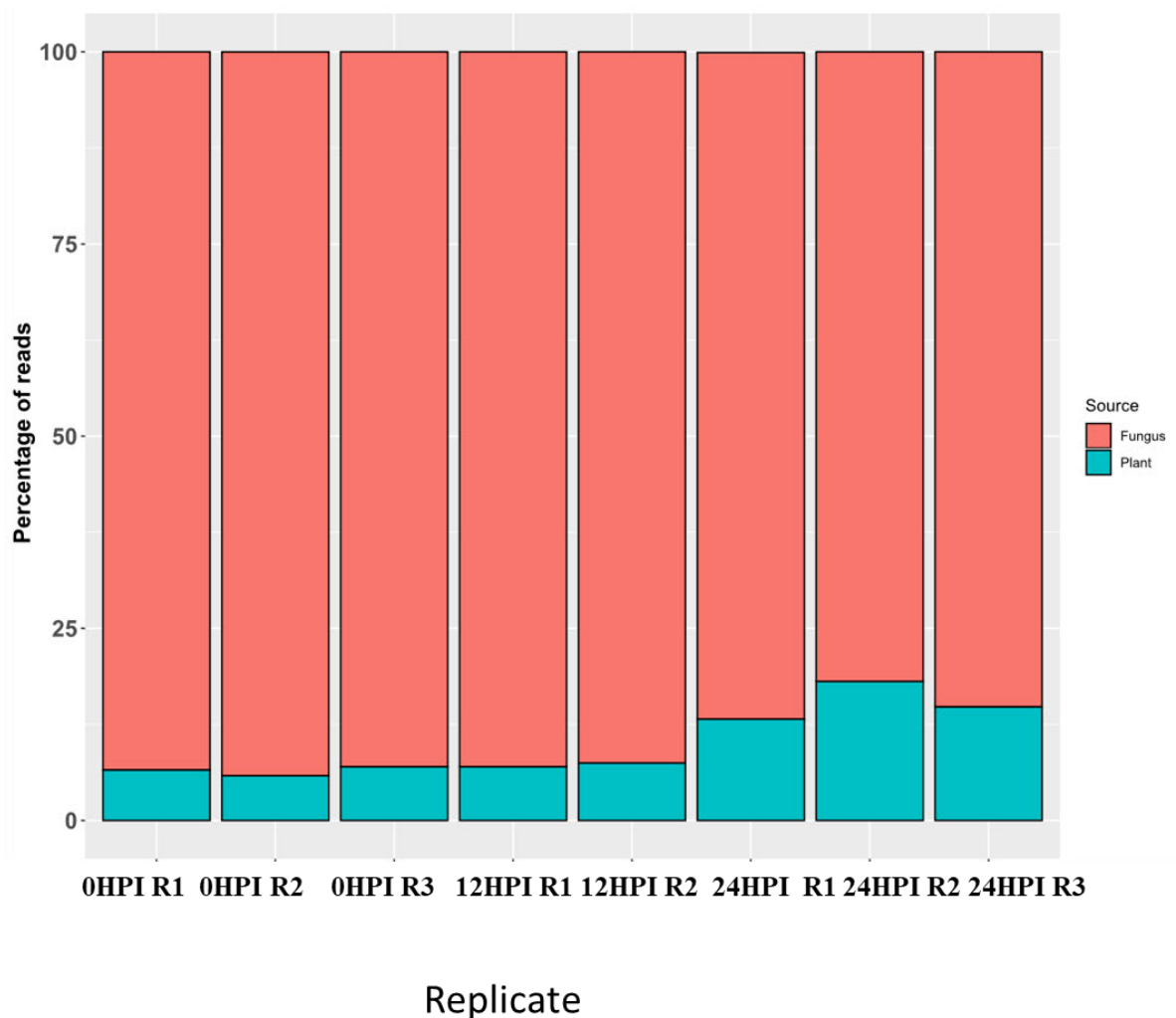


Fig. 3.2. Percentage of reads mapping to the *Sclerotinia sclerotiorum* and *Brassica napus* genomes. The x-axis shows the replicates of treatment and the y-axis shows the percentage of reads mapping to either *S. sclerotiorum* (orange) or *B. napus* (turquoise).

3.3.2 Assessment of sRNA population and characteristics

To understand the characteristics of the *S. sclerotiorum* sRNA population under different conditions, we examined the length distribution and 5' nucleotide bias of sRNA libraries. Non-specific RNA degradation results in a uniform distribution of short RNA sequences, whereas

dicer-dependent generation of sRNAs results in a peak between 20 and 25 nt [33, 59, 66]. The distribution of sRNA sequences followed a typical sRNA size distribution, with the majority of reads between 20 and 24 nt long. The most frequent read length was 22 nt, followed by 23 nt for both total and unique reads (Fig. 3.3). Highly abundant sRNAs in *S. sclerotiorum* had a bias toward uridine as the 5' nucleotide. This characteristic is attributed to the sorting of sRNAs into particular Argonaute proteins [67]. Interestingly, at 0 HPI, total reads had a 5' nucleotide bias of adenine at a length of 27 nt (Fig. 3.4). Previous studies have reported that *S. sclerotiorum* sRNAs exhibit a characteristic length distribution with most reads 22 and 23 nt in size with an enrichment of uracil as the 5' nucleotide [59, 68]. Therefore, most of these sRNAs are processed through the loading into Argonaute 1 proteins [40]. Assessing the abundance of both redundant and unique sRNA sequencing reads allows us to determine the complexity of the sRNAs expressed [69]. For example, if there are proportionally more redundant reads of a particular size than there are unique reads, it may indicate that sRNAs of that are dominated by high levels of expression of a few unique sequences. In *S. sclerotiorum*, we found that unique and redundant reads had proportionally similar sizes, indicating a fairly complex sRNA composition.

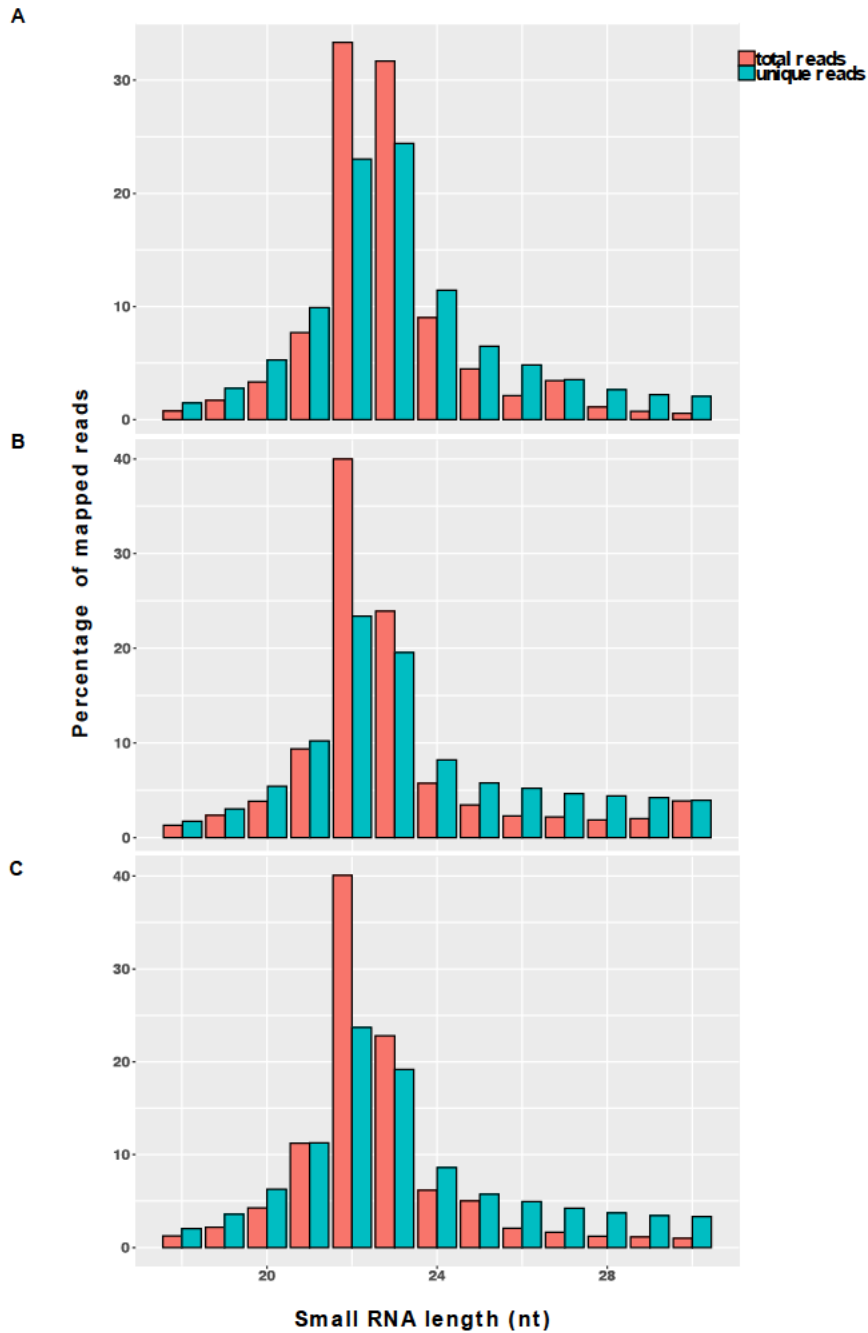


Fig. 3.3. length distribution of *Sclerotinia sclerotiorum* sRNAs. The percentage of reads (y-axis) according to nucleotide (nt) sequence length (x-axis) obtained in 0 HPI (A), 12 HPI, (B) and 24 HPI (C) pooled samples. The red bars indicate percentage of library for each length of total reads while the blue bars indicate percentage of library for each length of unique reads.

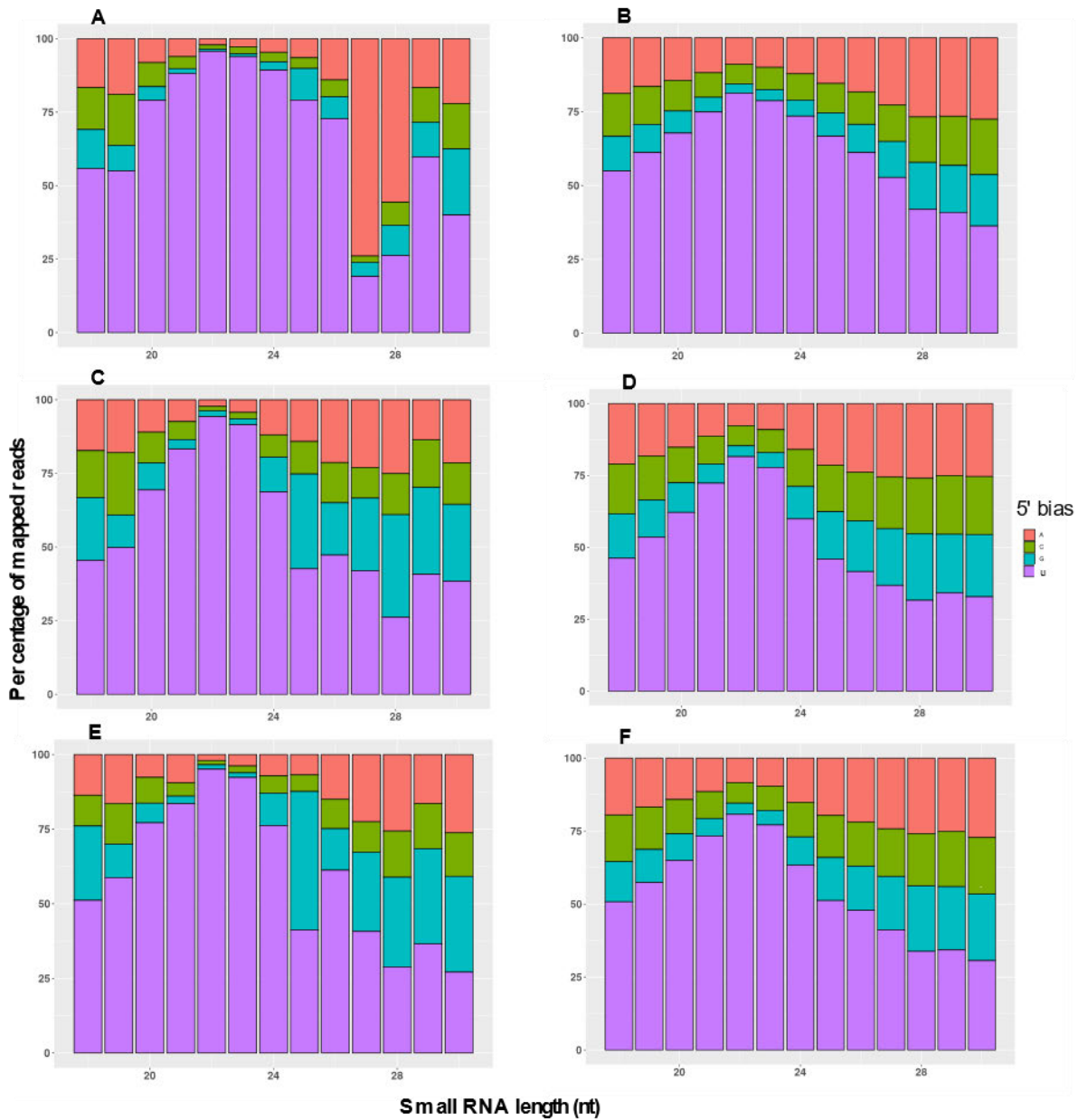


Fig. 3. 4. 5' nucleotide bias of *Sclerotinia sclerotiorum* small RNAs. The percentage of adenine (orange), cytosine (green), guanine (torquiose) and uridine (purple) in the 5' nucleotide (nt) position according to read length for 0 HPI total reads (A) and unique reads (B), 12 HPI total reads (C) and unique reads (D), 24 HPI total reads (E) and unique reads (F).

3.3.3 Annotation of sRNA generating loci in *Sclerotinia sclerotiorum*

Shortstack identified 1,073 Dicer-derived sRNA loci with 475 loci belonging to the negative strand of the chromosome and 598 from the positive strand. Among Dicer-derived loci, 245 originated from sense genic, 92 from antisense genic, 324 from repeat sense, and 670 from repeat antisense regions; a total of 354 loci had a read count of ≥ 100 RPM. We further investigated the origins of these loci and found almost 62 % overlapped with repeat elements, with 30 % genes, and the remaining with other regions including intergenic region and unannotated regions (Fig. 3.5).

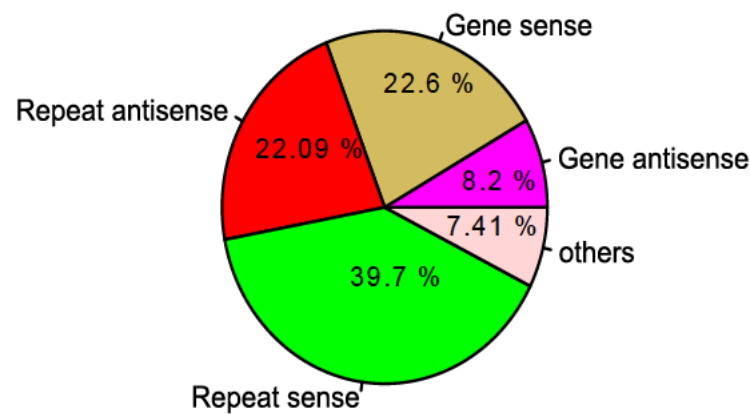


Fig. 3.5. Assessment of sRNA origin of 1,073 Dicer-derived loci in the *Sclerotinia sclerotiorum* genome.

3.3.4 Identification of *Sclerotinia sclerotiorum* miRNAs

To identify whether *S. sclerotiorum* encodes miRNAs, we employed the miRCat tool implemented in UEA sRNA workbench using default parameters [60]. It is also worth mentioning here that running Shortstack in -hp mode also predicts miRNA loci. The Shortstack

program did not predict any miRNA loci from virus-infected *S. sclerotiorum* in a previous study [70]. Here, we identified only one miRNA locus using the ShortStack program. The original location of this locus was Scaffold_74:155-245 with an annotated hairpin structure of UAAAGCGCAUCACUAGUAAUUCUUCGUGUUACUACCUAUCUCCAUCUGAAUGG GUAGAAGCACGAAUGGAUGAUGAUGAGAUCUAGGUUC and a mature sequence of UGAAUGGGUAGAAGCACGAAUGGAU (Fig. 3.6). In total, 30,849 sequence reads were found for this particular miRNA. Interestingly, a homology search of this miRNA sequence identified four closely matched miRNAs deposited in miRBase [71], including egr-miR-10241-5p [72], mdo-miR-7398l-3p [73], mdo-miR-7398-5p [73] and oni-miR-10798 [74]. In contrast, miRCat predicted a greater number of miRNA loci. Altogether, 1,313, 944, and 1,139 miRNA loci were predicted from 0, 12, and 24 HPI libraries, respectively, and these contained 1,293, 919, and 1,099 non-redundant miRNA sequences, respectively. However, only 44 and 64 miRNA sequences were common to 0 HPI + 12 HPI and 0 HPI + 24 HPI libraries, respectively. Furthermore, 16 sequences were common to all libraries. The expression analysis of 16 common miRNAs across all libraries is shown in Fig. 3.7B. The venn diagram showing the miRNAs identified uniquely or commonly at 0 HPI, 12 HPI and 24 HPI is shown in Fig. 3.8. The presence of miRNA* sequences provide strong evidence of Dicer-mediated generation of the mature miRNA sequence.

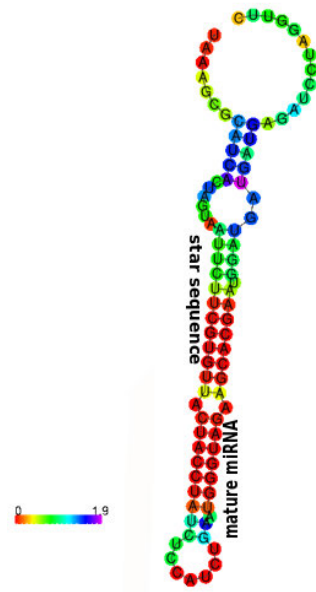


Fig. 3.6. The hairpin loop structure of the miRNA locus predicted by ShortStack. The different colours show the intensity of minimum free energy associated with the secondary structure.

3.3.5 Small RNAs are differentially expressed in *Sclerotinia sclerotiorum* mycelium after host inoculation

The individual raw reads for each library from Shortstack were normalized using DESeq2 in R [62]. We used the 354 abundant loci with sRNA reads of more than 100 RPM to find differentially expressed loci. Altogether, 34 and 57 loci were altered in their expression profiles, with a p-value of less than 0.01, across 12 and 24 HPI time points, respectively. From these loci, 25 were upregulated in at least one timepoint post infection (Fig. 3.7A) suggesting they were infection-induced.

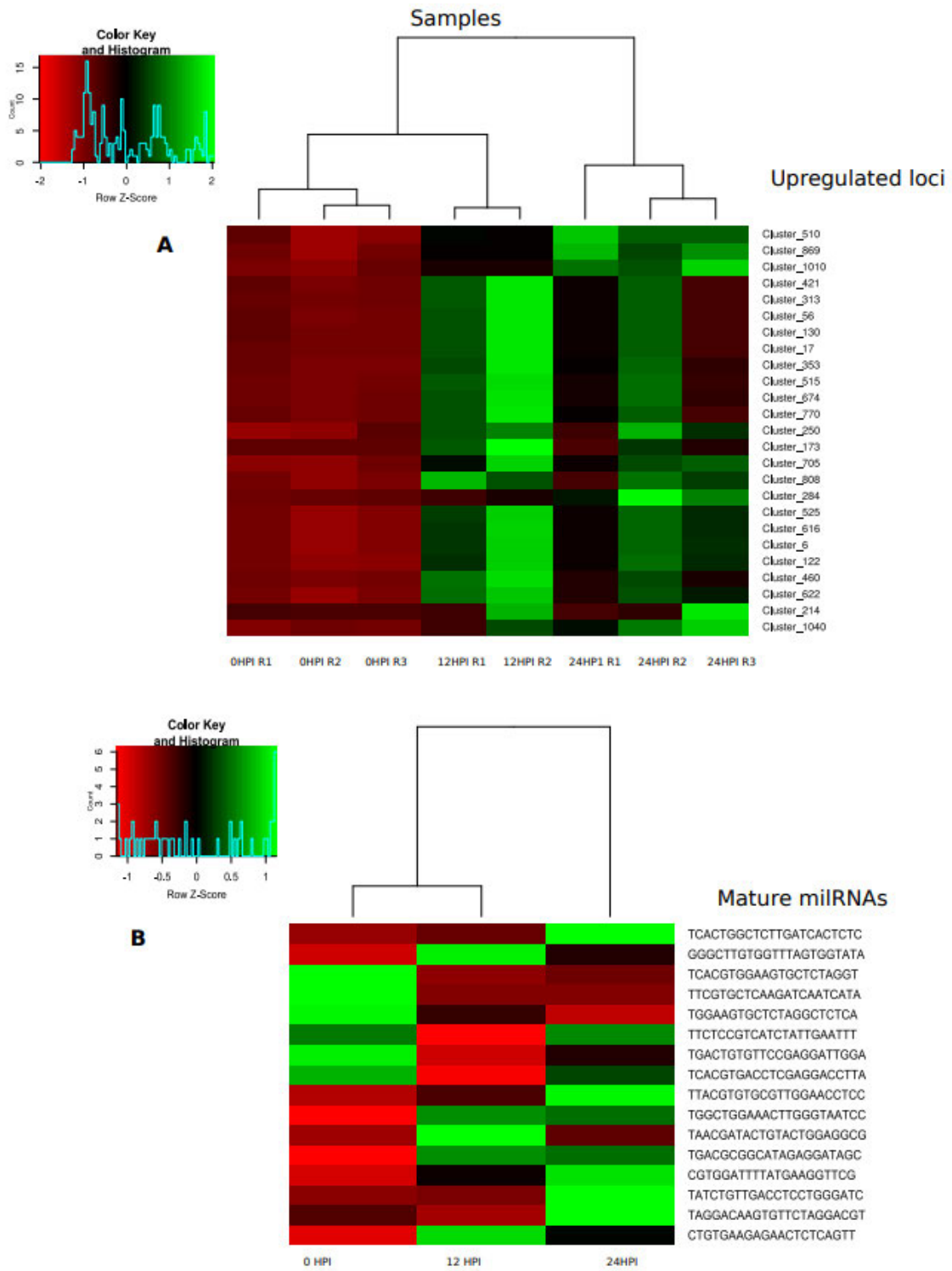


Fig. 3.7. Heatmap of sRNAs identified in the *Sclerotinia sclerotiorum* genome. A Heat map of normalized expression data from the 25 ShortStack loci commonly up-regulated across both time points, 12 and 24 hours post inoculation (HPI). B The expression profile of 16 mature miRNAs common to all three libraries normalized to reads per million.

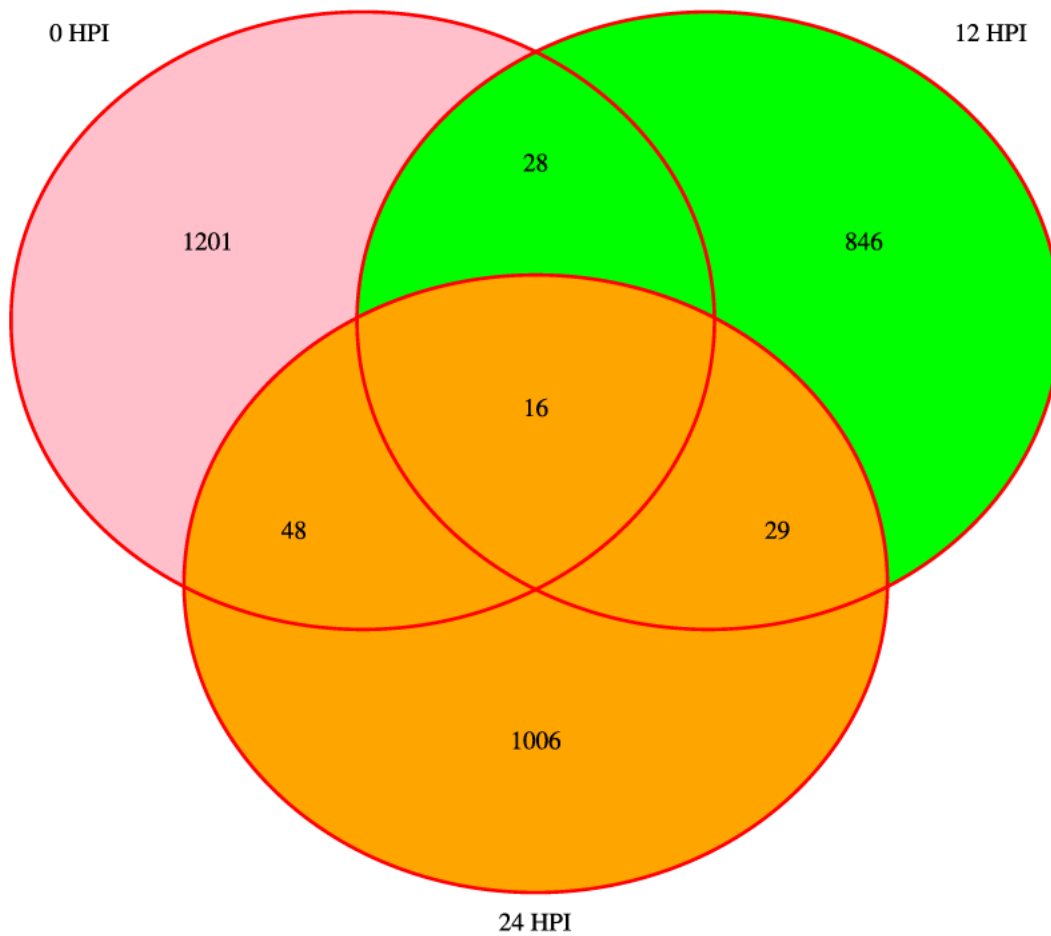


Fig. 3. 8. Venn diagram showing the number of miRNAs identified from *Sclerotinia sclerotiorum* from different libraries.

From our dataset, we found 24 miRNA loci with both mature and miRNA* sequences, and these contained 20 unique miRNAs. We found homology for 12 miRNA sequences in the miRBase database. Details of these loci are shown in Table 3.1.

Table. 3. 1. Description of 24 miRNA loci with mature miRNA and miRNA* sequences with homology to known miRNA sequences in miRBase identified in the *Sclerotinia sclerotiorum* CU8.24 genome, using the sRNA sequencing data and the miRCat software.

Loci	miRNA	Reads	Homology	Organism	Sample
scaffold_117	AATACGCGCTGCCGAGAATC	16	sja-miR-3482-3p	<i>Schistosoma japonicum</i>	12 HPI
scaffold_184	TCGGACTCGGCGTAATCTGCG	135	ata-miR166b-5p	<i>Aegilops tauschii</i>	12 HPI
scaffold_74	TTCTTCGTGTTACTACCTATCT	11885	NA		12 HPI
scaffold_169	TATATAGATCTTCGTAAGTAGG	3970	NA		24 HPI
scaffold_206	TCACTGGCTCTTGATCACTCTC	62437	bmo-miR-3260, ppc-miR-8256-5p, tca-miR-3872-5p	<i>Bombyx mori</i> , <i>Pristionchus pacificus</i> , <i>Tribolium castaneum</i>	24 HPI
scaffold_222	TAAGGCGGGAACATTTTTGAGG	77	cpo-miR-507a-3p	<i>Cavia porcellus</i>	24 HPI
scaffold_286	ACCGAATCTGGGTTTGAGAC	16	ptc-miR6459b, ptc-miR6459a-3p	<i>Populus tichocarpa</i>	24 HPI

scaffold_32	TCACTGGCTCTTGATCACTCTC	62437	bmo-miR-3260, ppc-miR-8256-5p, tca-miR-3872-5p	<i>Bombyx mori</i> , <i>Pristionchus</i> <i>pacificus</i> , <i>Tribolium</i> <i>castaneum</i>	24 HPI
scaffold_337	TCACTGGCTCTTGATCACTCTC	62437	bmo-miR-3260, ppc-miR-8256-5p, tca-miR-3872-5p	<i>Bombyx mori</i> , <i>Pristionchus</i> <i>pacificus</i> , <i>Tribolium</i> <i>castaneum</i>	24 HPI
scaffold_10	TCGTGCTTGATGTCTTTGAGGT	4	gma-miR862a, gma-miR862b	<i>Glycine max</i>	0 HPI
scaffold_133	TTTAGGTCTAGGTCGTCTGTCT	18	NA		0 HPI
scaffold_165	TGTGCGGCTGTGGGACTAAAGC	58	NA		0 HPI
scaffold_169	TATATAGATCTTCGTAAGTAGG	11616	NA		0 HPI
scaffold_18	TTTCCCTGAGGTATCCTGGCTA	37	crm-miR-64d-3p	<i>Caenorhabditis</i> <i>remanei</i>	0 HPI
scaffold_19	TCGATTTGAGATAGTGTCTC	100	NA		0 HPI

scaffold_194	TAGCACTTCTAGGGATTCCGCC	24	NA		0 HPI
scaffold_218	TCATGAAGGGTTATGGTGGGT	30	aca-miR-5449	<i>Anolis carolinensis</i>	0 HPI
scaffold_319	TGTGTGTGTATGATTCTATTGC	40	oni-miR10680, mmu-miR-1187, atr-miR8565g	<i>Oreochromis niloticus</i> , <i>Mus musculus</i> , <i>Amborella trichopoda</i>	0 HPI
scaffold_52	TCTCAAAGGATCTGTTTCAGCAC	112	vca-miR10208-5p, egr-miR-10276-3p	<i>Vriesea carinata</i> , <i>Echinococcus granulosus</i>	0 HPI
scaffold_56	TAGTAGTGACTCTTCCTCGGAT	45	gga-miR-6559-5p	<i>Gallus gallus</i>	0 HPI
scaffold_7	TGTTACGCTGCGGAATTTGACA	46	gga-miR-6579-5p	<i>Gallus gallus</i>	0 HPI
scaffold_71	TGACTGTGTTCCGAGGATTGGA	126	ath-miR2111b-3p	<i>Arabidopsis thaliana</i>	0 HPI
scaffold_74	TCTTCGTGTTACTACCTATCTC	17753	NA		0 HPI

3.3.4 Prediction of phased loci in *Sclerotinia sclerotiorum*

To identify phasiRNAs in *S. sclerotiorum* we used the UEA sRNA workbench [60]. Altogether, 26 phased loci were predicted from combined *in vitro* and infected samples. In total, 217 and 147 unique phasiRNAs were produced from these loci. Detection of loci producing phasiRNAs indicates that some *S. sclerotiorum* sRNAs could act in a similar fashion to miRNAs in higher organisms [70]. Details of predicted phased loci are shown in Table 3.2.

Table. 3. 2. Identification of phasiRNAs producing loci in *Sclerotinia sclerotiorum*

Scaffold	Start position	End position	Total sequences	Phased sequences	p-value
scaffold_133	135,378	135,629	5	4	1.56 x10 ⁻⁵
scaffold_134	125,419	125,670	3	3	8.20 x10 ⁻⁵
scaffold_143	104,980	105,231	60	10	7.58 x10 ⁻⁵
scaffold_149	58,362	58,613	47	9	6.46x 10 ⁻⁵
scaffold_230	26,437	26,688	21	7	1.10 x10 ⁻⁵
scaffold_231	121,018	121,269	31	8	1.84 x10 ⁻⁵
scaffold_242	18,700	18,951	17	6	3.79 x10 ⁻⁵
scaffold_242	95,820	96,071	6	4	4.55 x10 ⁻⁵
scaffold_309	28	279	59	10	6.48 x10 ⁻⁵
scaffold_353	311	562	34	8	3.91x10 ⁻⁵

scaffold_41	165,625	165,876	3	3	8.20 x10 ⁻⁵
scaffold_43	63,778	64,029	16	6	2.52 x10 ⁻⁵
scaffold_431	1,985	2,236	100	13	5.10 x10 ⁻⁵
scaffold_440	1,510	1,761	71	11	5.49 x10 ⁻⁵
scaffold_552	752	1,003	135	15	5.30 x10 ⁻⁵
scaffold_599	548	799	10	5	2.60 x10 ⁻⁵
scaffold_6	137,162	137,413	3	3	8.20 x10 ⁻⁵
scaffold_612	586	837	73	11	7.28 x10 ⁻⁵
scaffold_643	182	433	128	15	2.56 x10 ⁻⁵
scaffold_659	133	384	8	5	6.14x10 ⁻⁵
scaffold_668	224	475	127	15	2.30x10 ⁻⁵
scaffold_686	404	655	60	10	7.58x10 ⁻⁵
scaffold_713	448	699	55	10	3.33x10 ⁻⁵
scaffold_78	5,213	5,464	3	3	8.20x10 ⁻⁵
scaffold_82	29	280	59	10	6.48x10 ⁻⁵
scaffold_841	312	563	90	13	1.44x10 ⁻⁵

3.3.6 Endogenous targets of fungal sRNAs were mostly repetitive elements

For target prediction, we divided *S. sclerotiorum* sRNAs into two sets. The first consisted of 25 major sRNAs from loci up-regulated commonly across both 12 and 24 HPI timepoints and the second comprised 165 miRNAs, predicted using miRCat, with ≥ 100 reads found in all libraries. From the psRNA online target prediction tool, with a maximum expectation score of 3.5 [64], 343 targets were predicted in the *S. sclerotiorum* genome from 25 upregulated sRNAs. Among these targets, 54 were annotated as genes while 343 targets overlapped with repeat elements. We also used our fungal specific degradome sets (detailed in Chapter 4) to see whether there was any evidence of cleavage of endogenous targets. Interestingly, we found two transposable element targets from the analysis of the degradome dataset. Furthermore, we found 4,917 endogenous targets regulated by 165 miRNAs (Table 3.3). Among these targets, 475 were annotated as genes, while 4,621 were related to repeat elements. Furthermore, we found evidence for cleavage of 12 targets from our fungal specific degradome dataset (summarised in Supplementary File 3.1). One of the targets “*sscle_01g003340*” was predicted to be cleaved by four miRNAs which encodes a domain related to RNA-binding, a NAB2 type zinc finger. Other confirmed targets were “*sscle_03g028020*”, “*sscle_15g102960*”, “*sscle_03g030160*”, “*sscle_07g059010*”, which encode domains related to phosphate acyltransferases, eukaryotic protein of unknown function (DUF829), emopamil binding protein, guanine nucleotide exchange factor for Ras-like GTPases; N-terminal motif, respectively. The remaining four targets were “*sscle_15g106630*”, “*sscle_15g106740*”, “*sscle_08g066610*”, “*sscle_01g011410*” which had no known Pfam domains.

Table. 3.3. Summary of endogenous targets predicted from *Sclerotinia sclerotiorum* sRNAs. The number of targets that overlapped genes or repeat elements in the *S. sclerotiorum* CU8.24 genome using second approach with the input of both gene and repeat element files.

sRNA	Total Targets	Repeat elements	Genes	Degradome
Upregulated 25 sRNAs	343	307	54	2
Abundant 165 miRNAs	4,917	4,621	475	12

3.4. Discussion

Small RNAs are short non-coding RNAs derived from either from double-stranded or hairpin RNAs [13, 35]. Depending on the precursor and biogenesis pathways, sRNAs can be classified as siRNAs and miRNAs. siRNAs are formed from double-stranded precursors while miRNAs are formed from hairpin loop structures [10]. Besides this, there are other categories of siRNAs identified in plants and animals like hairpin-siRNAs, natural-antisense siRNAs, heterochromatic siRNAs and secondary siRNAs. PhasiRNAs are a subclass of siRNA which are produced by miRNA-mediated cleavage of mRNAs or non-coding RNAs [75]. In our study, we identified 1,073 Dicer-derived loci, 165 highly expressed miRNA loci, and 26 phased loci in the *S. sclerotiorum* genome that were identified from sRNA sequencing datasets from *in vitro* cultures and infected detached leaves.

In this study, *S. sclerotiorum* mycelial mats sampled from three different stages were used to develop sRNA libraries. The size distribution of total and unique reads was investigated. The

length distribution of sRNAs in *S. sclerotiorum* peaked at 22 and 23 nt, which is fairly similar to other fungi like *F. oxysporum* [31], *P. marneffeii* [76] and *A. flavus* [16]. Previous studies in *S. sclerotiorum* report that the most abundant sRNAs are 22 and 23 nt in length with a 5' bias of uridine [29, 44], which agrees with the findings of this study with 5' bias of uridine. The former study identified 374 highly abundant sRNAs during infection of two host plants *A. thaliana* and beans while the latter study has analysed the response of *S. sclerotiorum* during mycovirus infection. We found a 5' uridine bias in the most abundant, sRNAs suggesting that these sRNAs bind to AGO1 [67]. Interestingly, total reads of 0 HPI had a 5' nucleotide bias of adenine at a length of 27 nt. While there are no reports of such a bias in fungal sRNAs, 5' adenine biased siRNAs have been previously reported to be involved in RNA dependent DNA methylation in *A. thaliana*, with the loading into the Argonaute protein AGO4 [67].

sRNA mediated RNAi is an important gene regulation mechanism in plants, animals and fungi [11]. miRNAs in fungi, similar to plants and animals, have recently been identified to regulate different process like development, pathogenesis and reproduction [29, 31, 77]. It has been well demonstrated that sRNAs play key roles in development and stress responses. In this study, we identified 25 commonly-upregulated sRNAs at 12 and 24 HPI during *B. napus* infection. The expression patterns of these sRNAs suggest possible roles in growth and pathogenicity on *B. napus*. sRNAs have previously been shown to have a key role in regulation of transposable elements [15, 78]. Target identification of sRNAs including miRNAs revealed almost 90 % of the targets as repeat elements. sRNAs contribute to maintaining genome integrity by regulation of repeat elements. It is not surprising to see regulation of repetitive elements by fungal sRNAs, since one of the major functions of sRNAs is silencing of repeats/transposons to maintain genome integrity [40]. In the wheat pathogen *P. triticina* 32.19 % of endogenous targets of pathogen sRNAs were repeat elements [15]. This research further enhances our knowledge of

sRNA classes in *S. sclerotiorum* and how they change over time during *B. napus* interactions. Although there were hundreds of targets identified from *in silico* predictions, we were able to confirm the endogenous cleavage of 14 targets from this study using degradome sequencing data. Interestingly, domains related to binding factor, exchange factor and phosphate acetyltransferase were encoded by these genes. Fungal pathogen induced phosphate acetyltransferase have been reported to be involved in pathogen interactions with the host by affecting host cell wall and causing dehydration of leaves [79, 80]. Interestingly, gene *sscle_03g028020* was cleaved by miRNAs suggesting *S. sclerotiorum* miRNA endogenously regulate this class of genes. RNA-binding, NAB2-type zinc finger domain are involved in poly (A) tailing of mRNAs [81]. Depending on how mRNA binds to RNA binding protein determines its fate for further functions. Poly(A) tailing is especially important for regulation of transcript stability [82]. Gene *sscle_01g003340* identified here has a domain related to RNA binding proteins therefore fungal miRNAs may also play an important role in transcript stability. Emopamil binding protein is a membrane protein of the endoplasmic reticulum. In *B. cinerea* this protein has been shown to have a role in virulence of pathogen [83]. The gene *sscle_03g030160* is regulated by miRNAs and encodes an emopamil binding protein suggesting miRNAs identified here could play a role in virulence of *S. sclerotiorum*. Guanine nucleotide exchange factor for Ras-like GTPases; N-terminal motif are involved in the catalytic activity of the exchange factor in fungi [84, 85]. These classes of genes have been shown to be involved in vesicle trafficking, endocytosis and development processes in *F. graminearum* [85]. The gene *sscle_07g059010* identified here to be regulated by miRNAs, may be important for similar processes in *S. sclerotiorum*.

Fungi are one of the three lineages of eukaryotes, along with plants and animals [86]. Although the sRNA machinery is present in fungi, plants and animals, conserved miRNAs are often not

detected. So far, there are no homologs of *S. sclerotiorum* miRNAs were identified in plants and animals. In this study, we identified 13 miRNAs which have homologues in other kingdoms. This result indicated that some miRNAs might be conserved between kingdoms [87, 88]. These identified miRNAs homologues were reported to have a role in morphology, bi-directional movement, nutrient metabolism and in embryonic development in worms and mammals [72, 74].

We also showed that *S. sclerotiorum* might regulate the gene expression through the production of secondary siRNAs like phasiRNAs. In plants, phasiRNAs are either produced by mRNAs or non-coding RNAs [89]. The locus producing phasiRNAs are called PHAS genes. Endogenous phasiRNAs play important regulatory roles by silencing protein coding transcripts and are associated with the pathogen infection [90]. Here, we identified 26 PHAS genes with 147 phasiRNAs with 730 targets in *S. sclerotiorum*. In summary, this study enhances the knowledge of endogenous *S. sclerotiorum* sRNAs and their targets.

3.5 References

1. Dodds, P.N. and J.P. Rathjen, *Plant immunity: towards an integrated view of plant–pathogen interactions*. Nature Reviews Genetics, 2010. **11**(8): p. 539.
2. Galindo-Gonzalez, L. and M.K. Deyholos, *RNA-seq transcriptome response of flax (*Linum usitatissimum* L.) to the pathogenic fungus *Fusarium oxysporum* f. sp. *lini**. Frontiers in Plant Science, 2016. **7** (1766):p.1-22
3. Girard, I.J., et al., *RNA sequencing of Brassica napus reveals cellular redox control of Sclerotinia infection*. Journal of Experimental Botany, 2017. **68**(18): p. 5079-5091.

4. Pertea, M., et al., *Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown*. Nature Protocols, 2016. **11**(9): p. 1650.
5. Westermann, A.J., S.A. Gorski, and J. Vogel, *Dual RNA-seq of pathogen and host*. Nature Reviews Microbiology, 2012. **10**(9): p. 618-630.
6. Feng, H., et al., *Exploration of microRNAs and their targets engaging in the resistance interaction between wheat and stripe rust*. Frontiers in Plant science, 2015. **6**: p. 469.
7. Huang, J., M. Yang, and X. Zhang, *The function of small RNAs in plant biotic stress response*. Journal of Integrative Plant Biology, 2016. **58**(4): p. 312-327.
8. Weiberg, A., et al., *Fungal small RNAs suppress plant immunity by hijacking host RNA interference pathways*. Science, 2013. **342**(6154): p. 118-123.
9. Großhans, H. and W. Filipowicz, *Molecular biology: the expanding world of small RNAs*. Nature, 2008. **451**(7177): p. 414.
10. Guleria, P., et al., *Plant small RNAs: biogenesis, mode of action and their roles in abiotic stresses*. Genomics, Proteomics & Bioinformatics, 2011. **9**(6): p. 183-199.
11. Dang, Y., et al., *RNA interference in fungi: pathways, functions, and applications*. Eukaryotic Cell, 2011. **10**(9): p. 1148-1155.
12. Romano, N. and G. Macino, *Quelling: transient inactivation of gene expression in Neurospora crassa by transformation with homologous sequences*. Molecular Microbiology, 1992. **6**(22): p. 3343-3353.
13. Liu, H., et al., *RNA interference in the pathogenic fungus Cryptococcus neoformans*. Genetics, 2002. **160**(2): p. 463-470.

14. Kang, K., et al., *Identification of microRNA-Like RNAs in the filamentous fungus Trichoderma reesei by solexa sequencing*. PloS One, 2013. **8**(10): p. e76288.
15. Dubey, H., et al., *Discovery and profiling of small RNAs from Puccinia triticina by deep sequencing and identification of their potential targets in wheat*. Functional & Integrative Genomics, 2019. **19**(3): p. 391-407.
16. Bai, Y., et al., *sRNA profiling in Aspergillus flavus reveals differentially expressed miRNA-like RNAs response to water activity and temperature*. Fungal Genetics and Biology, 2015. **81**: p. 113-119.
17. Zhou, Q., et al., *Genome-wide identification and profiling of microRNA-like RNAs from Metarhizium anisopliae during development*. Fungal Biology, 2012. **116**(11): p. 1156-1162.
18. Silvestri, A., et al., *In silico analysis of fungal small RNA accumulation reveals putative plant mRNA targets in the symbiosis between an arbuscular mycorrhizal fungus and its host plant*. BMC Genomics, 2019. **20**(1): p. 1-18.
19. Yang, F., *Genome-wide analysis of small RNAs in the wheat pathogenic fungus Zymoseptoria tritici*. Fungal Biology, 2015. **119**(7): p. 631-640.
20. Wang, B., et al., *Puccinia striiformis f. sp. tritici mi croRNA-like RNA 1 (Pst-milR1), an important pathogenicity factor of Pst, impairs wheat resistance to Pst by suppressing the wheat pathogenesis-related 2 gene*. New Phytologist, 2017. **215**(1): p. 338-350.

21. Fei, S., et al., *Small RNA profiling of Cavendish banana roots inoculated with Fusarium oxysporum f. sp. cubense race 1 and tropical race 4*. Phytopathology Research, 2019. **1**(1): p. 22.
22. Jian, J. and X. Liang, *One small RNA of Fusarium graminearum targets and silences CEBiP gene in common wheat*. Microorganisms, 2019. **7**(10):p.425
23. Jin, X., et al., *Identification of Fusarium graminearum-responsive miRNAs and their targets in wheat by sRNA sequencing and degradome analysis*. Functional & Integrative Genomics, 2019:20 (1) p. 1-11.
24. Drinnenberg, I.A., G.R. Fink, and D.P. Bartel, *Compatibility with killer explains the rise of RNAi-deficient fungi*. Science, 2011. **333**(6049): p. 1592-1592.
25. Drinnenberg, I.A., et al., *RNAi in budding yeast*. Science, 2009. **326**(5952): p. 544-550.
26. Billmyre, R.B., et al., *RNAi function, diversity, and loss in the fungal kingdom*. Chromosome Research, 2013. **21**(6): p. 561-572.
27. Nakayashiki, H., N. Kadotani, and S. Mayama, *Evolution and diversification of RNA silencing proteins in fungi*. Journal of Molecular Evolution, 2006. **63**(1): p. 127-135.
28. Laurie, J.D., et al., *Genome comparison of barley and maize smut fungi reveals targeted loss of RNA silencing components and species-specific presence of transposable elements*. The Plant Cell, 2012. **24**(5): p. 1733-1745.
29. Xia, Z., et al., *Characterization of microRNA-like RNAs associated with sclerotial development in Sclerotinia sclerotiorum*. Fungal Genetics and Biology, 2020. **144**: p. 103471.

30. Shao, J., et al., *Identification of miRNAs and their target genes in Ganoderma lucidum by high-throughput sequencing and degradome analysis*. Fungal Genetics and Biology, 2020. **136**: p. 103313.
31. Chen, R., et al., *Exploring microRNA-like small RNAs in the filamentous fungus Fusarium oxysporum*. PloS One, 2014. **9**(8): p. e104956.
32. Zeng, W., et al., *Dicer-like proteins regulate sexual development via the biogenesis of perithecium-specific microRNAs in a plant pathogenic fungus Fusarium graminearum*. Frontiers in Microbiology, 2018. **9**: p. 818.
33. Axtell, M.J., *ShortStack: comprehensive annotation and quantification of small RNA genes*. RNA, 2013. **19**(6): p. 740-751.
34. Guo, J., et al., *Analysis of microRNAs, phased small interfering RNAs and their potential targets in Rosa rugosa Thunb.* BMC Genomics, 2019. **19**(9): p. 1-13.
35. Ma, Z., C. Coruh, and M.J. Axtell, *Arabidopsis lyrata small RNAs: transient MIRNA and small interfering RNA loci within the Arabidopsis genus*. The Plant Cell, 2010. **22**(4): p. 1090-1103.
36. Deng, P., et al., *Biogenesis and regulatory hierarchy of phased small interfering RNAs in plants*. Plant Biotechnology Journal, 2018. **16**(5): p. 965-975.
37. Fei, Q., R. Xia, and B.C. Meyers, *Phased, secondary, small interfering RNAs in posttranscriptional regulatory networks*. The Plant Cell, 2013. **25**(7): p. 2400-2415.
38. Mochama, P., et al., *Mycoviruses as triggers and targets of RNA silencing in white mold fungus Sclerotinia sclerotiorum*. Viruses, 2018. **10**(4): p. 214.

39. Hunt, M., et al., *Small RNA discovery in the interaction between barley and the powdery mildew pathogen*. BMC Genomics, 2019. **20**(1): p. 610.
40. Kusch, S., et al., *Small RNAs from cereal powdery mildew pathogens may target host plant genes*. Fungal Biology, 2018. **122**(11): p. 1050-1063.
41. Chen, Y., et al., *Characterization of RNA silencing components in the plant pathogenic fungus *Fusarium graminearum**. Scientific Reports, 2015. **5**: p. 12500.
42. Boland, G.J. and R. Hall, *Index of plant hosts of *Sclerotinia sclerotiorum**. Canadian Journal of Plant Pathology, 1994. **16**(2): p. 93-108.
43. Zhou, J., et al., *Identification of microRNA-like RNAs in a plant pathogenic fungus *Sclerotinia sclerotiorum* by high-throughput sequencing*. Molecular Genetics and Genomics, 2012. **287**(4): p. 275-282.
44. Derbyshire, M., et al., *Small RNAs from the plant pathogenic fungus *Sclerotinia sclerotiorum* highlight host candidate genes associated with quantitative disease resistance*. Molecular Plant Pathology, 2019. **20**(9): p. 1279-1297.
45. Mochama, P., et al., *Mycoviruses as triggers and targets of RNA silencing in white mold fungus *Sclerotinia sclerotiorum**. Viruses, 2018. **10**(4).
46. Denton-Giles, M., et al., *Partial stem resistance in *Brassica napus* to highly aggressive and genetically diverse *Sclerotinia sclerotiorum* isolates from Australia*. Canadian Journal of Plant Pathology, 2018: p. 1-11.

47. Perchepped, L., et al., *Nitric oxide participates in the complex interplay of defense-related signaling pathways controlling disease resistance to sclerotinia sclerotiorum in arabidopsis thaliana*. *Molecular Plant-Microbe Interactions*, 2010. **23**(7): p. 846-860.
48. Mwape, V.W., et al., *Analysis of differentially expressed Sclerotinia sclerotiorum genes during the interaction with moderately resistant and highly susceptible chickpea lines*. *BMC Genomics*, 2021. **22**(1): p. 1-14.
49. Martin, M., *Cutadapt removes adapter sequences from high-throughput sequencing reads*. *EMBnet. Journal*, 2011. **17**(1): p. 10-12.
50. Bolger, A. and F. Giorgi, *Cutadapt removes adapter sequences from high-throughput sequencing reads*.
51. Andrews, S., *FastQC: a quality control tool for high throughput sequence data*. 2010, Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom.
52. Derbyshire, M., et al., *The complete genome sequence of the phytopathogenic fungus Sclerotinia sclerotiorum reveals insights into the genome architecture of broad host range pathogens*. *Genome Biology and evolution*, 2017. **9**(3): p. 593-618.
53. Chalhoub, B., et al., *Early allopolyploid evolution in the post-Neolithic Brassica napus oilseed genome*. *Science*, 2014. **345**(6199): p. 950-953.
54. Bushnell, B.B.A.F., *Accurate, Splice-Aware Aligner*. United States. , *BBMap: A Fast, Accurate, Splice-Aware Aligner*. United States. . 2014.
55. Kalvari, I., et al., *Non-coding RNA analysis using the Rfam database*. *Current Protocols in Bioinformatics*, 2018. **62**(1): p. e51.

56. Nawrocki, E.P. and S.R. Eddy, *Infernal 1.1: 100-fold faster RNA homology searches*. *Bioinformatics*, 2013. **29**(22): p. 2933-2935.
57. Langmead, B. and S.L. Salzberg, *Fast gapped-read alignment with Bowtie 2*. *Nature Methods*, 2012. **9**(4): p. 357-359.
58. Quinlan, A.R. and I.M. Hall, *BEDTools: a flexible suite of utilities for comparing genomic features*. *Bioinformatics*, 2010. **26**(6): p. 841-842.
59. Derbyshire, M., et al., *Small RNAs from the plant pathogenic fungus *Sclerotinia sclerotiorum* highlight candidate host target genes associated with quantitative disease resistance*. *Molecular Plant Pathology*, 2019. **20**(9):p. 1279-1297
60. Mohorianu, I., et al., *The UEA small RNA workbench: a suite of computational tools for small RNA analysis*, in *MicroRNA Detection and Target Identification*. 2017, Springer. p. 193-224.
61. Lorenz, R., et al., *SHAPE directed RNA folding*. *Bioinformatics*, 2016. **32**(1): p. 145-147.
62. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2*. *Genome Biology*, 2014. **15**(12): p. 550.
63. Benjamini, Y. and Y. Hochberg, *Controlling the false discovery rate: a practical and powerful approach to multiple testing*. *Journal of the Royal statistical society: Series B (Methodological)*, 1995. **57**(1): p. 289-300.
64. Dai, X. and P.X. Zhao, *psRNATarget: a plant small RNA target analysis server*. *Nucleic Acids Research*, 2011. **39**(suppl_2): p. W155-W159.

65. Chen, H.-M., Y.-H. Li, and S.-H. Wu, *Bioinformatic prediction and experimental validation of a microRNA-directed tandem trans-acting siRNA cascade in Arabidopsis*. Proceedings of the National Academy of Sciences, 2007. **104**(9): p. 3318-3323.
66. Mueth, N.A., S.R. Ramachandran, and S.H. Hulbert, *Small RNAs from the wheat stripe rust fungus (Puccinia striiformis f. sp. tritici)*. BMC Genomics, 2015. **16**(1): p. 718.
67. Mi, S., et al., *Sorting of small RNAs into Arabidopsis argonaute complexes is directed by the 5' terminal nucleotide*. Cell, 2008. **133**(1): p. 116-127.
68. Lee Marzano, S.-Y., A. Neupane, and L. Domier, *Transcriptional and small RNA responses of the white mold fungus Sclerotinia sclerotiorum to infection by a virulence-attenuating hypovirus*. Viruses, 2018. **10**(12): p. 713.
69. Omidvar, V., et al., *Identification of miRNAs with potential roles in regulation of anther development and male-sterility in 7B-1 male-sterile tomato mutant*. BMC Genomics, 2015. **16**(1): p. 1-16.
70. Marzano, S.Y.L., A. Neupane, and L. Domier, *Transcriptional and small RNA responses of the white mold fungus sclerotinia sclerotiorum to infection by a virulence-attenuating hypovirus*. Viruses, 2018. **10**(12).
71. Kozomara, A., M. Birgaoanu, and S. Griffiths-Jones, *miRBase: from microRNA sequences to function*. Nucleic Acids Research, 2019. **47**(D1): p. D155-D162.
72. Bai, Y., et al., *Genome-wide sequencing of small RNAs reveals a tissue-specific loss of conserved microRNA families in Echinococcus granulosus*. BMC Genomics, 2014. **15**(1): p. 1-13.

73. Meunier, J., et al., *Birth and expression evolution of mammalian microRNA genes*. Genome Research, 2013. **23**(1): p. 34-45.
74. Eshel, O., et al., *Identification of male-specific amh duplication, sexually differentially expressed genes and microRNAs at early embryonic development of Nile tilapia (*Oreochromis niloticus*)*. BMC Genomics, 2014. **15**(1): p. 1-18.
75. Zhai, J., et al., *MicroRNAs as master regulators of the plant NB-LRR defense gene family via the production of phased, trans-acting siRNAs*. Genes & Development, 2011. **25**(23): p. 2540-2553.
76. Lau, S.K., et al., *Identification of microRNA-like RNAs in mycelial and yeast phases of the thermal dimorphic fungus *Penicillium marneffei**. PLoS Neglected Tropical Diseases, 2013. **7**(8): p. e2398.
77. Jin, Y., et al., *A fungal miRNA mediates epigenetic repression of a virulence gene in *Verticillium dahliae**. Philosophical Transactions of the Royal Society B, 2019. **374**(1767): p. 20180309.
78. Aravin, A.A., et al., *Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the *D. melanogaster* germline*. Current Biology, 2001. **11**(13): p. 1017-1027.
79. Fawke, S., et al., *Glycerol-3-phosphate acyltransferase 6 controls filamentous pathogen interactions and cell wall properties of the tomato and *Nicotiana benthamiana* leaf epidermis*. New Phytologist, 2019. **223**(3): p. 1547-1559.
80. Seifbarghi, S., et al., *Changes in the *Sclerotinia sclerotiorum* transcriptome during infection of *Brassica napus**. BMC Genomics, 2017. **18**(1): p. 1-37.

81. Kelly, S.M., et al., *Recognition of polyadenosine RNA by zinc finger proteins*. Proceedings of the National Academy of Sciences, 2007. **104**(30): p. 12306-12311.
82. Brockmann, C., et al., *Structural basis for polyadenosine-RNA binding by Nab2 Zn fingers and its function in mRNA nuclear export*. Structure, 2012. **20**(6): p. 1007-1018.
83. Gioti, A., et al., *A Botrytis cinerea emopamil binding domain protein, required for full virulence, belongs to a eukaryotic superfamily which has expanded in euascomycetes*. Eukaryotic Cell, 2008. **7**(2): p. 368-378.
84. Bravo-Plaza, I., et al., *Identification of the guanine nucleotide exchange factor for SARI in the filamentous fungal model Aspergillus nidulans*. Biochimica et Biophysica Acta (BBA)-Molecular Cell Research, 2019. **1866**(12): p. 118551.
85. Li, Y., et al., *FgMon1, a guanine nucleotide exchange factor of FgRab7, is important for vacuole fusion, autophagy and plant infection in Fusarium graminearum*. Scientific Reports, 2015. **5**(1): p. 1-13.
86. Woese, C.R., *Interpreting the universal phylogenetic tree*. Proceedings of the National Academy of Sciences, 2000. **97**(15): p. 8392-8396.
87. Pasquinelli, A.E., et al., *Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA*. Nature, 2000. **408**(6808): p. 86-89.
88. Floyd, S.K. and J.L. Bowman, *Ancient microRNA target sequences in plants*. Nature, 2004. **428**(6982): p. 485-486.

89. Hardcastle, T.J., S.Y. Müller, and D.C. Baulcombe, *Towards annotating the plant epigenome: the Arabidopsis thaliana small RNA locus map*. Scientific Reports, 2018. **8**(1): p. 6338.
90. Wu, F., et al., *Genome-wide identification and characterization of phased small interfering RNA genes in response to Botrytis cinerea infection in Solanum lycopersicum*. Scientific Reports, 2017. **7**(1): p. 1-10.

Chapter 4: Small RNA profiling in *Sclerotinia sclerotiorum* and target identification in *Brassica napus*

Abstract

Fungal plant pathogens secrete proteins and metabolites known as 'effectors' to suppress host immunity. Recently, some fungi were also found to secrete molecules known as small RNAs (sRNAs) that act similarly to protein and metabolite effectors. These short, non-coding RNAs regulate gene expression via a mechanism called RNA interference (RNAi), whereby complementary mRNAs are targeted for degradation. The fungus *Sclerotinia sclerotiorum* causes disease in more than 400 plant species, including the important crop species *Brassica napus*. However, little is known about the molecular mechanisms at play at the sRNA level during host infection. Recent analyses have shown that *S. sclerotiorum* may produce sRNAs with the potential to silence immunity genes in the model plant *Arabidopsis thaliana*. Using a high throughput sequencing technique, we identified several fungal sRNAs that potentially target plant immune response genes in *B. napus*. Targets of these sRNAs are enriched for plant immune response genes suggesting a role of fungal sRNA effectors during *B. napus* infection. We experimentally analysed the silencing of host genes through pathogen sRNAs using degradome sequencing, 5' RACE, and quantitative polymerase chain reaction.

4.1 Introduction

Small RNAs (sRNAs) are non-coding RNA molecules with a size ranging from 18-30 nucleotides (nt) and they have significant roles in plant growth and development [1]. sRNAs regulate gene expression by silencing target transcripts via sequence complementarity [2], a process commonly called RNA interference (RNAi). RNAi is mediated via a discrete set of proteins. During this process, the precursor double stranded RNAs (dsRNAs) are cleaved by

an enzyme called Dicer into sRNA duplexes of approximately 18-30 nt. The sRNA duplex is subsequently loaded into an RNA-induced silencing complex (RISC) in which the protein Argonaute is the major component which functions as a sRNA-guided endonuclease. One of the strands of the sRNA duplex guides the RISC to complementary mRNA transcripts via base-pairing. RNA dependent polymerase (RdRp) generates dsRNA from single-stranded transcripts from primary transcripts or by using sRNAs as primers to synthesize RNA complementary to the target mRNA [3]. Target mRNAs are commonly cleaved, but may also have their translation inhibited [4]. Also, target chromosomal gene sequences may be methylated by the action of sRNAs. All these processes effectively silence gene expression.

Among the sRNAs, there are two major groups, the microRNAs (miRNAs) and the short interfering RNAs (siRNAs). The miRNAs and siRNAs differ in terms of their precursors, with the former derived from hairpin-shaped single-stranded RNAs and the latter produced from double-stranded RNAs [4]. The endogenous roles of many sRNAs have been well investigated and characterized in plants. However, recent research has shown that sRNAs derived from pathogens may also enter plant cells to reduce gene expression via RNAi. So far, this phenomenon is described in only a handful of species including the broad host range pathogen *Botrytis cinerea* (a relative of *S. sclerotiorum*, which is also in the family Sclerotiniaceae) [5, 6], and other pathogens like *Puccinia striiformis* [7], *Fusarium graminearum* [8] and *F. oxysporum* [9]. In contrast to these examples, one study has shown a distinct lack of evidence for cross-kingdom RNAi between the fungus *Zymoseptoria tritici* and its host wheat. This indicates that the presence/absence of cross-kingdom RNAi may be variable between fungal species [10].

In *B. cinerea*, several sRNAs are known that can manipulate host plant defence responses in *A. thaliana* [6]. In *S. sclerotiorum*, a recent study has shown evidence of pathogen sRNA accumulation in host plants and potential targeting of immunity genes in *A. thaliana* [11]. Derbyshire et al. [11] found 374 highly abundant sRNAs in *S. sclerotiorum*, which mostly originated from repeat-rich plastic genomic regions. Predicted targets of these sRNAs in *A. thaliana* were enriched for plant immune response domains and had a statistical association with disease resistance from GWAS data, suggesting cross-kingdom interference through sRNAs from *S. sclerotiorum* might be utilised to colonise its host [11]. This leads to the opportunity for further investigation of the potential role of *S. sclerotiorum* sRNAs to manipulate the plant immune system in hosts other than *A. thaliana* such as canola (*B. napus*).

Despite these observations, RNAi from *S. sclerotiorum* to host plants is not well understood. In particular, it is not known whether *S. sclerotiorum* sRNAs are differentially expressed at different time points during infection. Therefore, we performed sRNA sequencing of an infection time course of either *S. sclerotiorum* mycelial mats on detached *B. napus* leaves or agar plugs on attached *B. napus* leaves. We combined differential expression profiling and target prediction to identify several candidate *S. sclerotiorum* sRNAs that are 1) significantly expressed during infection and 2) have potential host target genes with functional domains linked with plant immunity. Moreover, previous characterization of *S. sclerotiorum* effectors by Derbyshire et al [11] was based on *in silico* target prediction, whereas here we developed a high throughput method of target validation called degradome sequencing which is experimental evidence for sRNA-mediated cleavage.

4.2 Materials and Methods

4.2.1 Plant material and inoculation

Two different methods of sample collection were adopted for total RNA extraction: detached and attached leaf infection assays. Detached leaves were inoculated with fungal mycelial mats and attached leaves were inoculated with agar plugs. The reason we included the mycelial mat inoculation procedure was to maximise recovery of fungal tissue [12]. In addition, the *in-planta* infection assay was performed to understand the enrichment of sRNAs during the infection process. For the detached leaf assay, *B. napus* (cv AV Garnet) plants were grown for six weeks in a growth chamber with 16 h of light and 8 h of darkness. First leaves were excised from each of the plants and put into 9 cm Petri dishes containing filter paper soaked in sterile distilled water. The *S. sclerotiorum* isolate CU8.24 originally collected from South Stirling, WA was used for infection assays [13]. Fully mature sclerotia were cut into halves and placed into 9 cm Petri dishes containing potato dextrose agar (PDA). After germination from the sclerotium, mycelium was subcultured onto fresh PDA medium. After 48 hours of incubation, a 1 cm mycelial plug was used to inoculate liquid minimal medium containing 2 g/L NH_4NO_3 , 1 g/L KH_2PO_4 , 0.1 g/L $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 0.5 g/L yeast extract, 4 g/L DL-malic acid and 1 g/L NaOH [14]. Once the mycelial mat was well spread over the minimal medium plates, approximately 200 mg of mycelium was used to infect the detached leaves and they were incubated at 20°C after inoculation. After 12 hours and 24 hours, intact mycelia were carefully separated from the leaves, immediately frozen in liquid nitrogen and stored at -80 °C until RNA extraction. For attached leaf infection assays, plants were grown as mentioned above; however, the mycelial plug was directly placed onto leaves of six-week-old *B. napus* plants [15]. After infection, leaves were carefully covered with a polythene bag to increase humidity. Mock

inoculated (empty agar plugs) leaf samples and the infected area of inoculated leaves were collected 24 hours post inoculation (HPI) for small RNA and degradome sequencing analysis.

4.2.2 RNA extraction and construction of sRNA and degradome sequencing libraries

Total RNAs were extracted using the TRIzol™ (Thermo Fisher, Sydney, NSW, Australia) reagent following the manufacturer's protocol. After extraction, total RNAs were quantified using a NanoDrop spectrophotometer (Thermo Scientific) and Qubit (Invitrogen); the integrity of RNA samples was checked using agarose gel electrophoresis. Fourteen libraries of sRNAs were prepared from the total RNAs from three replicates each for *in vitro*, 24 HPI detached leaf assay, 24 HPI mock, 24 HPI infected samples and duplicates of the 12 HPI detached leaf assay. Libraries of detached leaf assays at 12 HPI and 24 HPI along with attached leaf samples (hereafter referred to as *in planta* samples) were used to identify abundant sRNAs while all the downstream analyses like qPCR, 5' RACE and degradome analysis were done comparing *in vitro*, mock and 24 HPI *in planta* samples. The sRNA sequencing was done using the TruSeq small RNA sample preparation kit (Illumina) according to the manufacturer's protocol. Single-end (50 bp) sequencing was performed on an Illumina HiSeq2500 by Novogene (Beijing, China).

To find putative cleavage sites, we performed degradome sequencing from mock and infected *in planta* samples at 24 HPI. Samples for degradome libraries were the same as those that were used for *in planta* small RNA sequencing. Three replicates each containing three leaf samples from different plants were pooled together and sequenced using the Illumina platform. Degradome sequencing was done as mentioned in [16]. In brief, construction of the degradome library was started from the degradation site (with a 5' monophosphate group) of the degraded mRNA. The sequencing adaptors were added to both ends of the degraded mRNAs and a

library with an insert size of around 200-400 (base pairs) bp was generated. Paired end sequencing (2 x 150bp) was performed on a Hiseq 2500 (Illumina). The read length was shorter than the insert fragment length. Therefore, the sequencing data of the degradome library did not contain adapters, and they were all the same length. The length of the original raw data was 50bp.

4.2.3 Filtering of reads and assignment for mapping

Raw reads of the sRNA sequencing data were trimmed using the Cutadapt v1.15 optimized for single-end reads with a setting of cutadapt -a (universal TruSeq adapter) -m18 -M30. The quality of filtered reads was checked by running in FastQC [17]. Reads with a length in the range of 18-30 nt were retained. Trimmed reads were assigned to the fungal and plant genome using the bbmap program v33 with the option ‘ambig2’ set as ‘toss’ [18]. With this option, reads that were ambiguous and mapped to both the fungal and plant genome with equal confidence were removed, keeping only fungal and plant unique reads for the prediction of fungal and plant sRNAs, respectively. A similar technique was applied to filter the degradome reads.

4.2.4 Prediction of fungal sRNAs

The adapter trimmed clean reads that potentially originate from structural RNAs were filtered out using the Rfam database [19] with the program Infernal v1.1.3, with the command ‘cmscan --nohmmonly --rfam --cut_ga --fmt 2 --oclan --oskip -o strRNA.out --tblout strRNA.tblout structuralRNA.cm CU8_24.fasta’ [20]. First, Rfam hits in the fungal genome were prepared from the Rfam clanin file (cmfetch Rfam.cm). sRNA libraries were then mapped to the Rfam hits in the fungal genome using Bowtie2 [21].

Three different software packages were used, ShortStack v3.8.5 [22], miRCAT v2 [23], and miRDeep2 v2.0.0.8 [24] for prediction of sRNA/miRNA loci. For ShortStack, all clean individual sRNA libraries without collapsing were aligned with the genome of the *S. sclerotiorum* strain CU8.24 [25] with the package Bowtie v2 [21], via the package ShortStack [22] allowing 0 mismatches keeping miconv 0.5 rpm. ShortStack first identifies significant alignment coverage based on the depth of reads. The significant alignments are then padded upstream and downstream. Overlapping regions are thereafter merged to form clusters. Loci were annotated as Dicer-derived at a minimum threshold of 80% of mapped reads within the range of 18-30 nt which is suggested to improve specificity of sRNA locus identification. The most abundant sRNA reads of a particular size class in the Dicer-derived loci were considered their major sRNAs. Size class distribution and 5' nucleotide bias of sRNA reads were considered to find loci producing sRNAs [22]. Loci with read counts of at least 100 RPM were retained and major sRNAs belonging to these loci were used for target analysis. The positions of highly expressed sRNA loci relative to genomic features was carried out using Bedtools v2.29.0 [26]. The genomic features include gene sense, gene antisense, repeat sense, and repeat antisense. The generic and repeat elements in *S. sclerotiorum* CU8.24 were previously annotated by Derbyshire et al. [25].

For miRDeep2, each library was first collapsed using the collapse.pl script that comes with miRDeep2. Then, a config.txt file was created with unique three letter IDs assigned to each individual collapsed library. Then, we mapped all the individual collapsed reads against the reference genome using Bowtie allowing no mismatches with the command: *mapper.pl config.txt -d -c -i -j -l 18 -m -p genome_index -s reads.fa -t reads_vs_genome.arf*. The output reads. fa and reads_vs_genome.arf were passed to the miRDeep2.pl script to identify the miRNA loci. The non-redundant miRNAs with at least 10 reads across all libraries were

retained and blasted against miRBase using the program miRProf. For miRCat, we extracted the mapped sRNA reads from binary alignment map (BAM) formatted files resulting from the Bowtie alignment and used these in conjunction with the reference genome keeping all the default parameters.

4.2.5 Normalization and differential expression analysis

We used DeSeq2 v1.22.1 to find differentially expressed sRNA loci [27]. Each sRNA from the set of highly expressed 156 loci from ShortStack was considered as a single entry with a raw read count from the ShortStack result. All *in planta* samples were compared to *in vitro* samples using the negative binomial test. Differentially expressed loci were defined as those with a p-value of < 0.05 after correction for false discovery rate using the Benjamini-Hochberg adjustment [28].

4.2.6 Prediction of fungal sRNA target genes in plant hosts using the psRNA target server

Target prediction of the highly expressed sRNAs and miRNAs was done using the psRNATarget (version 2017 release) server [29]. To find putative targets we used the server option '*B. napus* transcript_v5' maintaining all the default parameters except decreasing the expectation score from ≤ 5 to ≤ 3 . The psRNATarget server finds target sequences of sRNAs based on complementary matches according to a scoring scheme and calculates unpaired energy [29]. The unpaired energy determines the likelihood of cleavage of the sRNA targets.

4.2.7 Quantification of target genes using quantitative polymerase chain reaction

The expression levels of plant target genes potentially cleaved by upregulated sRNAs in plant samples were analysed by qPCR. Three μg of total RNAs from mock and infected *B. napus* leaf samples were converted to cDNA using the MMV reverse transcriptase kit. The cDNA samples were then diluted 1/20 before qPCR. The qPCR analysis was performed using the Bio-Rad Taq Universal SYBR Green Supermix according to the manufacturer's instructions. The thermocycler settings were 95 °C for 2 min, then 95 °C for 15 sec, 60 °C for 30 sec and 72 °C for 15 sec, and cycled 40 times, followed by 72 °C for 2 min. Three biological and three technical replicates were used for each transcript tested. The primer sequences for the target genes are listed in Supplementary Table 4.1. Expression levels were calculated using the ΔCt method [30], where the target gene was normalised to the *B. napus* housekeeping gene actin (*BnaC02g00690D*). Differential gene expression analysis between mock and treated samples was conducted using student's t-test (P value < 0.05) using the statistical software package in R 3.6.1.

4.2.8 5' RACE validation of target genes

5'RACE was conducted using the first choice RACE kit following the manufacturer's protocol (Applied Biosystems, Carlsbad, CA, USA) without adding calf intestinal Phosphatase enzyme. In brief, 5' RACE adapters were ligated to 5 μg of total RNA, which was reverse transcribed using the universal RT primer provided in the kit and the MMLV transcriptase. The first PCR was conducted on 1 μL of cDNA with a 5' outer RACE primer and gene-specific outer primer. The second nested PCR was done using the first PCR product with a 5' inner RACE primer and inner nested PCR primers. The PCR product was visualized on a 2% Agarose gel. The

amplified DNA fragment was gel purified and cloned into TOP TA vector and 5 independent clones were Sanger sequenced.

4.2.9 Degradome sequencing and analysis

The methods for degradome sequencing is detailed in Chapter 3 (3.2.6). We used PAREsnip2 [31] and CleaveLand 4 [32] for the analysis of degradome datasets. PAREsnip2 was used when a large set of sRNAs were used as input while CleaveLand 4 was used when a smaller set of sRNAs were used as input. We also compared the common targets identified from both programs. Clean reads that potentially originated from structural RNAs (rRNA, tRNA, snRNA, snoRNA) were removed by the method mentioned above for sRNA analysis. Only cleavage tags with a *P*-value under 0.05 and a minimum free energy (MFE) ratio higher than 0.7 and category ≤ 3 were considered significant.

4.2.10 Gene ontology enrichment analysis of overrepresented target gene domains

Gene Ontology (GO) term enrichment analysis was performed with TopGO v2.34.0 [33]. The background comprised all the genes annotated to a GO term in the entire set, while sampling frequency is the number of genes annotated to that GO term which were predicted to be the targets of fungal small RNAs in plants.

4.3 Results

4.3.1 Overview of small RNA sequencing from the fungal pathogen *Sclerotinia*

sclerotiorum

In addition to the mycelial mat samples (Chapter 3), this chapter further enhances the prediction of fungal sRNAs with *in planta* sRNA samples at 24 HPI. To explore the role of *S. sclerotiorum* sRNAs during *B. napus* infection we generated sRNA sequencing libraries from fungal

mycelium grown on detached leaves at 0, 12, and 24 HPI and samples of infected attached leaves at 24 HPI. The 0 HPI samples were *in vitro* mycelial mats and were used as controls to compare the expression levels of sRNAs to *in planta* samples. Approximately 6 billion raw sRNA reads were obtained from the 14 libraries. The summary about the sRNA reads has been given in Chapter 3 (Fig 3.1) In brief, the most frequent read length was 22 nt followed by 23 nt. We did not find any difference in size distribution between *in vitro* and infected samples. All highly abundant sRNAs in *S. sclerotiorum* had a bias toward uridine as the 5' nucleotide with peak size of 23 and 24 nt (Fig. 4.1).

Similarly, previous studies reported that *S. sclerotiorum* exhibits a characteristic sRNA length distribution with a peak in abundance of reads at 22-23 nucleotides and a bias toward 5' uridine [10, 34]. This characteristic sRNA size distribution and sequence composition may be attributed to the sorting of sRNAs into particular Argonaute proteins [35]. The frequency distribution of 5' nucleotide biases might, therefore, be an initial identification for potential sRNA candidates given their loading preferences for Argonaute proteins [36]. The clean sRNA reads were aligned to the fungal genome using Bowtie with 0 mismatches. On average, 70% of reads were mapped to the genome ranging from 66-80% depending on the sample.

4.3.2 Prediction of fungal small RNAs

ShortStack analysis yielded 12,824 putative clusters of sRNAs. Among these, 6,554 loci were classed as Dicer-derived with 6,008 non-redundant major RNA sequences. Altogether, 288 highly expressed fungal sRNA loci (≥ 100 reads per million) were identified from the fungal genome. We used Bedtools to identify the possible origins of these loci using existing CU8.24 gene and repeat annotations. Interestingly, 70% of these highly expressed loci overlapped with transposable elements (Fig. 4.2). In total, 156 non-redundant major sRNAs were found in these

loci, which were abundant in *in planta* samples (with at least 100 RPM of each locus). From our dataset we did not find any miRNA loci from the ShortStack program, while 680 (Supplementary file 4.1) and 80 miRNAs (Supplementary File 4.2) were predicted using miRDeep2 and miRCat with at least 10 RPM, among which 83 sRNA loci overlapped with the loci identified using ShortStack (Supplementary Table 4.2). Such variability in prediction of miRNAs has been reported in previous studies and it may be because the three programs we used were developed on plant and animal but not fungal data. It could also be due to differences in the search algorithms between the software, which have different advantages and disadvantages for identifying miRNAs [34]. Therefore, we used three different programs to favour sensitivity.

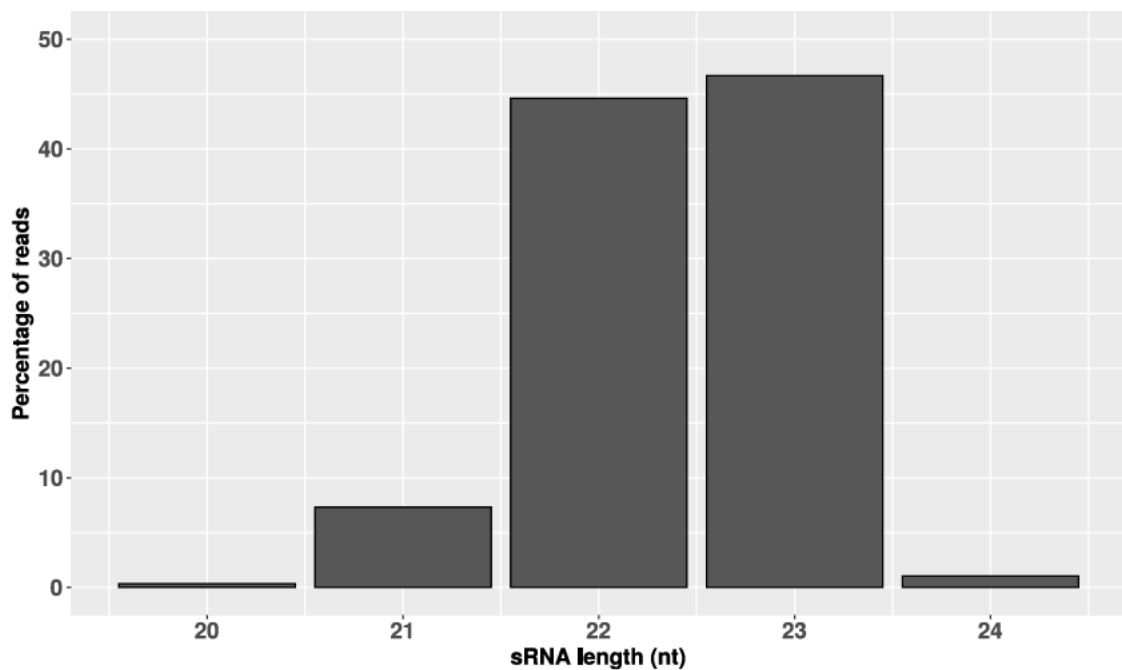


Fig. 4. 1. Characteristics of *Sclerotinia sclerotiorum* small RNAs. Histogram of read size of 288 highly expressed sRNAs from pooled mock and infected samples.

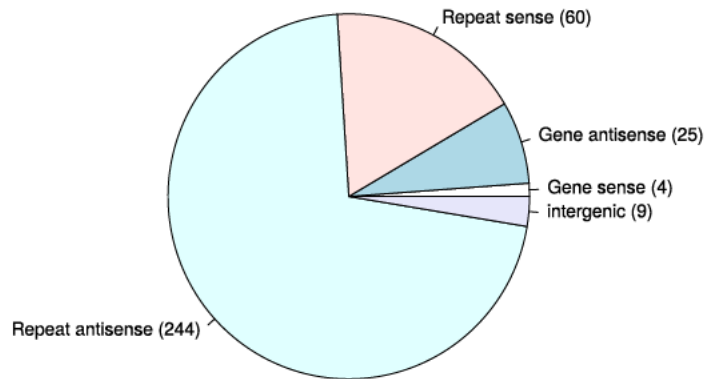


Fig .4. 2. A pie chart showing the origin of 288 sRNA loci in the *Sclerotinia sclerotiorum* isolate CU8.24 genome.

4.3.3 Identification of miRNA-like loci in *Sclerotinia sclerotiorum*

The identification of miRNAs is governed by the ability of their surrounding sequences to form hairpin secondary structures [37]. There is no systematic database for the deposition of fungal miRNAs and it is not clear whether they are evolutionarily related to the plant and animal miRNAs; therefore, fungal sRNA loci that form hairpins are usually called miRNA-like RNAs (milRNAs) rather than miRNAs. Here, we used two different software packages for the identification of milRNA loci in *S. sclerotiorum*. Altogether, 172 putative milRNA loci were predicted using miRCat. Eighty loci with non-redundant sequences had an abundance of at least 10 RPM. Interestingly, four loci, Scaffold_19:346184-346203, Scaffold_206:1453-1474, Scaffold_32:79430-79451, and Scaffold_337:10265-10286 were found to harbour miRNA* sequences. The miRNA* sequence originates from the other side of the hairpin, which normally remains inactive and is degraded during the silencing process [38]. The program miRDeep2 predicted 680 milRNA loci with more than 10 RPM. We were interested to see whether these milRNAs had close homologs in other organisms based on the sRNA sequences stored in

miRBase. Interestingly, 2 miRNA families, miR-423 and miR-7178, were found with the sequences TGAGGGGCAGAGAGCGAG and TAATTTCTGATGCTCTTC, with 0 mismatches. Interestingly, the former miRNA was previously reported to be present in various vertebrates including humans [39] and mice [40] and has 23 isoform entries in miRBase, and the latter one was reported to be present in mammals [41].

We also found homologs with mismatches of three or fewer base pairs in 45 miRNA families. A total of 27 sequences were annotated as novel miRNAs as per the annotation criteria, and these are thereafter described as novel miRNAs (Fig. 4.3).

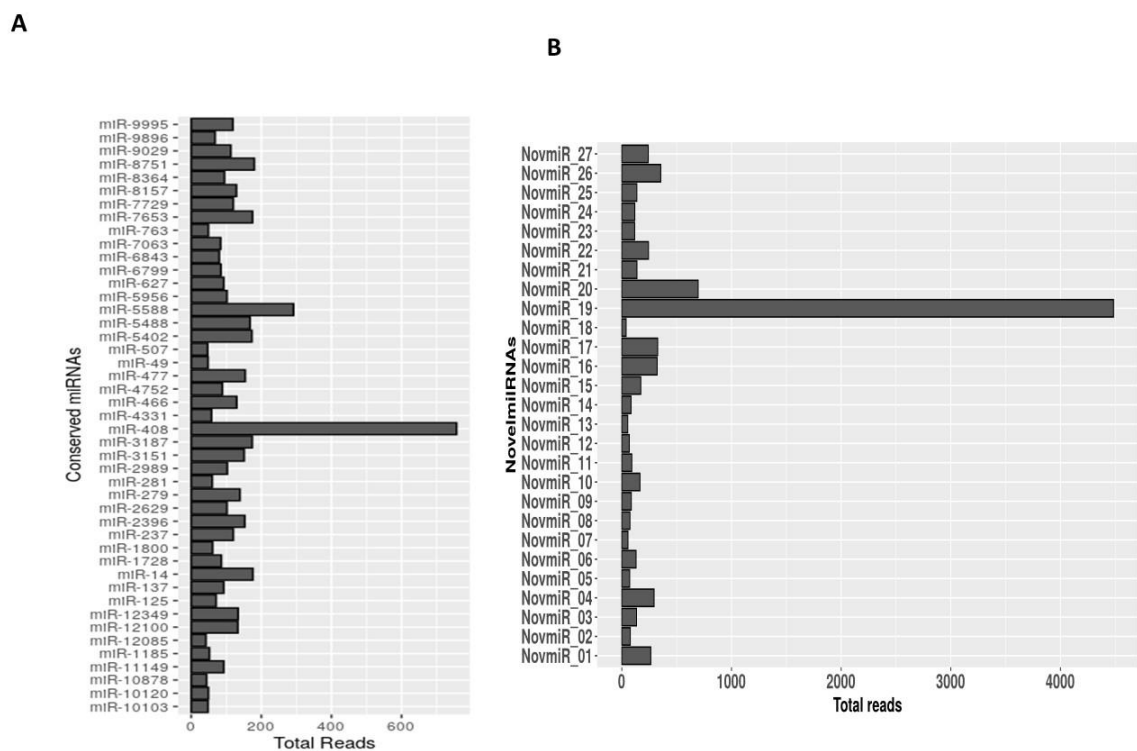


Fig .4.3. Prediction of *Sclerotinia sclerotiorum* miRNA-like loci with at least 100 reads per million (RPM). (A) Bar graph showing miRNA families that had close homologues in the miRBase database with three or fewer mismatches. (B) Novel miRNA loci with no homologues in miRbase.

4.3.4 Small RNAs are differentially expressed in *Sclerotinia sclerotiorum* mycelium during host infection

To identify differentially expressed sRNAs we used raw read counts of the 156 highly expressed major sRNA loci from ShortStack. Among these loci, 63 were upregulated after *S. sclerotiorum* infection at 24 HPI. Similarly, 67 loci were downregulated, while 26 were not differentially expressed before and after inoculation. A heatmap of upregulated sRNA loci during infection of *B. napus* is shown in Fig. 4.4.

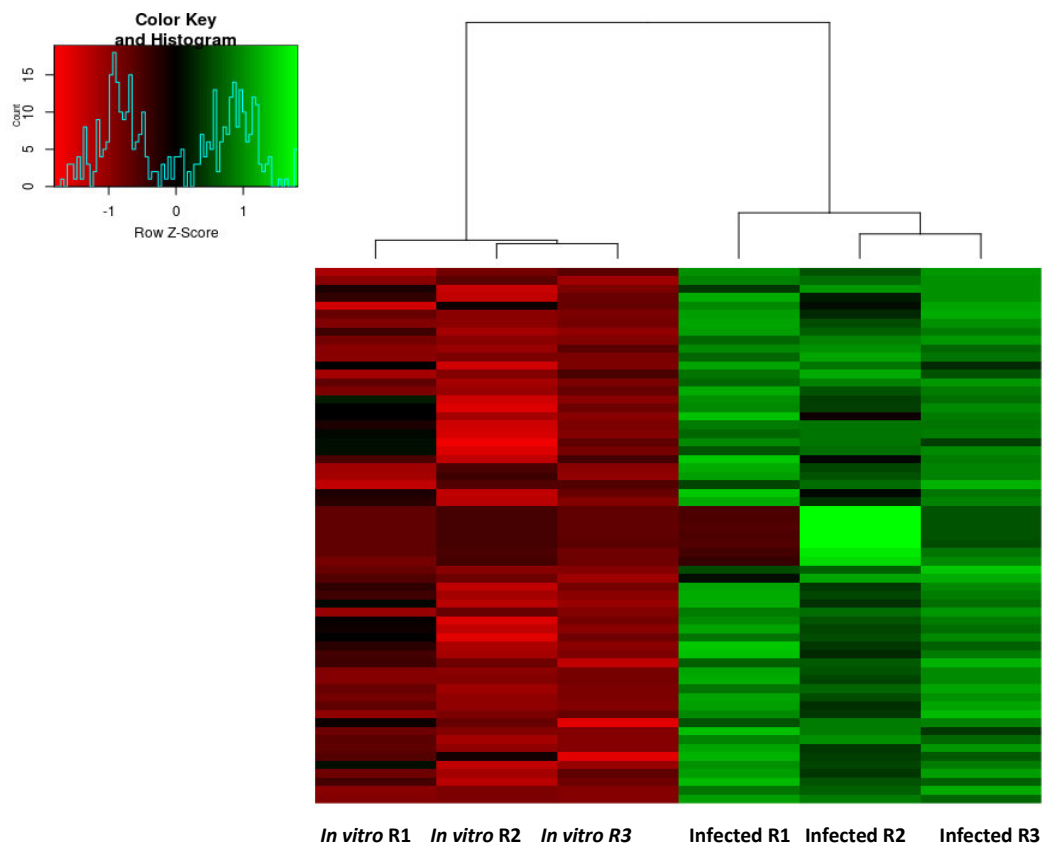


Fig .4. 4. Heatmap showing 63 upregulated sRNA loci of *Sclerotinia sclerotiorum* isolate CU8.24 during infection of *Brassica napus* leaves (24 HPI) relative to *in vitro* controls. The three biological replicates for *in vitro* and infected samples are plotted with normalized counts from DESeq2.

4.3.5 Small RNA target genes in the host are enriched with plant immune response domains

Cross-kingdom RNAi has been recently reported in several pathosystems. To investigate the presence of cross-kingdom RNAi in the *S. sclerotiorum* and *B. napus* pathosystem, the most abundant 156 sRNAs were used to find target genes in the host using the psRNA target server. Altogether, 2,163 targets were predicted from these fungal sRNAs in *B. napus*. The detail about these targets were given in Supplementary Table 4.3. We performed GO term enrichment analysis of target genes. The terms xylan biosynthetic process (GO:004592), regulation of transcription (GO:0006355), signal transduction (GO:0007165), photosynthesis (GO:0015979), response to abscisic acid (GO:0009737), and regulation of ARF protein signals (GO:0032011) were among the most interesting biological processes enriched in these target genes (Fig. 4.5). Based on the results we chose our top four candidates, which were *BnaC06g14210D* (protein kinase), *BnaC01g35070D* (AP2/ERF), *BnaC08G20040D* (MORN motif), *BnaA02g11730D* (plastocyanin-like). Among these four candidate genes that were putatively cleaved by fungal sRNAs, two of them having domains AP2/ERF and protein kinase, that have been shown to be involved in plant immune response [42, 43]. However, the MORN motif and plastocyanin are indirectly linked to plant response to biotic stress like pathogen attack. A MORN motif is found in phosphatidylinositol monophosphate kinase (PIP2K) which is a key enzyme for PI- signalling pathway [44]. The PI-signalling pathway has been shown to be one of the components of the plant response to pathogen infection [45]. The plastocyanin-like domain is mainly involved in regulating leaf chlorophyll content and photosynthesis in plants [46]. Leaf colour has been shown as an indication of plant response during biotic stress [47].

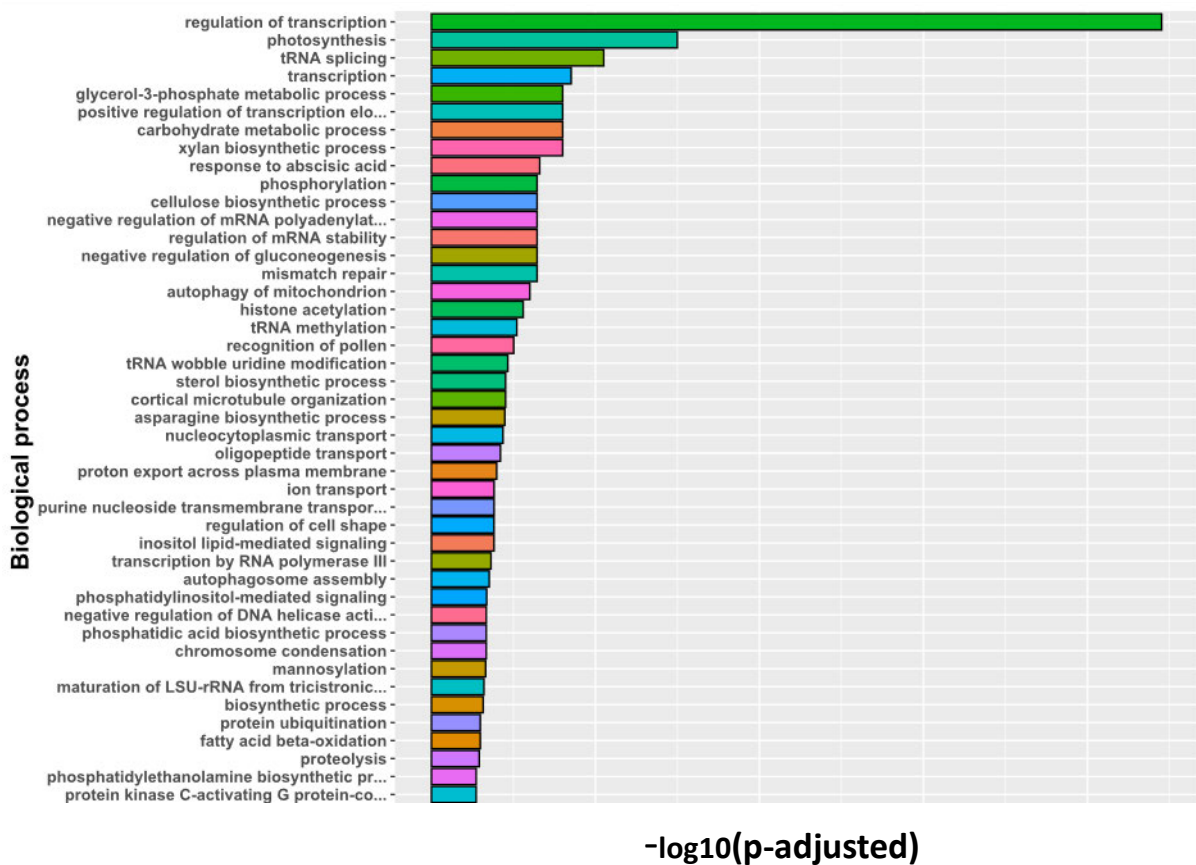


Fig .4.5. GO term enrichment of 156 highly expressed *Sclerotinia sclerotiorum* small RNA targets in *Brassica napus* obtained from the psRNA target server. The y-axis shows the biological process, while x-axis shows the $-\log_{10}(\text{p adjusted})$ values.

4.3.6 qPCR shows that fungal sRNA target genes are repressed during infection

Reverse transcriptase qPCR was performed to quantify transcript levels of *BnaA01g35070D* (AP2/ERF), *BnaC06g14210D* (protein kinase) and *BnaC08g20040D* (MORN motif) in mock and 24 HPI infected samples. We could not detect the expression of *BnaA02g11730D* (Plastocyanin like) through qPCR. We found that transcript levels of predicted target genes evaluated here at 24 HPI compared with the mock sample (uninoculated attach leaves) were reduced. The AP2/ERF, protein kinase, and MORN motif genes had mean $2^{-\Delta\text{Ct}}$ of 4.5 and 0.85, 0.05 and 0.01, 6 and 1.79, for mock and infected samples, respectively (Fig. 4.6).

Although *BnaC01g35070D* (AP2/ERF) and *BnaC08g20040D* (MORN motif) are repressed, they are not significantly decreased compared to the mock (p=0.7 and p= 0.8 respectively). However, the gene *BnaC06g14210D* (protein kinase) expression was significantly decreased in the infected sample compared to the mock (p= 0.03). We infer from this that reduced transcript levels of these genes may be due to gene silencing by fungal sRNAs.

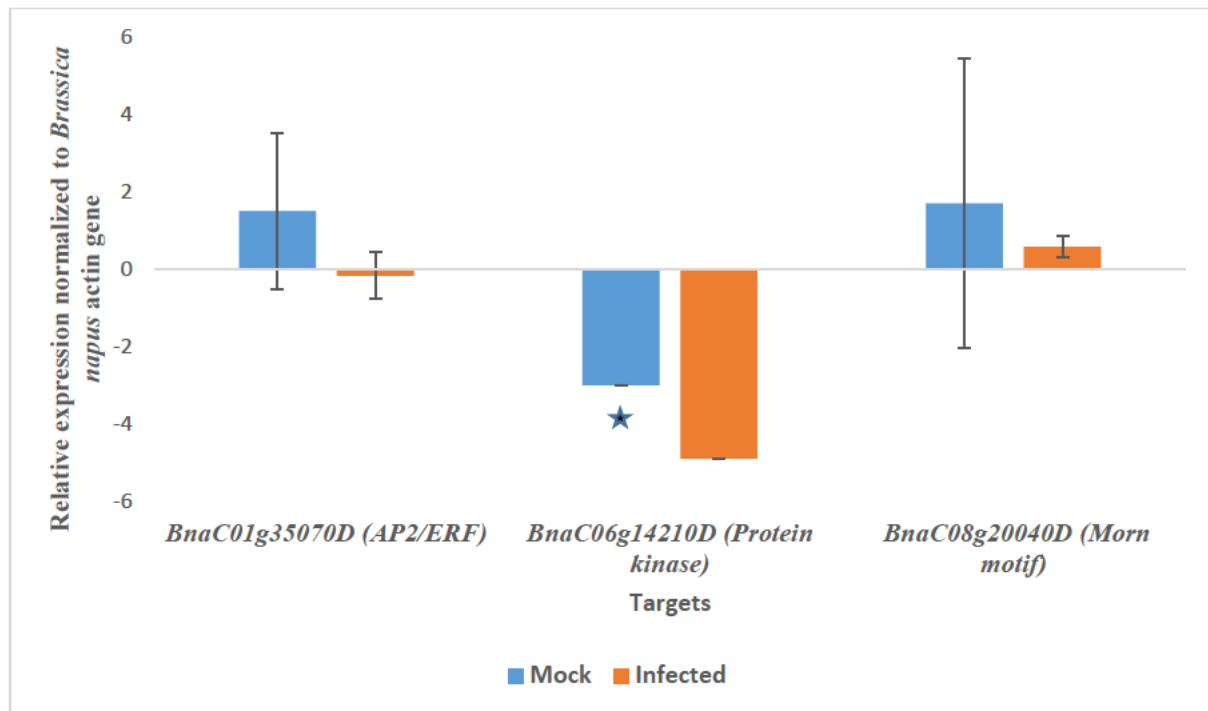


Fig. 4.6. Expression analysis of fungal sRNA target genes in *Brassica napus*. The x-axis displays gene ID and domains found in the genes of interest, *BnaA01g35070D* (AP2/ERF), *BnaC06g14210D* (protein kinase) and *BnaC08g20040D* (MORN motif). The y-axis represents the relative expression levels mean $\log(2^{-\Delta Ct})$ normalized to the *B. napus* actin gene. The error bar shows the standard error values from three biological replicates.

4.3.7 5' RACE reveals cleavage of putative target genes

To confirm sRNA-mediated cleavage of these target genes, 5'RACE experiments were performed based on the highly confident targets from psRNA target server. The 5'RACE products were visualized using gel electrophoresis (Supplementary Fig 4.1). Fig. 4.7 shows the

mapping of the cleavage site for four *S. sclerotiorum* sRNAs using 5'RACEA. The expected amplicon sizes of genes *BnaC06g14210D* (protein kinase), *BnaC01g35070D* (AP2/ERF), *BnaC08g20040D* (MORN motif) and *BnaA02g11730* (plastocyanin-like), were 261, 407, 251, and 401 bp respectively. The cleavage site from 5' RACE matched to the predicted cleavage site from psRNA target server.

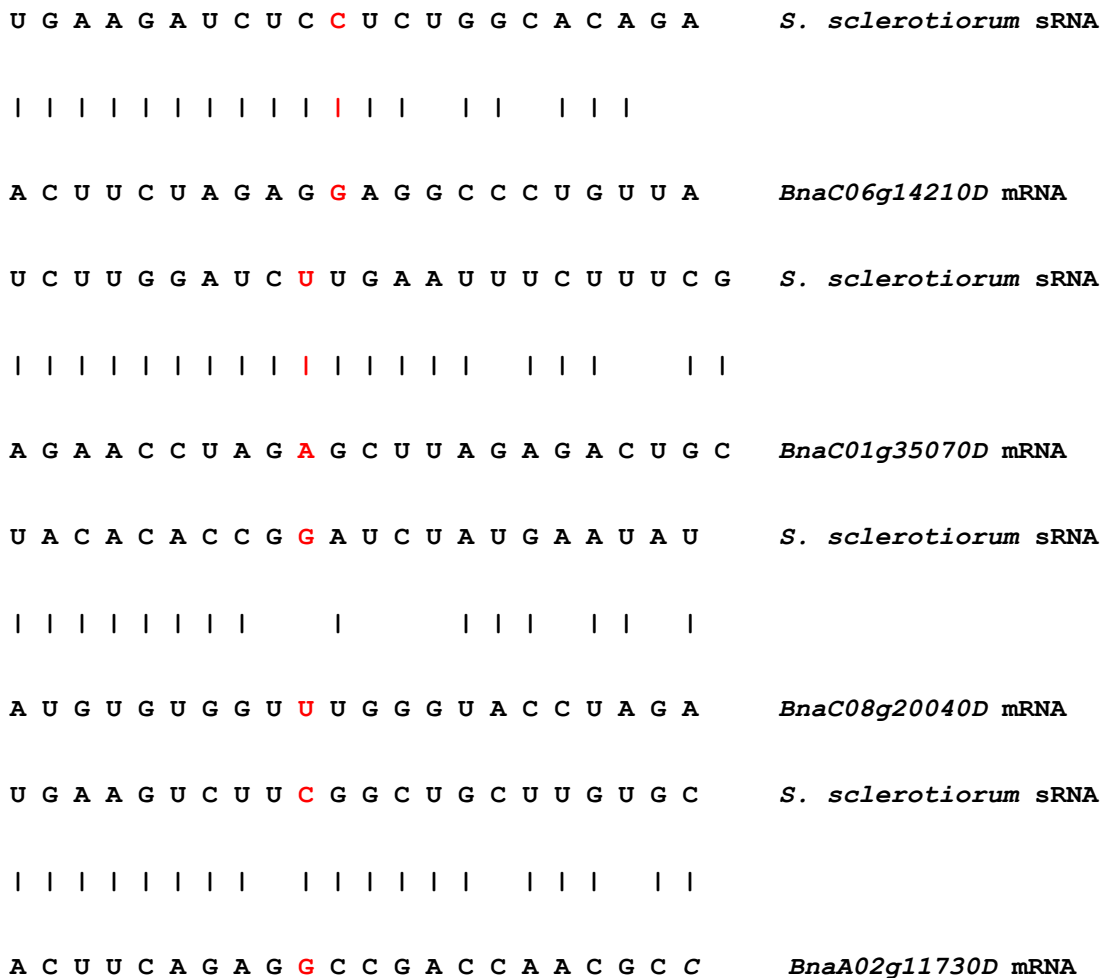


Fig .4. 7. Mapping of the cleavage sites for four *Sclerotinia sclerotiorum* small RNAs targets using 5'RACE. *BnaC06g14210D* encodes a protein kinase, *BnaC01g35070D* encodes an AP2/ERF transcription factor, *BnaC08g20040D* encodes MORN motif and *BnaA02g11730D* encodes plastocyanin-like gene. The red coloured letters indicate the target cleavage site.

4.3.8 High throughput fungal sRNA target identification in *Brassica napus* using degradome sequencing

To identify sRNA targets in *B. napus* plants at a global level, two degradome libraries (24 HPI attached leaf samples and mock (uninoculated attached leaf samples), which captured cleaved mRNAs were sequenced. A total of 37,833,599 and 27,708,104 reads for mock and infected samples were obtained. Among these reads, 214,978 and 229,289 reads possibly derived from structural RNAs were filtered out for target prediction. We used two different sRNA datasets to find their targets. At first, we used the 156 highly expressed sRNAs to find targets from the infected samples using CleaveLand 4. Altogether, 24 cleaved products were obtained from the infected library, which are likely to be targets of fungal sRNAs. The sRNA sequence, target gene ID and their target descriptions are shown in Table 4.1. Among these 24 cleaved products from psRNA only eight targets were identified from psRNA target server. In order to gain a global overview, we also used all Dicer-derived major sRNAs from ShortStack along with miRNAs identified from the miRCAT and miRprof programs. We further tested the mRNA cleavage of gene *BnaC05g04050D*, which encodes a zinc finger domain containing protein, by a sRNA with the sequence UUUGCUUCCGGACUGUUCUCC with 5' RACE. Fig. 4.8 shows the 5' RACE product of expected size of 431 bp cloned into TOPO-TA vector and confirmed with a restriction digest.

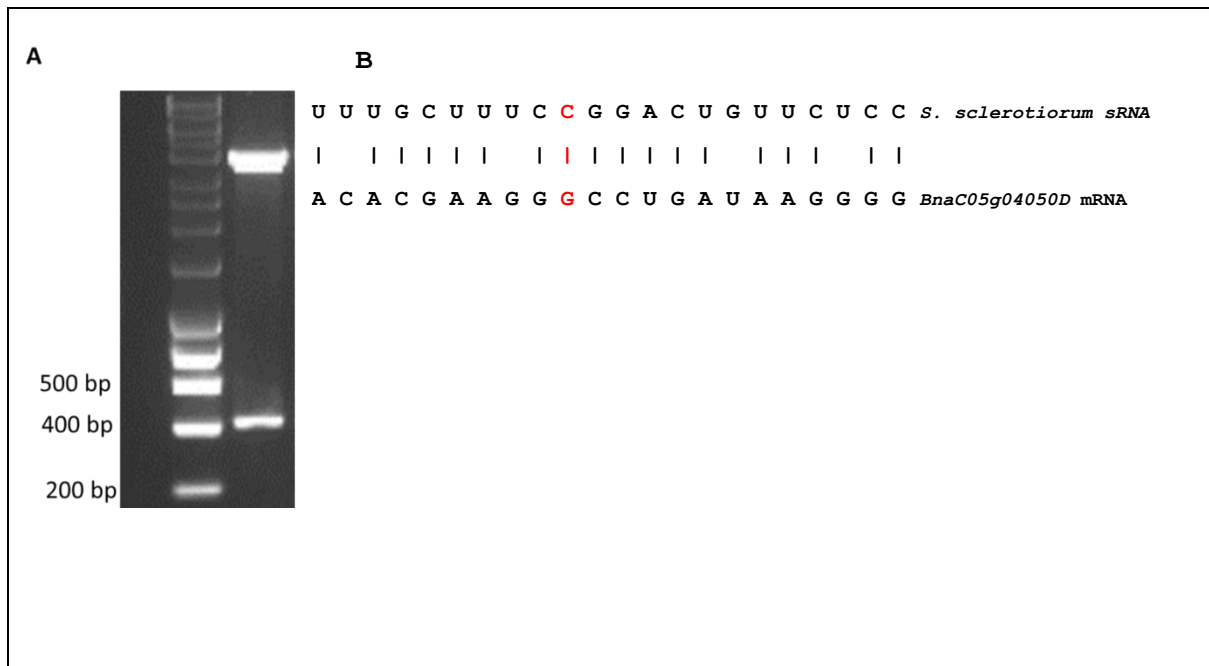


Fig .4.8. Validation of sRNA-mediated cleavage of *Brassica napus* zinc finger encoding transcript. (A) The 5'RACE clones in TOPO-TA vector confirmed by restriction digest with the upper band showing the backbone of the vector and the lower band shows the 5'RACE product size released after the digestion with *EcoRI*. (B) Mapping of the target cleavage site (indicated in with the red letter) in the *BnaC05g04050D* gene, which contains a zinc finger domain.

Our second sRNA dataset includes all Dicer-derived major sRNAs and miRNAs. Altogether, we found 1,330 target genes in our infected degradome samples (Supplementary Table 4.4) while only 149 targets were cleaved in the mock samples. T-plots of 3 target genes from this set are shown in Fig. 4.9. The level of coverage of these genes was quite high in the infected sample in compared to the mock sample revealing these genes were genuine targets of the *S. sclerotiorum* sRNAs. These genes, *BnaA06g0070D*, *BnaA0606190D*, and *BnaC08g121710D* contained the domains “Thaumatic”, “Leucine-rich repeat”, and “UDP-glycosyltransferase”, respectively.

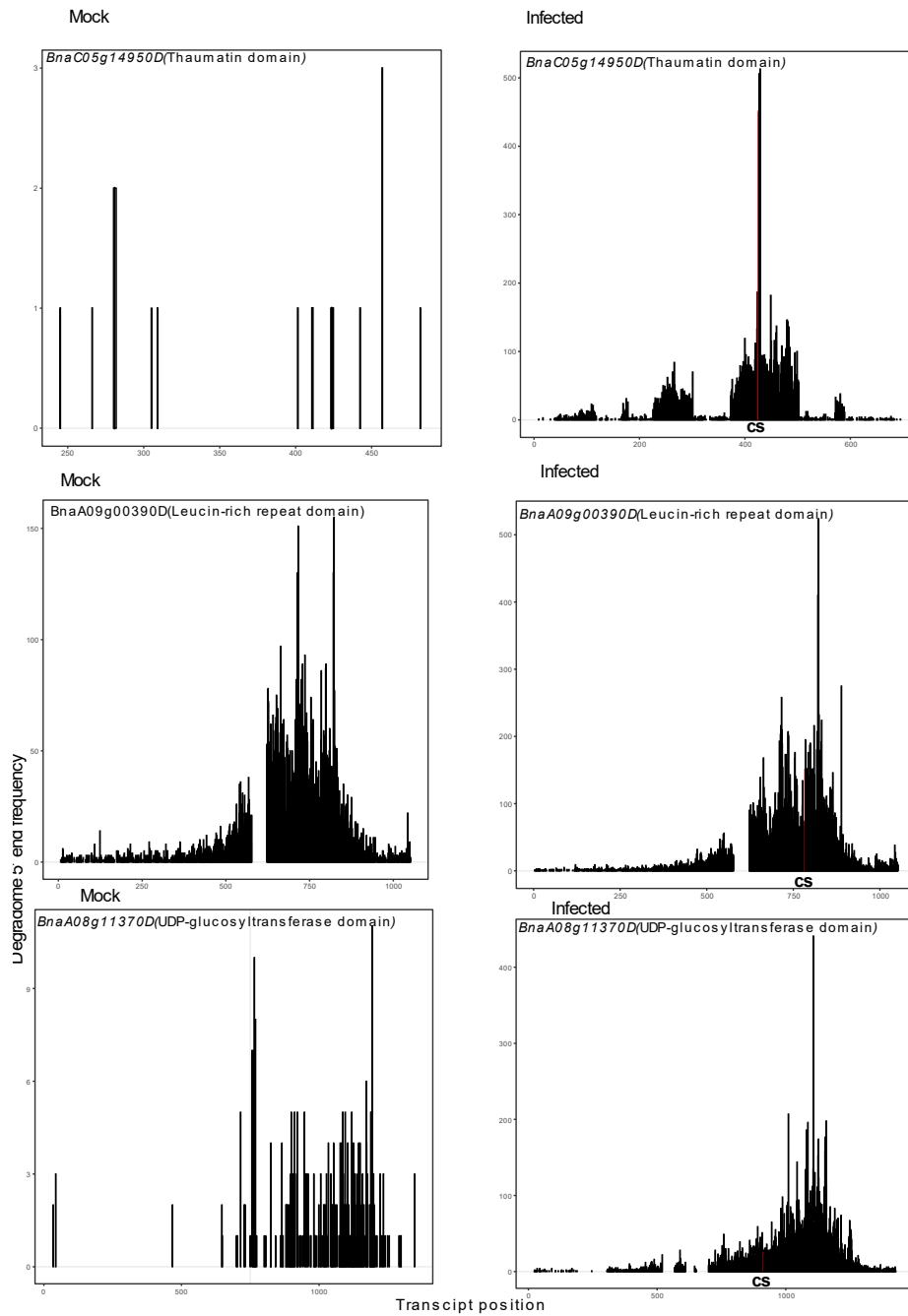


Fig. 4.9. Experimental evidence of sRNA-mediated cleavage of *Brassica napus* immune related genes by degradome sequencing. T-plots showing transcript position on the x-axis and degradome coverage on the y-axis, CS shows the cleavage site in the gene.

Table. 4.1. The list of *Sclerotinia sclerotiorum* smallRNAs and their *Brassica napus* gene targets with their predicted cleavage site position in base pairs using psRNAtarget server and their validated cleavage site by degradome sequencing.

Small RNA sequences	Target gene	Target description	psRNAtarget predicted cleavage site	Degradome validated target cleavage site
UACACACCGGAUCUAUGAAUUAU	<i>BnaA03g10610D</i>	Serine acetyl transferase, N-terminal	379	379
UACACACCGGAUCUAUGAAUUAU	<i>BnaC06g18990D</i>	Translocation protein Sec63	ND	1782
UACACGUAUCUCGGCUUCUGCUC	<i>BnaC07g22520D</i>	ACT domain	ND	587
UCCGGGUAGGUGUAGAACGUCU	<i>BnaC05g44780D</i>	Fork-head associated (FHA) domain	267	267
UAUUUGAUCUUGCCUCUGAUGUA	<i>BnaC04g08440D</i>	UDP-glucuronosyl/UDP-glucosyltransferase	ND	730
UAUUUGAUCUUGCCUCUGAUGUA	<i>BnaC06g20590D</i>	AP2/ERF domain	638	638
UAAUCCGUAUCUAGUUCUGAAUG	<i>BnaA10g25730D</i>	Protein of unknown function DUF1997	ND	296

UUAUCGGUCUGAUCUGGUUGAUC	<i>BnaC02g39910D</i>	Tubulin/FtsZ, GTPase domain	ND	235
UUAUCGGUCUGAUCUGGUUGAUC	<i>BnaC07g44840D</i>	D-isomer specific 2-hydroxyacid dehydrogenase, catalytic domain	ND	1610
UGCUCUUCGCUCUAAAUCUGAGG	<i>BnaA09g02360D</i>	Thioredoxin-like fold	ND	835
UUGAAUCGGGCUCCACAAGG	<i>BnaA05g25200D</i>	AP2/ERF domain	1037	1037
UAGGAAUUCUGUUUCUGUGACGU	<i>BnaA05g03420D</i>	Chitin-binding, type 1, conserved site	ND	342
UAGGAAUUCUGUUUCUGUGACGU	<i>BnaC03g36710D</i>	Not available	ND	3804
UUUGCUUCCGGACUGUUCUCC	<i>BnaC03g73610D</i>	ABC transporter, transmembrane domain	ND	1282
UUUGCUUCCGGACUGUUCUCC	<i>BnaA04g19600D</i>	Protein kinase, catalytic domain	ND	1326
UUUGCUUCCGGACUGUUCUCC	<i>BnaC05g04050D</i>	Zinc finger, RING-type	ND	353
UAUUGUGGGCUGACUGUAUUC	<i>BnaA04g02500D</i>	Glycoside hydrolase, family 17	ND	699

UCUUUCGGUCGAGUUGGUGCU	<i>BnaC03g01350D</i>	Pyridine nucleotide-disulphide oxidoreductase, NAD-binding domain	ND	709
UGUACCCUAUUGUAGACCGAUA	<i>BnaCnng11930D</i>	Zinc finger, FYVE-type	ND	1155
UAAGGAUCUUGGGGCUUUUCUCA	<i>BnaA01g18400D</i>	UDP-glucuronosyl/UDP-glucosyltransferase	685	685
UAGAGGCUUUCUGAAUGUAGGA	<i>BnaC09g01870D</i>	UDP-glucuronosyl/UDP-glucosyltransferase	617	617
UCGCGGAUGAGAUAGGUGUGUA	<i>BnaA09g29790D</i>	UDP-glucuronosyl/UDP-glucosyltransferase	1513	509
UCGCGGAUGAGAUAGGUGUGUA	<i>BnaC05g50780D</i>	UDP-glucuronosyl/UDP-glucosyltransferase	ND	506
UGAAGCUCGGGAUCAUCAUCUG	<i>BnaA05g16610D</i>	Hypothetical protein	1328	1328

4.4 Discussion

Some fungi have evolved to produce pathogenic sRNAs that migrate to host cells and silence host genes to promote disease. These sRNAs can function as effectors due to their contribution to disease promotion [6]. This research showed that *S. sclerotiorum* produces several sRNAs which may silence *B. napus* transcripts. A length distribution and a 5' nucleotide bias dictates the biogenesis and function of sRNAs. The sRNAs identified here exhibit typical size

distribution between 21 and 24 nucleotides and a 5' uracil bias which agrees with the previous reported result in other filamentous fungi [48-50]. This also corroborates the previous studies which conducted sRNA studies in *S. sclerotiorum* [36, 37, 51]. sRNAs with a strong bias of 5' uracil is associated with the loading in AGO1 in Arabidopsis [35]. This suggests that *S. sclerotiorum* sRNAs might have the potential to be loaded into Argonaute proteins of hosts during infection.

Many sRNA loci in fungi that have been studied previously are located in regions of the genome, containing genes, repeat elements and structural RNAs. Among these genomic regions, most sRNAs mapped to repetitive regions [7, 11]. We found that almost 88% of the highly expressed sRNAs of *S. sclerotiorum* CU8.24 resided in repeat elements suggesting that *S. sclerotiorum* sRNAs effectors are associated with transposable elements. In *Puccinia triticina*, one quarter of the sRNAs originate from repeat elements [7] and previous sRNA profiling in *S. sclerotiorum* also identified the repeat elements as hot spots for sRNA evolution [11]. sRNAs have also been reported to regulate the expression and movement of transposable elements [52]. Effector genes in necrotrophic broad host range pathogens are also associated with repeat sequences [53-55]. Mapping of sRNA loci with mostly repeat elements might be significant for their evolution to counter the immune systems of plants.

Fungal sRNAs can be involved in silencing of immune-related genes of the host [6-8]. We predicted targets of *S. sclerotiorum* sRNAs in the *B. napus* genome. Interestingly, we found *S. sclerotiorum* sRNA targets with diverse GO terms such as transcription factors, photosynthesis, xylan biosynthetic process, carbohydrate metabolic process related genes, protein kinases, response to abscisic acid, and photophosphorylation. In plants, cell wall associated kinases and xylans are involved in pathogen recognition via cell expansion and act as the first line of

defences in plants [56]. Cell wall associated kinases are receptor like kinases that have a cytoplasmic protein kinase domain, and their expression is triggered by numerous environmental stimuli including pathogen attack [57]. Xylan plays an important role in the integrity of the plant cell wall and increase cell wall recalcitrance to enzymatic digestion which helps plants to defend against pathogen attack [58]. *A. thaliana* mutants impaired in cell wall composition exhibited enhanced resistance to *B. cinerea* infection indicating the role of these classes of genes in the response to pathogens [59]. AP2/ERF transcription factors are only present in plants and are involved in various biological processes including regulation of disease resistance pathways [43]. ERF genes have been shown to regulate disease-related stimuli, such as ethylene (ET), jasmonic acid (JA), salicylic acid (SA), and infection by pathogens. Altogether, 321 AP2/ERF transcription factors were reported in *B. napus* with roles in biotic stress. Abscisic acid (ABA) is involved in plant responses to a broad range of pathogens [60]. ABA biosynthesis in plant is crucial for effective disease resistance against necrotrophic fungal pathogens [61, 62]. The *B. napus* genes that were predicted fungal sRNA targets were also significantly enriched in photosynthesis-related domains suggesting that *S. sclerotiorum* sRNAs may affect chlorophyll related processes and suppress the plant defence ability [47]. One of the strategies of pathogen to hijack the plant immune system is by weakening the chloroplast-mediated responses and to promote pathogenesis [63]. *B. napus* photosynthesis related genes were downregulated during *S. sclerotiorum* infection [64]. Prediction and validation of these genes which exhibited domains related to plant immune responses suggest *S. sclerotiorum* sRNAs like another fungal effector might suppress host immunity. It is worth mentioning that five of the targets are UDP- glycotransferase genes. glycotransferase genes are pathogen-responsive in Arabidopsis and are necessary for resistance to *Pseudomonas syringae* pv tomato [65]. Further investigation on their impact of silencing them on the plant would be interesting. From our datasets we also found Thaumatin protein as

the highest coverage cleaved genes from the infected samples. Thaumatin protein has been reported to have a role in host defence mechanism. This suggest that *S. sclerotiorum* sRNAs manipulate host defence related genes in *B. napus* [66].

Identification of sRNA and their targets facilitates a better understanding of the biological functions and regulatory mechanisms during *S. sclerotiorum* infection of hosts. Computational, degradome sequencing, and 5'RACE are three methods of target predictions. Computational target prediction suffers from a high rate of false positives, and 5' RACE is time consuming and a low throughput method. In this study also, we found discrepancy in target identified from degradome and psRNA targets. Not all targets predicted from degradome are present in the set of psRNA target and vice-versa. Such difference might be attributed to the assumptions and rules governed to find the cleavage site. psRNA target server predicts the cleavage site only at 10th and 11th position in mRNAs complementary to sRNAs regardless of fragment abundance. However, the position on the cleavage site from degradome sequencing depends on fragment abundance of cleaved products. In contrast, degradome evidence of transcript cleavage at particular sites corresponding to sRNA sequences identified through degradome sequencing is high throughput [67]. To date, only a couple of studies investigated cross-kingdom silencing using degradome sequencing [68, 69] while other studies have been based on *in silico* target prediction followed with experimental evaluation with 5' RACE [69]. Here, we developed degradome sequencing to support our hypothesis of cross-kingdom silencing in the interaction between *S. sclerotiorum* and *B. napus*. While we found cleavage among a number of genes from our degradome datasets, we might have missed many targets that were silenced through other mechanisms like translational repression. We also found some genes were detected in 5' RACE experiments but not in degradome sequencing. This might be due to the low signal of those genes in degradome sequencing.

4.5 References

1. Borges, F. and R.A. Martienssen, *The expanding world of small RNAs in plants*. Nature reviews Molecular Cell Biology, 2015. **16**(12): p. 727-741.
2. Großhans, H. and W. Filipowicz, *Molecular biology: the expanding world of small RNAs*. Nature, 2008. **451**(7177): p. 414.
3. Chen, Y., et al., *Characterization of RNA silencing components in the plant pathogenic fungus Fusarium graminearum*. Scientific Reports, 2015. **5**: p. 12500.
4. Guleria, P., et al., *Plant small RNAs: biogenesis, mode of action and their roles in abiotic stresses*. Genomics, Proteomics & Bioinformatics, 2011. **9**(6): p. 183-199.
5. Weiberg, A., et al., *Fungal small RNAs suppress plant immunity by hijacking host RNA interference pathways*. Science, 2013. **342**(6154): p. 118-123.
6. Wang, M., et al., *Botrytis small RNA Bc-siR37 suppresses plant defense genes by cross-kingdom RNAi*. RNA Biology, 2017. **14**(4): p. 421-428.
7. Dubey, H., et al., *Discovery and profiling of small RNAs from Puccinia triticina by deep sequencing and identification of their potential targets in wheat*. Functional & Integrative Genomics, 2019. **19**(3): p. 391-407.
8. Jian, J. and X. Liang, *One small RNA of Fusarium graminearum targets and silences CEBiP gene in common wheat*. Microorganisms, 2019. **7**(10).

9. Fei, S., et al., *Small RNA profiling of Cavendish banana roots inoculated with Fusarium oxysporum f. sp. cubense race 1 and tropical race 4*. *Phytopathology Research*, 2019. **1**(1): p. 22.
10. Kettles, G.J., et al., *sRNA profiling combined with gene function analysis reveals a lack of evidence for cross-kingdom RNAi in the wheat–Zymoseptoria tritici pathosystem*. *Frontiers in Plant Science*, 2019. **10**: p. 892.
11. Derbyshire, M., et al., *Small RNAs from the plant pathogenic fungus Sclerotinia sclerotiorum highlight host candidate genes associated with quantitative disease resistance*. *Molecular Plant Pathology*, 2019. **20**(9): p. 1279-1297.
12. Seifbarghi, S., et al., *Changes in the Sclerotinia sclerotiorum transcriptome during infection of Brassica napus*. *BMC Genomics*, 2017. **18**(1): p. 1-37.
13. Denton-Giles, M., et al., *Partial stem resistance in Brassica napus to highly aggressive and genetically diverse Sclerotinia sclerotiorum isolates from Australia*. *Canadian Journal of Plant Pathology*, 2018: p. 1-11.
14. Perchepped, L., et al., *Nitric oxide participates in the complex interplay of defense-related signaling pathways controlling disease resistance to sclerotinia sclerotiorum in arabidopsis thaliana*. *Molecular Plant-Microbe Interactions*, 2010. **23**(7): p. 846-860.
15. Regmi, R., et al., *Identification of B. napus small RNAs responsive to infection by a necrotrophic pathogen*. *BMC Plant Biology*, 2021. **21**(1): p. 1-21.
16. Ma, Z., C. Coruh, and M.J. Axtell, *Arabidopsis lyrata small RNAs: transient MIRNA and small interfering RNA loci within the Arabidopsis genus*. *The Plant Cell*, 2010. **22**(4): p. 1090-1103.

17. Andrews, S., *FastQC: a quality control tool for high throughput sequence data*. 2010, Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom.
18. Bushnell, B.B.A.F., *Accurate, Splice-Aware Aligner*. United States. , *BBMap: A Fast, Accurate, Splice-Aware Aligner*, United States.. 2014.
19. Kalvari, I., et al., *Non-coding RNA analysis using the Rfam database*. *Current Protocols in Bioinformatics*, 2018. **62**(1): p. e51.
20. Nawrocki, E.P. and S.R. Eddy, *Infernal 1.1: 100-fold faster RNA homology searches*. *Bioinformatics*, 2013. **29**(22): p. 2933-2935.
21. Langmead, B. and S.L. Salzberg, *Fast gapped-read alignment with Bowtie 2*. *Nature Methods*, 2012. **9**(4): p. 357-359.
22. Axtell, M.J., *ShortStack: comprehensive annotation and quantification of small RNA genes*. *RNA*, 2013. **19**(6): p. 740-751.
23. Moxon, S., et al., *A toolkit for analysing large-scale plant small RNA datasets*. *Bioinformatics*, 2008. **24**(19): p. 2252-2253.
24. Friedländer, M.R., et al., *miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades*. *Nucleic Acids Research*, 2012. **40**(1): p.
25. Derbyshire, M., et al., *The complete genome sequence of the phytopathogenic fungus *Sclerotinia sclerotiorum* reveals insights into the genome architecture of broad host range pathogens*. *Genome Biology and Evolution*, 2017. **9**(3): p. 593-618.
26. Quinlan, A.R. and I.M. Hall, *BEDTools: a flexible suite of utilities for comparing genomic features*. *Bioinformatics*, 2010. **26**(6): p. 841-842.

27. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2*. *Genome Biology*, 2014. **15**(12): p. 550.
28. Benjamini, Y. and Y. Hochberg, *Controlling the false discovery rate: a practical and powerful approach to multiple testing*. *Journal of the Royal statistical society: series B (Methodological)*, 1995. **57**(1): p. 289-300.
29. Dai, X. and P.X. Zhao, *psRNATarget: a plant small RNA target analysis server*. *Nucleic Acids Research*, 2011. **39**(suppl_2): p. W155-W159.
30. Allan, J., et al., *The host generalist phytopathogenic fungus Sclerotinia sclerotiorum differentially expresses multiple metabolic enzymes on two different plant hosts*. *Scientific Reports*, 2019. **9**(1): p. 1-15.
31. Thody, J., et al., *PAREsnip2: a tool for high-throughput prediction of small RNA targets from degradome sequencing data using configurable targeting rules*. *Nucleic Acids Research*, 2018. **46**(17): p. 8730-8739.
32. Addo-Quaye, C., W. Miller, and M.J. Axtell, *CleaveLand: a pipeline for using degradome data to find cleaved small RNA targets*. *Bioinformatics*, 2008. **25**(1): p. 130-131.
33. Alexa, A., *Rahnenfuhrer, J. topGO: enrichment analysis for gene ontology (R package version 2.40. 0)*. *Bioconductor*. 2020.
34. Marzano, S.Y.L., A. Neupane, and L. Domier, *Transcriptional and small RNA responses of the white mold fungus sclerotinia sclerotiorum to infection by a virulence-attenuating hypovirus*. *Viruses*, 2018. **10**(12).
35. Mi, S., et al., *Sorting of small RNAs into Arabidopsis argonaute complexes is directed by the 5' terminal nucleotide*. *Cell*, 2008. **133**(1): p. 116-127.

36. Kusch, S., et al., *Small RNAs from cereal powdery mildew pathogens may target host plant genes*. Fungal Biology, 2018. **122**(11): p. 1050-1063.
37. Zhou, J., et al., *Identification of microRNA-like RNAs in a plant pathogenic fungus Sclerotinia sclerotiorum by high-throughput sequencing*. Molecular Genetics and Genomics, 2012. **287**(4): p. 275-282.
38. Zhang, W., et al., *Multiple distinct small RNAs originate from the same microRNA precursors*. Genome Biology, 2010. **11**(8): p. 1-18.
39. Kasashima, K., Y. Nakamura, and T. Kozu, *Altered expression profiles of microRNAs during TPA-induced differentiation of HL-60 cells*. Biochemical and Biophysical Research Communications, 2004. **322**(2): p. 403-410.
40. Mineno, J., et al., *The expression profile of microRNAs in mouse embryos*. Nucleic Acids Research, 2006. **34**(6): p. 1765-1771.
41. Meunier, J., et al., *Birth and expression evolution of mammalian microRNA genes*. Genome Research, 2013. **23**(1): p. 34-45.
42. Tena, G., M. Boudsocq, and J. Sheen, *Protein kinase signaling networks in plant innate immunity*. Current Opinion in Plant Biology, 2011. **14**(5): p. 519-529.
43. Gutterson, N. and T.L. Reuber, *Regulation of disease resistance pathways by AP2/ERF transcription factors*. Current Opinion in Plant biology, 2004. **7**(4): p. 465-471.
44. Ma, H., et al., *MORN motifs in plant PIPKs are involved in the regulation of subcellular localization and phospholipid binding*. Cell Research, 2006. **16**(5): p. 466-478.
45. Hung, C.-Y., et al., *Phosphoinositide-signaling is one component of a robust plant defense response*. Frontiers in Plant Science, 2014. **5**: p. 267.

46. Schottler, M.A., H. Kirchhoff, and E. Weis, *The role of plastocyanin in the adjustment of the photosynthetic electron transport to the carbon metabolism in tobacco*. *Plant Physiology*, 2004. **136**(4): p. 4265-4274.
47. Menzies, I.J., et al., *Leaf colour polymorphisms: a balance between plant defence and photosynthesis*. *Journal of Ecology*, 2016. **104**(1): p. 104-113.
48. Mueth, N.A., S.R. Ramachandran, and S.H. Hulbert, *Small RNAs from the wheat stripe rust fungus (*Puccinia striiformis* f. sp. *tritici*)*. *BMC Genomics*, 2015. **16**(1): p. 718.
49. Vetukuri, R.R., et al., *Evidence for small RNAs homologous to effector-encoding genes and transposable elements in the oomycete *Phytophthora infestans**. *PLoS One*, 2012. **7**(12): p. e51399.
50. Yang, F., *Genome-wide analysis of small RNAs in the wheat pathogenic fungus *Zymoseptoria tritici**. *Fungal Biology*, 2015. **119**(7): p. 631-640.
51. Mochama, P., et al., *Mycoviruses as triggers and targets of RNA silencing in white mold fungus *Sclerotinia sclerotiorum**. *Viruses*, 2018. **10**(4): p. 214.
52. Aravin, A.A., et al., *Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the *D. melanogaster* germline*. *Current Biology*, 2001. **11**(13): p. 1017-1027.
53. Dallery, J.-F., et al., *Gapless genome assembly of *Colletotrichum higginsianum* reveals chromosome structure and association of transposable elements with secondary metabolite gene clusters*. *BMC Genomics*, 2017. **18**(1): p. 1-22.

54. Syme, R.A., et al., *Comprehensive annotation of the Parastagonospora nodorum reference genome using next-generation genomics, transcriptomics and proteogenomics*. PLoS One, 2016. **11**(2): p. e0147221.
55. Wang, Q., et al., *Characterization of the two-speed subgenomes of Fusarium graminearum reveals the fast-speed subgenome specialized for adaption and infection*. Frontiers in Plant Science, 2017. **8**: p. 140.
56. Verica, J.A. and Z.-H. He, *The cell wall-associated kinase (WAK) and WAK-like kinase gene family*. Plant Physiology, 2002. **129**(2): p. 455-459.
57. Kohorn, B.D. and S.L. Kohorn, *The cell wall-associated kinases, WAKs, as pectin receptors*. Frontiers in Plant Science, 2012. **3**: p. 88.
58. Wierzbicki, M.P., et al., *Xylan in the middle: understanding xylan biosynthesis and its metabolic dependencies toward improving wood fiber for industrial processing*. Frontiers in Plant Science, 2019. **10**: p. 176.
59. Miedes, E., et al., *The role of the secondary cell wall in plant resistance to pathogens*. Frontiers in Plant Science, 2014. **5**: p. 358.
60. García-Andrade, J., et al., *The role of ABA in plant immunity is mediated through the PYR1 receptor*. International Journal of Molecular Sciences, 2020. **21**(16): p. 5852.
61. Ton, J. and B. Mauch-Mani, *β -amino-butyric acid-induced resistance against necrotrophic pathogens is based on ABA-dependent priming for callose*. The Plant Journal, 2004. **38**(1): p. 119-130.

62. Adie, B.A., et al., *ABA is an essential signal for plant resistance to pathogens affecting JA biosynthesis and the activation of defenses in Arabidopsis*. *The Plant Cell*, 2007. **19**(5): p. 1665-1681.
63. Kuźniak, E. and T. Kopczewski, *The chloroplast reactive oxygen species-redox system in plant immunity and disease*. *Frontiers in Plant Science*, 2020. **11**: p. 1798.
64. Girard, I.J., et al., *RNA sequencing of Brassica napus reveals cellular redox control of Sclerotinia infection*. *Journal of Experimental Botany*, 2017. **68**(18): p. 5079-5091.
65. Mathilde, L.M., et al, *Pathogen-responsive expression of Glycosyltransferase genes UGT73B3 and UGT73B5 is necessary for resistance to Pseudomonas syringae pv tomato in Arabidopsis*, 2005. **139**(4): p. 1890-1901.
66. Liu, J.J., et al, *The superfamily thaumatin-like proteins:its origin, evolution, and expression towards biological function*, 2010. **29**(5):p. 419-436.
67. Llave, C., et al., *Target validation of plant microRNAs, in MicroRNAs in Development*. 2011, Springer. p. 187-208.
68. Shao, J., et al., *Identification of milRNAs and their target genes in Ganoderma lucidum by high-throughput sequencing and degradome analysis*. *Fungal Genetics and Biology*, 2020. **136**: p. 103313.
69. Ma, X., J. Wiedmer, and J. Palma-Guerrero, *Small RNA bidirectional crosstalk during the interaction between wheat and Zymoseptoria tritici*. *Frontiers in Plant Science*, 2020. **10**: p. 1669.

Chapter 5: Identification of *Brassica napus* small RNAs responsive to infection by a necrotrophic pathogen

Abstract

Small RNAs are short non-coding RNAs that are key gene regulators controlling various biological processes in eukaryotes. Plants may regulate discrete sets of sRNAs in response to pathogen attack. *Sclerotinia sclerotiorum* is an economically important pathogen affecting hundreds of plant species, including the oilseed *Brassica napus*. However, there are limited studies on how regulation of sRNAs occurs in the *S. sclerotiorum* and *B. napus* pathosystem. We identified different classes of sRNAs from *B. napus* using high throughput sequencing of replicated mock and infected samples at 24 hours post-inoculation. Overall, 3,999 sRNA loci were highly expressed, of which 730 were significantly upregulated during infection. These 730 up-regulated sRNAs targeted 64 genes, including disease resistance proteins and transcriptional regulators. A total of 73 conserved miRNA families were identified in our dataset. Degradome sequencing identified 2,124 cleaved mRNA products from these miRNAs from combined mock and infected samples. Among these, 50 genes were specific to infection. Altogether, 20 conserved miRNAs were differentially expressed, and 8 transcripts were cleaved by the differentially expressed miRNAs miR159, miR5139, and miR390, suggesting they may have a role in the *S. sclerotiorum* response. A miR1885-triggered disease resistance gene-derived secondary sRNA locus was also identified and verified with degradome sequencing. We also found further evidence for silencing of a plant immunity related ethylene response factor gene by a novel sRNA using 5'-RACE and RT-qPCR. The findings in this study expand the framework for understanding the molecular mechanisms of the *S. sclerotiorum* and *B. napus* pathosystem at the sRNA level.

5.1 Introduction

Small RNAs (sRNAs) are short non-coding RNAs, ranging in size from 18-30 nucleotides (nt), that are important for gene expression regulation and genome stability in eukaryotes [1]. There are three major sRNA classes, microRNAs (miRNAs), short interfering RNAs (siRNAs) and P-element Induced Wimpy (PIWI) associated RNAs (piRNAs); while the latter only occur in animals [2], the former two are found in plants. Different types of sRNAs have different biogenesis pathways [3].

sRNAs in plants silence gene expression through the RNA interference (RNAi) pathway. RNA dependent RNA polymerases (RdRps), Dicer-like proteins (DCLs), and Argonauts (AGOs) are the main RNAi pathway enzymes. RdRps facilitate the formation of dsRNAs, which are processed into sRNAs by DCLs. In general, AGO is thought to guide one of the strands (the guide strand) of dicer-processed sRNAs to silence complementary targets [4] while the other strand (the passenger ‘sRNA*’ strand), is often quickly degraded. However, recent studies have shown that the passenger strand can also have important gene silencing roles in plants [5-7]. In addition to complementary pairing with transcripts, siRNAs can also regulate gene expression epigenetically by RNA-directed methylation of complementary DNA [1].

Plants have many classes of siRNA [1, 8-10]. The main siRNA classes are the hairpin-siRNAs (hp-siRNAs), natural-antisense siRNAs (natsiRNAs), secondary siRNAs and heterochromatic siRNAs (hetsiRNAs) [1]. These classes have distinct biogenesis pathways and may involve DCL protein mediated cleavage of duplexes from either very long hairpins (hp-siRNAs) or double stranded RNAs generated from single stranded precursors by RdRp enzymes [1]. In addition to silencing mRNAs, miRNA-mediated cleavage of mRNAs or non-coding RNA precursors may also produce secondary siRNAs. These are often described as phased siRNAs

(pha-siRNAs) as they appear at precise 21-22 nucleotide intervals from the miRNA cleavage site. Loci that produce pha-siRNAs are known as 'PHAS' loci [11]. The secondary RNAs produced by PHAS loci may silence the gene from which they are derived or they may act in *trans* to silence the expression of other genes; the latter type of secondary siRNAs is known as trans-acting siRNAs (ta-siRNAs) and their biogenesis loci are often referred to as 'TAS' genes.

While a large number of miRNA-triggered secondary siRNAs have been identified in the genomes of plants, only a few have been experimentally validated [12-15]. Four miRNA-triggered ta-siRNA families have been characterized in the model plant *Arabidopsis thaliana* [13]. Of these, the miRNA390-triggered TAS3 genes were found to be conserved across various plant species. The formation of pha-siRNAs depends on several protein components, including SUPPRESSOR OF GENE SILENCING 3 (SGS3), RDR6 and DCL4 [13]. Studies on PHAS loci in different plants have shown that miRNAs trigger pha-siRNA production from many types of transcript, including noncoding RNAs, and the mRNAs of disease resistance and pentatricopeptide repeat genes [16]. Nucleotide-binding site leucine-rich repeat (NBS-LRR) genes form the largest set of genes identified so far that can potentially produce pha-siRNAs upon binding of specific miRNAs [17].

Small RNAs in plants regulate genes associated with various biological processes such as seed germination [18], organ development [19] and maturation [20], signal transduction [21], and stress response [22]. Plants under pathogen attack may employ various sRNA-regulated immune pathways [23, 24]. For example, while studying the sRNAome in wheat cultivars during *Puccinia graminis* infection, Gupta et al. (2012) reported that miR408 exhibits different expression patterns in susceptible and resistant cultivars after a two-day course of infection [25]. Some immunity-related sRNAs have also been functionally characterised. For example, in the model plant *A. thaliana*, miR393 targets different auxin signalling genes to confer

antibacterial resistance [22] and miR408 is a negative regulator of plantacyanins and laccase genes [26]. These latter genes have roles in stress responses, cell-to-cell signalling and maintaining plasticity and vigour of the cell wall. In addition, overexpression of miR7695 results in an incremental increase in resistance in rice against the blast fungus *Magnaporthe oryzae* [27]. During viral infection of plants, changes in the accumulation of miRNAs result in production of different pha-siRNAs [28]. In legumes and tomato, a number of miRNA families are involved in triggering pha-siRNAs by binding to the transcripts of NB-LRR genes [16, 29]. In tomato, the abundance of secondary siRNAs from disease resistance genes was lower during bacterial and viral infection, suggesting that pha-siRNA production is important for fine-tuning defence responses [29]. Recently Cui et al. (2020) demonstrated the role of miR1885-mediated ta-siRNA expression in maintaining plant growth and immunity in *B. napus* upon viral infection [30]. The roles of pha-siRNAs in plant response to bacterial and viral infection have been investigated in several studies but little is known about their roles in responding to pathogenic fungi. One of the few studies on this subject was characterized by Wu et al. (2017). This study characterized pha-siRNAs produced by tomato in response to *Botrytis cinerea* infection [14]. It was found that many pathogen-responsive tomato pha-siRNAs downregulate transcription factors, which is suggestive of a broad role in the regulation of gene expression regulation [15].

Canola (*Brassica napus*) is an economically important oilseed crop grown worldwide [31]. Sclerotinia stem rot (SSR), caused by the fungus *Sclerotinia sclerotiorum*, is an important disease that causes large economic losses in canola [32]. Some studies have been conducted in *Brassica* spp. to identify plant-specific miRNAs [33-35] under biotic and abiotic stresses. There have been two studies on *B. napus* miRNA expression upon *S. sclerotiorum* infection [36, 37]. However, these studies were performed with single sRNA libraries at 3, 12- and 48-

hours post-inoculation (HPI) without any replicates, which limits the proper understanding of differential expression of small RNAs in the *B. napus* response to *S. sclerotiorum*. Furthermore, in comparison to mature miRNAs deposited in miRBase for other plants like *Medicago truncatula*, *O. sativa* and *A. thaliana*, the number of miRNAs for *B. napus* is quite low, suggesting many miRNAs in *B. napus* are yet to be discovered. Finally, little is known about the triggers of PHAS loci in the *B. napus* genome and their functions in gene regulation in response to pathogens.

To assess differential expression of sRNAs, identify new pathogen-responsive miRNAs and characterise the role of secondary sRNAs in the *B. napus* response to *S. sclerotiorum*, we developed replicated sRNA libraries for mock-inoculated and *S. sclerotiorum* inoculated leaves 24 hours post-inoculation to characterize different classes of sRNAs. To identify targets of these sRNAs, we also performed high throughput degradome sequencing and, for one gene, 5'RACE and RT-qPCR.

For the first time, we identified a large number of *B. napus* sRNAs up-regulated in response to *S. sclerotiorum* infection after 24 HPI. We also found evidence of pathogen-responsive activation of novel PHAS loci likely involved in regulation of disease resistance proteins. Our follow-up 5'-RACE and qPCR studies provided further evidence of sRNA-directed regulation of a gene involved in ethylene signalling.

5.2 Methods

5.2.1 Biological materials

An Australian *S. sclerotiorum* isolate (CU8.24) originally collected from South Stirling WA was used for infection assays [38]. Mature sclerotia were cut into halves and placed onto 9 cm Petri dishes containing potato dextrose agar (PDA). After germination from the sclerotium, mycelium was subcultured onto fresh PDA medium. After 48 hours of incubation, mycelial plugs were placed on fully expanded second or third leaves of one-month-old *B. napus* plants (AV Garnet) The plants were grown for a month in a growth chamber with 16 h of daylight and 8 h of darkness. After infection, plants were carefully covered with a polythene bag to increase the humidity, thereby facilitating the infection process. Twenty-four HPI, a characteristic necrotic lesion was observed on the infected leaves. The infected tissues were carefully excised using sterilized scissors and immediately frozen in liquid nitrogen and stored at -80 °C until RNA extraction for sequencing. For mock samples, PDA only agar plugs were used without any fungal mycelium. Three leaves from three different plants were pooled together for each replicate. For small RNA sequencing, three biological replicates were sequenced separately while two degradome libraries were sequenced by pooling all infected replicates as one library and all mock replicates as another library.

5.2.2 Total RNA extraction and sequencing

Total RNAs were extracted using the TRIZOL reagent following the manufacturer's protocol (Invitrogen Carlsbad, CA, USA). After extraction, total RNAs were quantified using a NanoDrop spectrophotometer, and Qubit. The integrity of RNA samples was checked using agarose gel electrophoresis. Three to five µg and 25-30 µg of total RNA were sent to Novogene (Singapore) for small RNA and degradome sequencing respectively. The sRNA sequencing was done using the NEBNext® Multiplex Small RNA Library Prep Kit for Illumina® with

single end 50 bp reads according to the manufacturer's protocol. Degradome sequencing was done as mentioned in [39]. In brief, the construction of a degradome library was started from the degradation site (with monophosphate group) of the degraded mRNA. The sequencing adaptors were added to both ends of the degradation library and a library size of around 200-400 bp was selected. The sequencing was performed on a Hiseq 2500.

5.2.3 Analysis of small RNA sequencing data

Raw reads were trimmed using cutadapt software (version 1.15) optimized for single-end reads with a setting of cutadapt -a (universal True Seq adapter) -m18 -M30 [40]. The quality of filtered reads was checked by running in FastQC [41]. Reads with a length in the range of 18-30 nt were retained. Trimmed infected reads were assigned to the fungal [42] and plant [43] reference sequences using bbsplit software in the bbmap [44] program keeping ambig2 option set as toss. The option ambig2 removes all the reads that map to both references with equal confidence. The reads that were unique to the *B. napus* genome were kept for prediction of *B. napus* sRNAs.

For prediction of *B. napus* sRNA biogenesis loci, clean reads were aligned to the reference genome of *B. napus* allowing for a maximum of two mismatches. We used ShortStack [45] to gain an overall idea of sRNA producing loci from the *B. napus* genome and to characterize highly expressed sRNA loci after infection. Each library was used as a single entity without collapsing for input into ShortStack. For conserved miRNA prediction, we matched the clean reads against miRBase (version 22) (<http://www.miRBase.org>) using the miRProf program in UEA small RNA workbench [46]. The reads that matched to mature miRNAs in the miRBase database with 0 mismatches were considered as conserved miRNAs. The remaining reads that did not match miRBase were parsed for the prediction of novel miRNA-producing loci using the miRDeep2 program [47]. Differential expression analysis was done using the Bioconductor

package DESeq2 in R with estimate variance – mean dependencies. We used the raw cluster abundance of 6 individual libraries from ShortStack to find differentially expressed sRNA loci. The sRNAs with a Benjamini-Hochberg corrected p-value of < 0.05 were considered as differentially expressed sRNAs [48]. For the differential expression analysis of conserved miRNAs, we used the raw counts for the individual miRNA sequences identified from the miRprof program.

The phasing patterns of loci were predicted with a Perl script from the PhaseTank software (version 1.0) [49]. To find the miRNA triggered phased initiator loci, complementary cleavage sites of predicted miRNAs on PHAS loci were searched using the psRNA target server assuming that the 10th nucleotide position on the miRNA is a cleavage start position of its targeted PHAS loci [15].

5.2.4 Analysis of sRNA targeting using *in silico* target prediction and degradome sequencing data

To determine whether reads originated either from the plant or the fungus we used bbsplit [44] to categorise infected degradome reads as fungal or plant-specific reads. The filtered reads were separated from potential structural RNAs by filtering against the RFAM database [50] using the program Infernal (version 1.1.3) [51].

We used either the psRNA target server with a default setting and an expectation score of 5 for computational prediction of sRNA targets [52] or PARESnip2 [53] to validate the cleavage sites from Degradome datasets following the rules of Fahlgren and Carrington [54]. We retained the targets with category number 0-3. Category-0 are targets with a degraded product having degradome peaks more than one read and the maximum on the transcript where there

is only one maximum. Category-1 are those having degradome peaks greater than one read and are the maximum on the transcript, but there is more than one maximum. Category-2 peaks are those that have reads more than one and are above the average fragment abundance on the transcript. Category-3 signals are those that have greater than one read and are below or equal to the average fragment abundance on the transcript. Further verification of PHAS locus activation by specific miRNAs was also investigated using the degradome sequencing tags with PARESnip2. To gain a more detailed understanding of silencing by different sRNA classes, we used four different datasets: the one highly expressed major RNA per locus from the ShortStack program, the conserved miRNAs from miRBase, novel miRNAs annotated from miRDeep2 and phasiRNA identified by PhaseTank.

5.2.5 Gene ontology enrichment analysis

Gene ontology enrichment analysis was conducted on sRNA target transcripts with the topGO program from R version 3.6.1 Bioconductor package. GO term enrichment tests were performed separately on mock and infected samples. In each case, the background set was all GO terms in the *B. napus* genome and the foreground set was any gene with evidence of degradome targeting. The mock and infected samples were compared to identify genes that were enriched in the mock and depleted in the infected sample or vice versa. GO terms with a p-value < 0.05 were considered as significantly enriched or depleted [55].

5.2.6 Five prime rapid amplification of cDNA ends of a cleaved target

We conducted a 5'-RACE experiment on one of the ethylene response factor genes from our degradome dataset that is potentially cleaved by a plant sRNA. The reason for choosing this

gene is directed by previous pieces of literature where these classes of genes were shown to be crucial for defence responses in plants against pathogen attack [56] and high confidence complementary site between this siRNA and the target gene according to psRNA target server. Furthermore, the sRNA targeting this gene was hitherto uncharacterised, and it is not a miRNA or phasiRNA. The 5' RACE methods have been described in Chapter 4. Two independent samples collected from independent infection assays were used to conduct 5'RACE using the first choice RACE kit following the manufacturer's protocol (Applied Biosystems, Carlsbad, CA, USA) without adding calf intestinal Phosphatase enzyme. One sample was the same as the one used for degradome sequencing while the other was not.

5.2.7 Quantitative polymerase chain reaction

The expression levels of a sRNA target gene were analysed by RT-qPCR. One to five ug of total RNAs from mock and infected *B. napus* leaf samples were converted to cDNA using the MMLV reverse transcriptase kit (Sigma-Aldrich). The cDNA samples were then diluted 1/20 before qPCR. The qPCR analysis was performed using the Bio-Rad Taq Universal SYBR Green Supermix according to the manufacturer's instructions. The thermocycler settings were 95 °C for 2 min, then 95 °C for 15 sec, 60 °C for 30 sec and 72 °C for 15 sec, and cycled for 40 times, followed by 72 °C for 2 min. Three biological and three technical replicates were used for each sample. Relative expression was calculated as per $\log(2^{-\Delta C_t})$ method normalized to the *B. napus* housekeeping actin gene. The primers and adapters used for 5' RACE and qPCR experiments are listed in Supplementary Table 5.1.

5.3. Results

5.3.1 Overview of sequencing results

To determine the role of *B. napus* sRNAs during *S. sclerotiorum* infection we sequenced six sRNA libraries on the Illumina platform from three replicates each of mock and infected samples at 24 HPI when SSR symptoms manifested on leaves. A total of 152,090,773 raw reads were obtained from the six libraries. We retained 126,887,984 (83.24 %) high quality reads after adapter trimming and length filtering (18-30 nt) from these six libraries (Table 5.1). Assignment and removal of ambiguous reads (that map to both plant and fungal genome) resulted in 41,797,278 unique *B. napus* sRNA reads that match best to the *B. napus* genome across all libraries. The reads that potentially originated from structural RNAs (rRNAs, snRNAs, snoRNAs) accounted for ~ 5 % of this total. The clean, high-quality mappable reads were then aligned to the *B. napus* genome. The overall alignment rate was 88.7 % with the highest percentage mapping in the mock samples (above 98 %), while in infected samples an average of ~78.9 % of reads mapped, ranging from 77 to 86.3 % between replicates. Among the mapped reads, ~86 % were mapped to more than one genomic locus revealing that these sRNAs originated from genomic repeats. From our dataset, we found 14 % of sRNA reads that uniquely mapped to a single genomic locus. Table 5.1 provides an overall summary of the sequencing data.

Table 5. 1. Summary of the sequencing data

Sample	Raw reads	Clean filtered reads*	Unique <i>B. napus</i> reads [†]	Structural RNA [‡]	Clean mappable reads ^{††}	Uniquely mapped reads	Reads mapping to multiple locations	Total mapped to <i>B. napus</i>	Total mapped to <i>B. napus</i> (%)
Mock	23,754,736	20,144,362	8,944,885	318,393	8,626,492	1,047,207	7,416,613	8,463,820	98.1
Mock	22,377,118	18,983,632	7,767,860	311,538	7,456,322	1,103,293	6,242,937	7,346,230	98.5
Mock	31,509,749	26,430,696	11,421,524	450,137	10,971,387	1,535,228	9,301,292	10,836,520	98.8
Infected	26,483,807	20,626,924	3,779,014	202,125	3,576,889	445,577	2,310,074	2,755,651	77
Infected	23,182,766	18,818,235	3,866,998	223,567	3,643,431	462,630	2,214,409	2,677,039	73.5
Infected	24,782,597	21,884,135	6,016,997	508,508	5,508,489	570,565	4,180,632	4,751,197	86.3
Total	152,090,773	126,887,984	41,797,278	2,014,268	39,783,010	5,164,500	31,665,957	36,830,457	87.3

*Reads after size and adapter filtering

[†]Reads that mapped best to the *B. napus* genome from Bbsplit

[‡]Reads that potentially originated from structural RNAs (ribosomal RNA, snRNAs, etc)

^{††}Reads after removing potential structural RNAs that were used for ShortStack

To determine the grouping of infected and mock libraries we performed a principal component analysis on mean normalized counts from DESeq2 (Fig 5.1A). The principal component analysis showed the replicated datasets were well grouped for two treatment groups, i.e. mock and infected, suggesting large overall differences between these treatments. Mock and infected samples were separated along principal component 1, which explained 99 % of the variance. There was some spread between the infected samples along principal component 2. However, variance between these samples along this axis is negligible, since only 1 % of the variance was explained by PC2.

5.3.2 Characteristic features of the *Brassica napus* small RNA population

To determine which sRNAs were induced in response to infection, we produced three sequencing replicates each from mock and infected samples. The following metrics are based on the pooled biological replicates for each treatment, mock and infected. Size class distribution and 5' nucleotide bias are two important characteristics to determine the origin and activities of sRNAs. To determine whether there may be a difference in the composition of sRNA origins in mock and infected samples, we analysed the nucleotide length and 5' nucleotide bias of these sRNAs. Interestingly, we found a difference in length distribution between mock and infected samples (Fig 5.1B), suggesting that upon infection, sRNA biogenesis mechanisms are altered. In mock samples, almost 50 % of total reads belonged to 24 and 23 nucleotide (nt) sRNAs followed by 21 nt. Adenine was enriched as the 5' nucleotide in 24 nt sRNAs while cytosine was more abundant in 23 nt sRNAs. A 5' nucleotide bias toward uracil was present mostly in 22 nt sRNAs (Fig 5.1C).

In infected samples, size classes were more uniform than in mock samples (Fig 5.1B). The most abundant read size was 21 nt with a slight 5' uracil bias, followed by 23 nt with a slight cytosine bias (Fig 5.1C). We also found a peak at 26 nt in infected samples with a 5' guanine bias. Similarly, size classes of 18, 19, 20 and 22 nt were also more abundant in infected samples. However, among non-redundant reads in both samples, most were 24 nt. Apart from size distribution, the ratio of total to unique reads is also an important feature of an sRNA library [57]. The lower complexity of 21 nt sequences in the infected samples in comparison to 24-nt sequences indicates that a small number of unique reads of 21 nt are highly expressed while there are many different 24 nt sequences. Such features have been attributed to 24 nt heterochromatin sRNAs [58]. Small RNA size distributions for non-redundant reads in the mock and infected sample is shown in Supplementary Fig. 5.1.

Previous reports presented similar data with a 5' uracil bias in 21 nt and a 5' adenine bias in 24 nt sRNAs. The 24 nt 5' adenine biased siRNAs have been previously shown to be involved in RNA dependent DNA methylation in *A. thaliana* with preferential loading into AGO4, while the 21 nt 5' uracil biased sRNAs have preferential loading into AGO1 [59]. Overall, our results suggest a marked shift in the types of sRNAs expressed from mock to infected *B. napus* leaves.

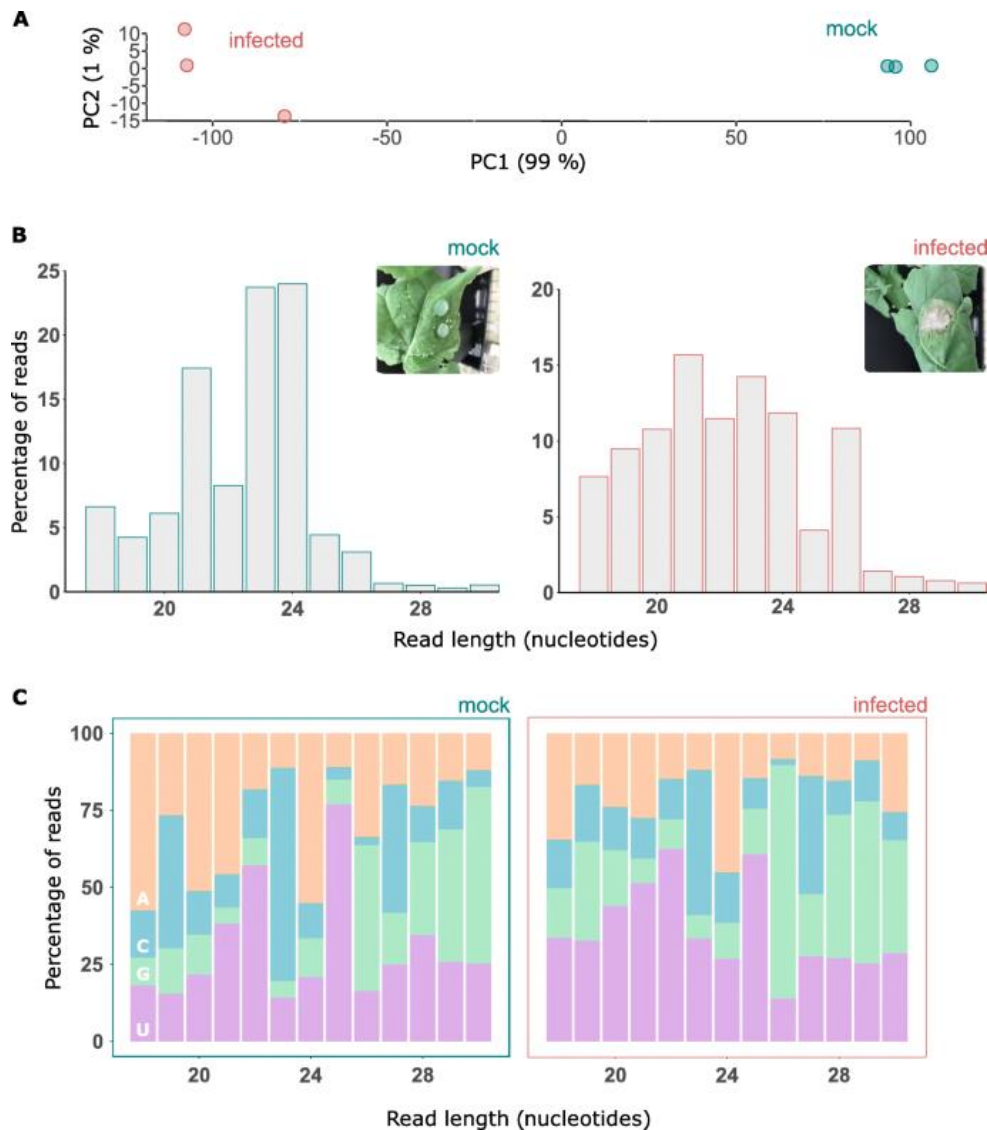


Fig. 5. 1. Changes in size class and 5' nucleotide of *Brassica napus* sRNAs in response to *Sclerotinia sclerotiorum* infection. (A) A principal component analysis based on normalized read counts from DESeq2. The x-axis shows principal component 1, which explained 99 % of the variance, and the y-axis shows principal component 2, which explained 1 % of the variance. The infected samples are depicted with red circles and the mock samples depicted

in turquoise. (B) Histogram of read sizes from pooled mock and infected samples. Inset: a representative picture of *Brassica napus* leaves under both treatments (24 hours post-inoculation (HPI)). Left: for the pooled replicates of the mock sample, y axis depicts the percentage of reads across all three replicates and the x axis read length in nucleotides. Right: for the pooled replicates of the infected sample, same information as for the mock sample. (C) Results from pooled replicates for the mock (left) and infected (right). Showing percentage of reads (y-axis) of each size class (x-axis) that had each of the four nucleotides (AGCU) in their 5' position.

5.3.3 A total of 730 unique *Brassica napus* small RNAs are upregulated in response to *Sclerotinia sclerotiorum* infection

To assess what *B. napus* sRNAs accumulate in response to infection with *S. sclerotiorum*, we performed a differential expression analysis with DESeq2. We did not only consider differential expression of miRNAs but the entire sRNA-ome in *B. napus*. ShortStack predicted 121,977 sRNA loci, 104,421 of which were likely Dicer-derived. Among these loci, 3,999 were highly expressed, with at least 100 raw major RNA sequencing reads (Supplementary Table 5.2). If these sRNAs were responding to infection, we hypothesised that they might be more expressed in infected samples as compared to mock samples. We found 915 loci significantly altered in their expression in infected samples compared to the mock samples. Among these, 730 were upregulated in *B. napus* after *S. sclerotiorum* infection; these loci produced 565 unique sRNAs based on the major sRNAs predicted by ShortStack (Fig 5.2A).

These 730 upregulated sRNAs were mostly enriched for 20 and 21 nt sequences, while the 185 downregulated major sRNAs were enriched for 24 nt sequences (Fig 5.2B). Uracil was enriched at the 5' ends of all the size classes for upregulated sRNAs except 24 nt, which had a 5' adenine bias (Fig 5.2C). However, downregulated sRNAs exhibited a 5' cytosine bias at 20 nt and 21 nt (Fig 5.2C). Size classes 22 and 24 nt shared a common 5' bias of uracil and adenine respectively in both sample groups. Overall, our data add weight to the hypothesis that sRNA

classes with distinct biogenesis and targeting pathways were expressed in response to *S. sclerotiorum* challenge.

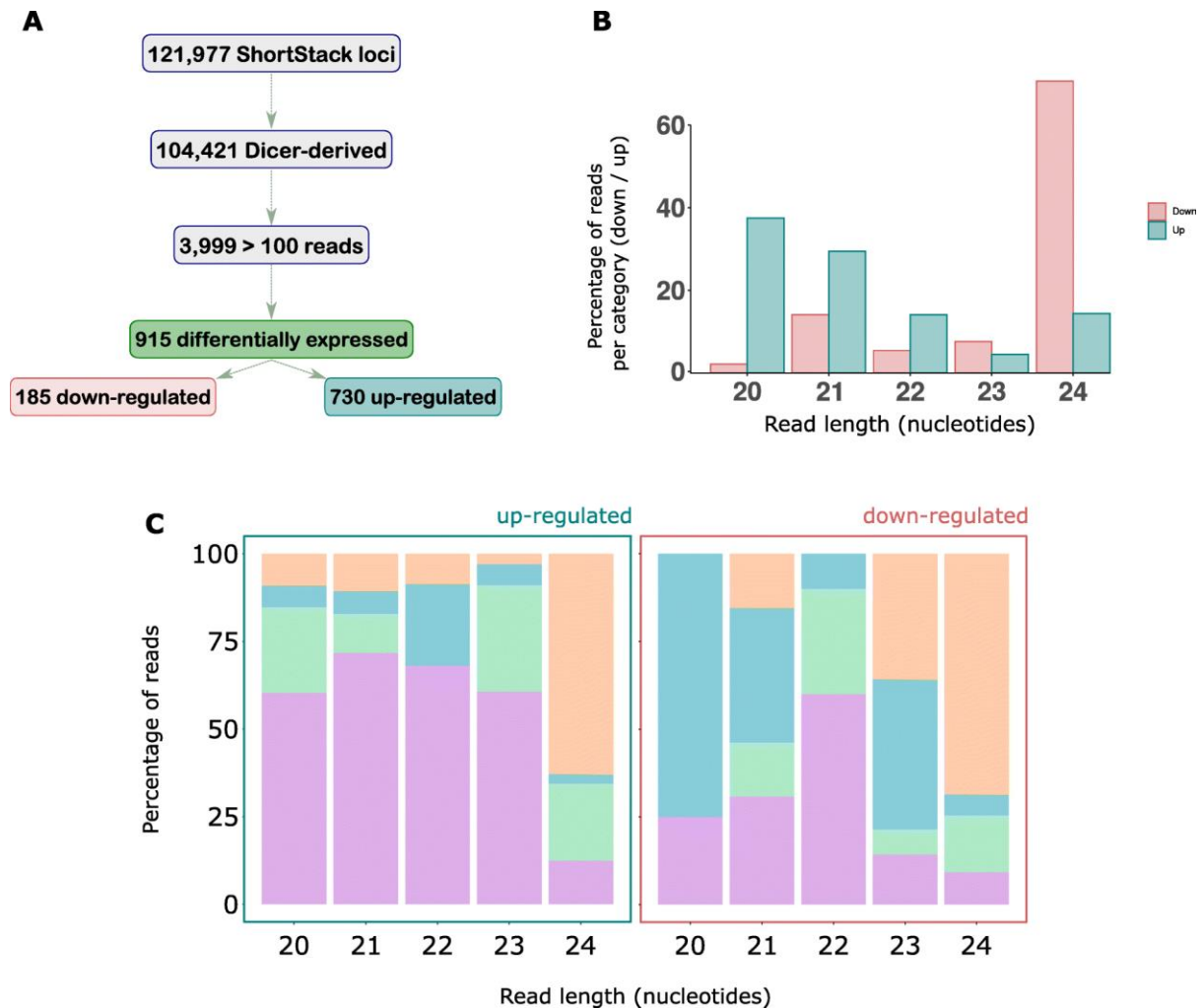


Fig. 5. 2. Small RNA population identified from *Brassica napus* genome in response to *Sclerotinia sclerotiorum* infection. (A) A flow chart showing total, Dicer-derived and highly expressed loci with a major RNA reads of ≥ 100 reads, as predicted by ShortStack, and differentially expressed loci identified from DESeq2. (B) Histogram of read sizes from upregulated and downregulated sRNA loci, y-axis depicts the percentage of reads in each category (upregulated or down-regulated) and the x axis read length in nucleotides. (C) Histogram of read sizes showing percentage of reads (y-axis) of each size class (x-axis) that had each of the four nucleotides (AGCU) in their 5' position for upregulated loci (left) and downregulated loci (right).

5.3.4 Stress responsive genes are targeted by up-regulated small RNAs

We used the degradome sequencing data to investigate targets of sRNAs upregulated during infection. A total of 64 target genes were identified from upregulated sRNAs this way (Table 5.2). Representative T-plots for four of these genes that were identified in infected libraries, which were assigned to different PARESnip2 confidence categories, are presented in Fig 5.3. Among these 64 targets, 10 were found in both libraries, resulting in 29 and 15 unique targets for mock and infected samples, respectively. Genes that were possibly regulated by small RNAs up-regulated during infection were annotated with transcription factor-related InterPro terms such as ‘ABC transcription factor’, ‘heat shock response’, ‘disease resistance protein-like’, ‘Zinc finger 1 domain’ and ‘leucine zipper domain’. This suggests that *B. napus* transcriptional regulatory networks may be modified by sRNAs specifically induced during infection.

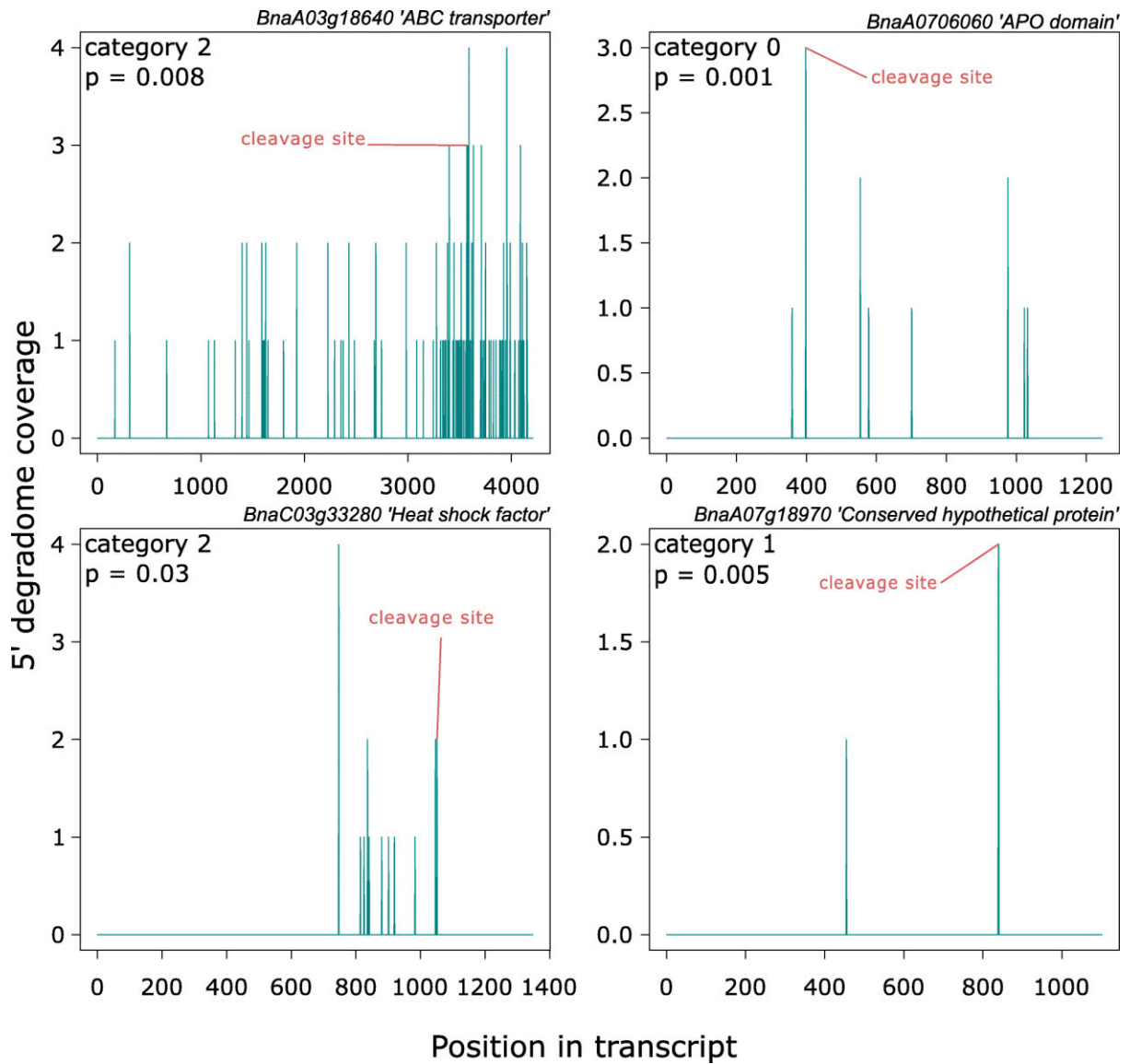


Fig. 5. 3. Representative Target plots (T-plots) for infection-specific targets of upregulated sRNAs. The x-axis shows the transcript position in the target genes and the y-axis shows the 5' read coverage at different positions; cleavage sites predicted by PARESnip2 are labeled in red. The category and p value given by PARESnip are shown in the top left-hand corners of graphs and the gene IDs and their putative functions are shown above.

Table 5.2 Predicted degraded *Brassica napus* targets of upregulated *B. napus* small RNAs based on degradome sequencing data analysed using PARESnip2.

Gene ID	Category*	Cleavage position	Alignment score	Duplex MFE ¹	Perfect MFE	MFE Ratio	p-value [‡]	Library [¶]
BnaA08g20190D	3	1211	3.5	-23.6	-32.9	0.717325	0.019978	M
BnaA02g03600D	0	748	4	-20.6	-28.9	0.712803	0.000756	M, I
BnaC03g43110D	2	219	4	-24.2	-32.3	0.749226	0.039024	M
BnaC05g00280D	2	1144	3	-24.6	-29.6	0.831081	0.016924	M
BnaA10g11740D	3	955	4	-25.8	-35.2	0.732955	0.032497	M
BnaA01g22910D	3	848	4	-28.4	-39.6	0.717172	0.038079	M
BnaCnng67150D	3	68	3	-31	-37.9	0.817942	0.01312	M
BnaA08g28020D	3	917	3.5	-29.1	-40.8	0.713235	0.033058	M
BnaA07g08860D	2	1153	1	-32.6	-39.2	0.831633	0.002904	M
BnaA09g27990D	2	1156	1	-32.6	-39.2	0.831633	0.002897	M,I
BnaC05g21250D	2	1162	1	-32.6	-39.2	0.831633	0.002885	M,I
BnaC07g11360D	2	1141	1	-32.6	-39.2	0.831633	0.002929	M
BnaA03g36860D	3	336	4	-29.6	-41.5	0.713253	0.039506	M
BnaA06g36560D	1	345	2	-33.8	-41.5	0.814458	0.008645	M
BnaC07g17320D	3	1383	2	-33.8	-41.5	0.814458	0.008237	M
BnaA04g18170D	3	388	4	-33.7	-42.3	0.79669	0.025798	M
BnaA10g18410D	3	1347	4	-31.8	-39.7	0.801008	0.021698	M
BnaA02g06410D	2	627	4	-27.5	-39.2	0.701531	0.009812	M
BnaA04g07950D	0	2282	3.5	-31.4	-43.5	0.721839	0.001638	M,I
BnaA07g25390D	0	2174	3.5	-31.4	-43.5	0.721839	0.001718	M,I
BnaA08g17390D	0	2282	4	-30.6	-43.5	0.703448	0.003145	M,I
BnaA09g26170D	0	2333	4	-30.6	-43.5	0.703448	0.003069	M,I
BnaC03g59640D	2	2267	4	-30.6	-43.5	0.703448	0.015732	M
BnaC05g23210D	0	2324	4	-30.6	-43.5	0.703448	0.00308	M,I
BnaC06g27170D	0	2168	3.5	-31.4	-43.5	0.721839	0.001722	M,I
BnaCnng25410D	0	2267	3.5	-31.4	-43.5	0.721839	0.001648	M,I
BnaC01g07210D	3	37	4	-33.7	-43.2	0.780093	0.030433	M
BnaAnng05290D	3	792	4	-26.4	-34.4	0.767442	0.009756	M

BnaA06g11000D	3	299	3	-27.3	-33.2	0.822289	0.042121	M
BnaA06g11010D	3	299	3	-27.3	-33.2	0.822289	0.042404	M
BnaC04g20940D	3	215	0	-33.2	-33.2	1	0.005071	M
BnaC08g38450D	1	1481	3.5	-27.1	-37.5	0.722667	0.001168	M
BnaCnng05480D	2	1980	3.5	-24.2	-34.3	0.705539	0.020979	M
BnaC07g37000D	2	487	0.5	-35.5	-35.9	0.988858	0.003279	M
BnaA09g16090D	3	3478	3.5	-28.4	-35.2	0.806818	0.037582	M
BnaA04g26610D	2	5713	4	-23.8	-33.9	0.702065	0.01421	M
BnaC04g50670D	2	5713	4	-23.8	-33.9	0.702065	0.013297	M
BnaC01g18190D	2	466	3	-22.1	-31	0.712903	0.013158	M
BnaCnng58300D	3	891	4	-26.5	-34.4	0.770349	0.018106	M
BnaA02g03600D	1	748	4	-20.6	-28.9	0.712803	0.003782	I,M
BnaC03g33280D	3	1046	3	-24.3	-28.8	0.84375	0.002259	I
BnaC05g38210D	2	80	4	-25.5	-35.3	0.72238	0.003571	I
BnaA07g18970D	1	840	3.5	-34.6	-46.4	0.74569	0.002773	I
BnaA07g26110D	2	1696	4	-30.9	-39.8	0.776382	0.036546	I
BnaA09g27990D	2	1156	1	-32.6	-39.2	0.831633	0.00145	I,M
BnaC05g21250D	2	1162	1	-32.6	-39.2	0.831633	0.001444	I,M
BnaA07g06060D	0	398	3.5	-31.5	-40.6	0.775862	0.000816	I
BnaA04g07950D	0	2282	3.5	-31.4	-43.5	0.721839	0.001638	I,M
BnaA07g25390D	0	2174	3.5	-31.4	-43.5	0.721839	0.001718	I,M
BnaA08g17390D	0	2282	4	-30.6	-43.5	0.703448	0.002753	I,M
BnaA09g26170D	0	2333	4	-30.6	-43.5	0.703448	0.002686	I,M
BnaC05g23210D	0	2324	4	-30.6	-43.5	0.703448	0.002695	I,M
BnaC06g27170D	0	2168	3.5	-31.4	-43.5	0.721839	0.001722	I,M
BnaCnng25410D	0	2267	3.5	-31.4	-43.5	0.721839	0.001648	I,M
BnaC03g30850D	3	1850	3.5	-23.1	-29.2	0.791096	0.035487	I
BnaA05g25860D	3	743	4	-20.8	-29.5	0.705085	0.017563	I
BnaA09g37590D	2	3003	4	-24.1	-33	0.730303	0.049431	I
BnaA09g04340D	3	45	4	-25.9	-33.3	0.777778	0.01457	I
BnaA08g00270D	3	152	4	-28	-36	0.777778	0.021834	I
BnaAnng16680D	2	735	3.5	-34.1	-46.4	0.734914	0.016373	I

BnaA06g04910D	3	2188	4	-24	-30.4	0.789474	0.011989	I
BnaA03g18640D	2	3633	2.5	-25.5	-30.4	0.838816	0.001911	I
BnaC08g21330D	3	587	4	-26.8	-37.1	0.722372	0.013281	I
BnaC02g30040D	2	367	2	-27.8	-34.3	0.810496	0.007331	I

* Category derived from PARESnip2 based on rules of Falgren and Carrington, 2010.

[†]MFE = minimum free energy.

[‡]P value based on randomisation test implemented in PARESnip2.

[‡]Library in which this tag was detected; M = mock library, I = infected library.

5.3.5 Identification of conserved miRNAs in *Brassica napus*

Several miRNAs are evolutionarily conserved in the plant kingdom [60]. We assessed whether conserved *B. napus* miRNAs were expressed during *S. sclerotiorum* infection. Therefore, all six clean libraries were searched against miRBase (Release 22.1). From our libraries, we identified 73 conserved miRNA families with 529 mature miRNA sequences. We found that 61 miRNA families had more than one sequence while 12 miRNA families had only one mature sequence predicted (Fig 5.4A). Among these miRNA families, miR156 had 42 isomiR sequences followed by miR159, and miR166 with 28 and 25 sequences, respectively. Most of these miRNA sequences were 21 nt long, followed by 20, 19 and 18 nt (Fig5. 4B). There was a 5' uracil bias in 18-22 nt long miRNA sequences, which agreed with previously reported results in miRNA studies in different plant species (Fig 5.4C). Using degradome sequencing, we found that from the 73 conserved miRNA families, 718 and 1,406 cleaved products (Supplementary Table 5.3) were obtained from infected and mock libraries, respectively. Four levels of degradome cleavage site confidence, based on read mapping characteristics, are described in [47, 56]. Category 0 is the most confident, followed by categories 1, 2 and 3.

Among the 718 cleavage events in the infected sample, 507 were in category 0, followed by category 2, 3, and 1 with 90, 87, and 34 events, respectively, based on the abundance of fragment transcripts in the library. Thus, most of the conserved miRNA targets identified with degradome sequencing were of relatively high confidence [61]. Several target genes were likely silenced by more than one miRNA family. Among these target genes, 158 non-redundant transcripts were found in infected samples. Among the 158 targets in infected samples, miR160, miR164, miR167, and miR396 were predicted to target more than 10 genes each. Similarly, miR156, miR6030, miR400, miR393, miR172 and miR171 were predicted to target eight genes (Supplementary Table 5.4). We found an additional 43 conserved miRNA families compared with *B. napus* miRNAs recorded in miRBase. These miRNAs were reported previously in several studies to regulate gene expression in plants during biotic and abiotic stress [62-65].

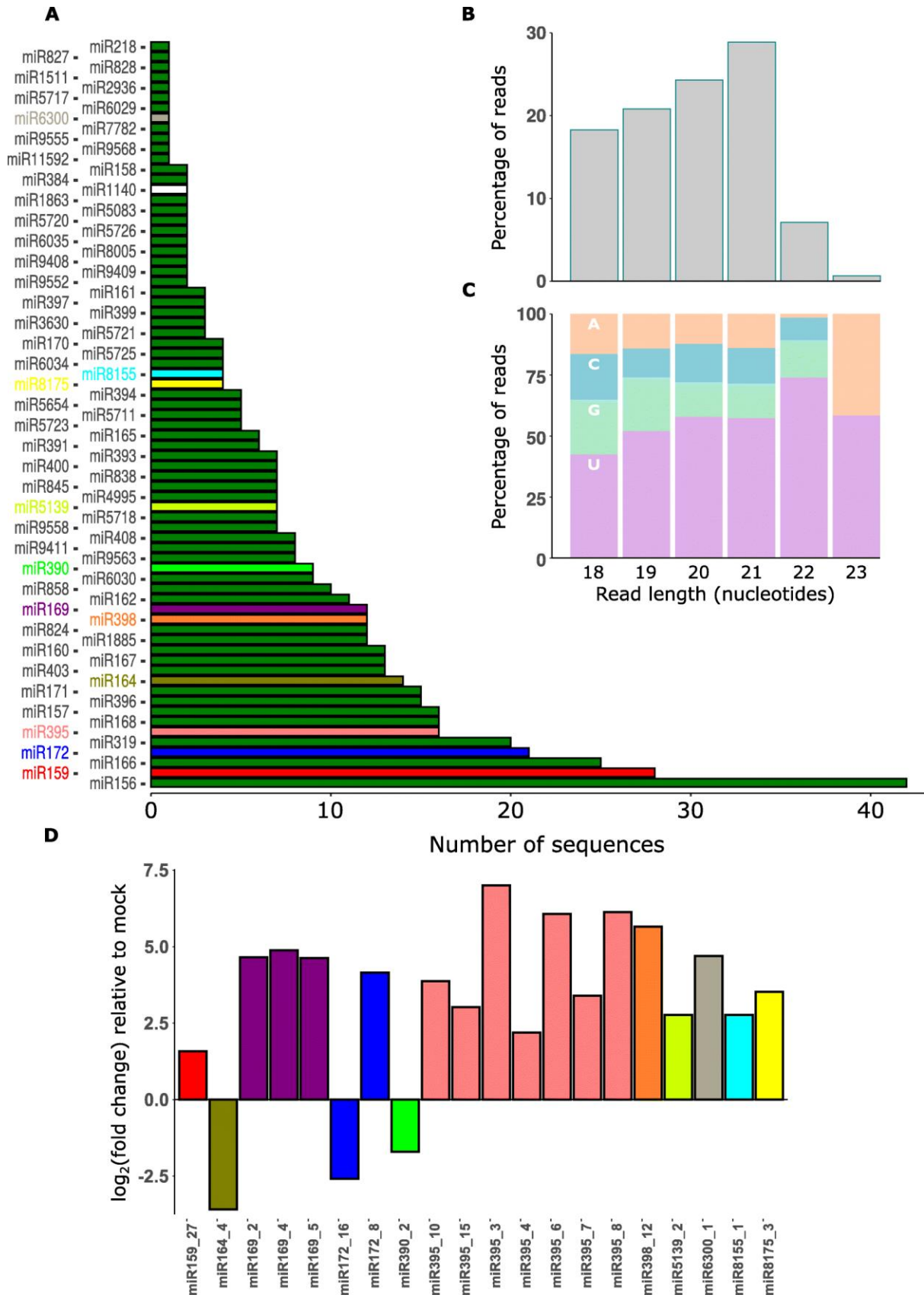


Fig. 5. 4. Prediction of infection-responsive microRNAs from the *Brassica napus* genome. (A) Histogram of the 73 conserved miRNA families. The y-axis shows the identified conserved miRNA family name and the x-axis shows the number of sequences (isomiRs) identified for each miRNA family; red bars are the significantly differentially expressed miRNA families. An isomiR is one of a family of highly similar miRNA sequences derived from either the guide or passenger strand. Green miRNA families were not differentially expressed. Those in other colours contained differentially expressed miRNAs and the colours correspond with D. (B) Histogram of read sizes from 529 conserved miRNAs. The y axis depicts the percentage of reads and the x-axis read length in nucleotides. (C) Histogram of the 529 miRNAs showing percentage of reads (y-axis) of each size class (x-axis) that had each of the four nucleotides (AGCU) in their 5' position. (D) 20 differentially expressed miRNAs with $\log_2(\text{fold change})$ on the y-axis. Colours correspond with miRNA families in A.

5.3.6 Pathogen responsive miRNAs cleave plant immune response genes in the infected sample

To determine whether any degradation of transcripts was specific to the infected samples, we filtered out all the genes from the infected sample that were also targeted in the mock sample, resulting in 50 targets (Supplementary Table 5.5). Altogether, 172 InterPro domains were found in these genes. These genes had functions such as transcriptional regulation, disease resistance, and posttranscriptional gene silencing. We found several miRNAs that were shown to have a role in plant and pathogen interactions from this set also.

5.3.7 Seventeen miRNAs belonging to 9 miRNA families were significantly upregulated during infection

We performed differential expression analysis on the individual conserved miRNAs with raw read counts from the miRprof analysis. Among the conserved miRNAs, only 20 miRNAs were

differentially expressed (Fig 5.4D; Supplementary Table 5.6). Among these, seventeen miRNAs belonging to 9 miRNA families were upregulated during infection and three miRNAs were downregulated. Among these upregulated 9 miRNA families, miR395 had 7 isomiRs, miR169 had 3 isomiRs while miR159, miR172, miR398, miR6300, miR8155, miR8175, and miR5139 had one isomiR each. The up-regulated miRNA sequences exhibited $\log_2(\text{fold change})$ values of between 1.58 and 6.13. A total of 16 of the 17 miRNAs had a $\log_2(\text{fold change})$ of more than 2.

The three downregulated miRNAs belonged to the miRNA families miR164, miR72, and miR390. These miRNAs exhibited $\log_2(\text{fold change})$ values during infection of between -1.7 and -3.59; two exhibited $\log_2(\text{fold change})$ values below -2. Interestingly, miR172 had two isomiRs with different expression patterns, with one upregulated and the another downregulated during infection.

Among the 17 upregulated miRNAs, we found 5 genes potentially cleaved by a member of the miR159 family, miR159_27 ($\log_2(\text{fold change}) = 1.56$; P-adjusted = 0.034), and a member of the miR5139 family, miR5139_2 ($\log_2(\text{fold change}) = 2.77$; P-adjusted = 0.0042), in the infected samples. Surprisingly, we found 3 transcripts cleaved by downregulated miR390 member, miRNA390_2 ($\log_2(\text{fold change}) = -2.94$; P-adjusted = 0.04), in the infected sample, while there was no cleavage of these genes in the mock samples.

5.3.8 RNA structure-aided prediction algorithms identify 135 novel *Brassica napus*

micro RNA loci

After filtering out the exact matches of conserved miRNAs to miRBase, 135 novel miRNA producing loci that did not have any hits in miRBase were identified from the *B. napus* genome.

Among these miRNAs, 67 loci were found to have both passenger strand (miRNA* or ‘star’) and mature strand reads revealing the confidence of these novel miRNAs as per annotation criteria. A detailed description of the novel miRNA loci is given in Supplementary Table 5.7.

From our study, we did not find any cleavage signal from the novel miRNAs predicted from miRDeep2. We used the same set of miRNAs to predict targets using the psRNA target server. From psRNA target, 12,104 genes were putatively targeted by these miRNAs. Several psRNA targets might be false positives since it is entirely based on a theoretical *in silico* procedure, whilst a degradome signal is a better reflection of the biological cleavage. It remains to be confirmed whether these novel miRNAs have genuine targets or not.

5.3.9 Nine *Brassica napus* PHAS loci are differentially expressed in response to *Sclerotinia sclerotiorum* infection

PHAS loci have not been very well characterised in *B. napus*. Therefore, we aimed to identify expressed PHAS loci in the *B. napus* genome from our sequencing data set. We found 26 PHAS loci in the *B. napus* genome. The genes associated with predicted PHAS loci were annotated by aligning PHAS locus sequences to the NCBI Nucleotide Collection (nr/nt). Among the 26 PHAS genes, about half were related to disease resistance proteins (five genes), non-coding RNAs (five genes), and chloroplast related (three genes). In addition, single genes were found for metal tolerance, pentatricopeptide repeat, cop9 signalosome complex subunit, and photosystem II protein D1. Nine PHAS genes were not homologous to any sequences in NCBI. A total of 182 pha-siRNAs were produced from these loci. Among these siRNAs, 41 were highly expressed, with a read abundance of more than 100 reads.

Since miRNAs are key triggers of pha-siRNA expression, we used the psRNA target server to find the cleavage sites in PHAS loci from the conserved miRNAs we identified. We found six PHAS loci potentially triggered by conserved miRNAs (Table 5.3). All excised PHAS clusters with their corresponding pha-siRNAs are shown in Supplementary File 5.1. The miR390-triggered PHAS gene *TAS3* was found to be conserved across different species. In this study, we found two genes possibly targeted by miR390, one of which had sequence similarity to *TAS3* in *A. thaliana*.

We were able to confirm likely cleavage of one of these TAS loci during infection using degradome sequencing (category 2, $P = 0.0019$; Fig 5.5A). This locus was triggered by conserved miR 1885. However, we did not find degradome signals for another five miRNA-triggered loci which were predicted to be cleaved with conserved miRNAs as evident while using psRNA target server. Recently, it has been shown experimentally that miR1885 plays a key role in targeting PHAS loci residing within NBS-LRR genes to trigger ta-siRNA production [30]. Accordingly, we found that this locus had homology to NBS-LRR proteins. We identified 10 likely ta-siRNAs produced from this TAS locus. From the degradome signal, one of the ta-siRNAs silenced a gene BnaC05g49720D, a galactose oxidase, beta-propeller, (category 2, $P = 0.016$; Fig 5.5B) in the infected sample. Possibly, *B. napus* miRNAs regulate gene expression in response to *S. sclerotiorum* infection through the production of miRNA triggered ta-siRNAs.

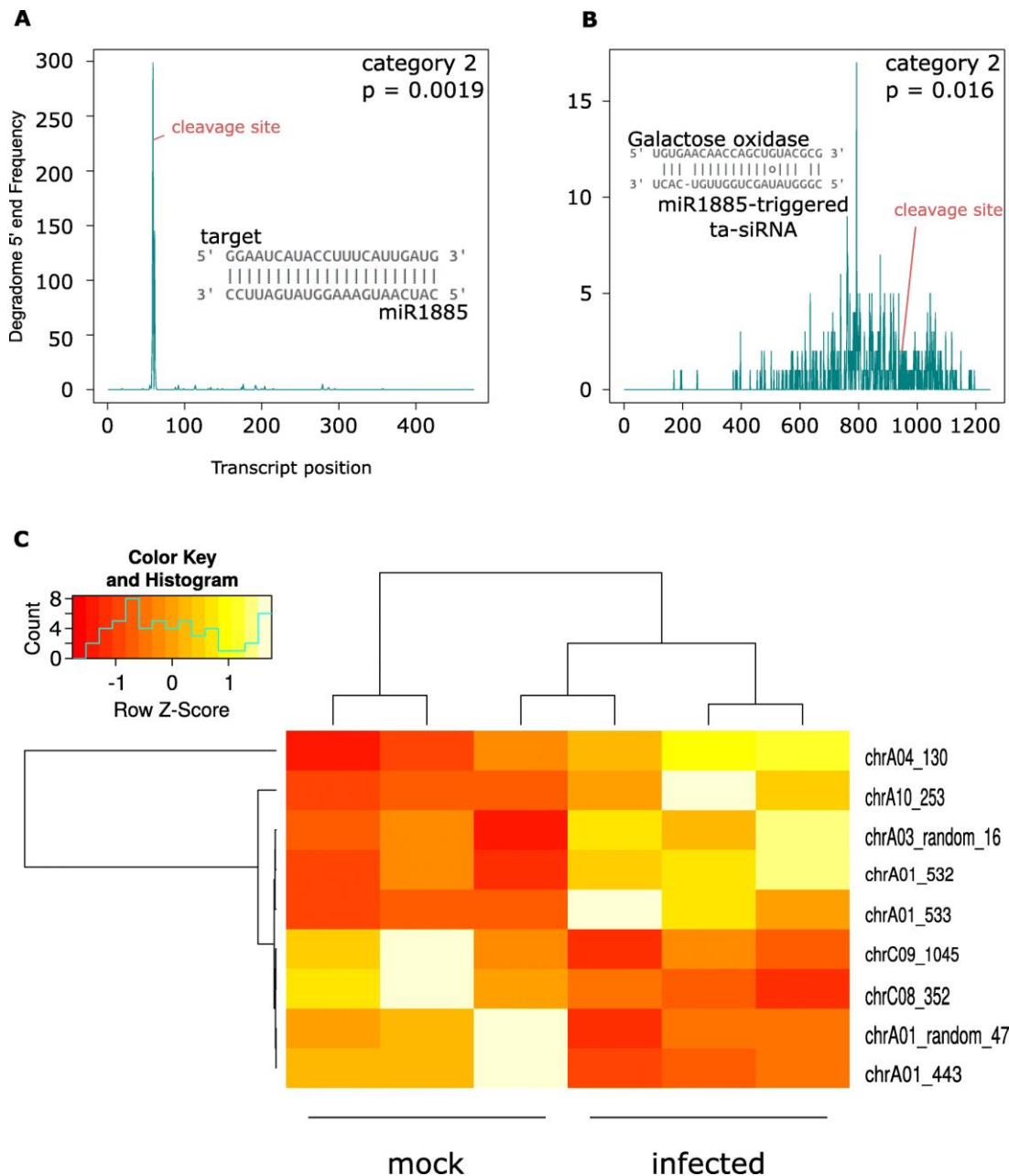


Fig. 5. 5. . Degradome validation of a TAS gene triggered by miRNA1885. (A) Target plot (T-plot) of the cleavage site on a TAS gene putatively targeted by miRNA1885. The transcript position in the target gene is on the x-axis and the y-axis shows the 5' read coverage at each position; the cleavage site predicted by PARESnip2 is labelled in red. The category and p value of the PARESnip2 test are given in the top left-hand corner of the graph. (B) T-plot of cleavage site on a galactose oxidase gene putatively targeted by one of the ta-siRNAs produced by the miRNA 1885-triggered TAS gene. The x-axis shows the transcript position in the target gene and the y-axis shows read 5' coverage at each position; the cleavage site predicted by PARESnip2 is labelled in red. The category and

p value from PARESnip2 are given in the top right-hand corner of the graph. (C) A heat map of 9 differentially expressed PHAS loci plotted with normalized counts from DESeq2.

Differential expression analysis of PHAS loci showed five loci were upregulated and four were downregulated during infection. Among upregulated loci, three were related to disease resistance with a log₂fold change ranging from 0.78 to 1.87. The remaining two genes were chloroplast and photosystem II protein D1 with a log₂fold change of 0.57 and 1.73, respectively. Among downregulated loci, two loci were non-coding RNAs with a log₂fold change of -0.7 and -0.96, one was related to COP9 signalosome complex subunit (log₂fold change= (-0.82) and the remaining one was not characterized. Fig 5.5C shows a heat map of nine differentially expressed PHAS loci.

To gain a global overview of genes targeted by pha-siRNAs we used the psRNA target server to find the targets of 41 highly expressed pha-siRNAs. We found 5,918 transcripts that might be regulated by this class of sRNA. We did GO term enrichment analysis of these targets and found regulation of several biological processes (Supplementary Table 5.8). The terms ‘posttranscriptional gene silencing’ (GO:0035194), ‘cellular potassium ion homeostasis’ (GO:0030007), ‘regulation of ARF protein signal transduction’ (GO:0032012) and ‘threonyl-tRNA amino acylation’ (GO:0006435), ‘oxidation-reduction process’ (GO:0055114), and ‘regulation of transcription DNA-templated’ (GO:0006355), ‘carbohydrate metabolic process’ (GO:0005975)’ were significantly enriched.

Table 5.3. An overview of the characteristics of PHAS loci identified using PhaseTank.

PHAS locus ID	Length	Phased ratio	Phased abundance	Phased number	Phased score	Triggering miRNA*	Description of BLAST hit	Differential expression
chrCnn_random_466	791	0.398	3845	16	52.538		Uncharacterized	
chrA04_130	917	0.448	274	19	47.821	miR4995, miR6035	chloroplast	
chrC01_614	331	0.568	3417	7	32.332		Uncharacterized	Upregulated
chrA03_random_16	434	0.6	630	8	30.96		Disease resistance protein	Upregulated
chrA01_443	245	0.79	877	5	26.769	miR390	non coding RNA	Downregulated
chrUnn_random_22	621	0.468	159	11	26.078		chloroplast	
chrC09_754	621	0.468	159	11	26.078		chloroplast	
chrA01_random_47	245	0.769	877	5	26.05		non coding RNA	Downregulated
chrA07_478	287	0.621	823	6	25.004		metal tolerance protein	
chrA01_532	371	0.666	111	7	21.953		Disease resistance protein	Upregulated
chrC09_1169	392	0.747	1217	4	21.216	miR838	Uncharacterized	
chrA02_413	455	0.381	2189	7	20.533		Uncharacterized	
chrAnn_random_753	308	0.531	547	6	20.073	miR390	non coding RNA	
chrA09_223	476	0.314	579	10	19.96	miR1885	Disease resistance protein**	
chrC05_811	392	0.393	559	8	19.894	miR838	non coding RNA	
chrA01_533	329	0.746	288	4	16.888		Disease resistance protein	Upregulated
chrC02_807	371	0.668	109	5	15.662		Disease resistance protein	
chrAnn_random_716	413	0.7	182	4	14.562		Uncharacterized	
chrA10_253	329	0.376	198	7	13.935		Photosystem II protein D1	Upregulated
chrC01_385	266	0.444	237	5	12.145		non coding RNA	
chrC09_1045	266	0.4	152	5	10.041		COP9 signalosome complex subunit	Downregulated
chrC09_1044	434	0.373	788	4	9.949		Pentatricopeptide repeat	
chrCnn_random_695	391	0.35	234	5	9.558		Uncharacterized	
chrA04_52	413	0.359	206	5	9.558		Uncharacterized	
chrC08_352	392	0.46	118	4	8.772		Uncharacterized	Downregulated

*Based on psRNA target comparison of conserved miRNA sequences with all PHAS loci.

**Evidence for cleavage from degradome sequencing.

We also specifically investigated the targets of the miR1885-triggered ta-siRNAs. We found 1,601 targets of these sRNAs with psRNA target. GO term enrichment analysis showed that these ta-siRNAs possibly regulate protein phosphorylation (GO: 0006468), transcription factors (GO: 0045944), vesicle mediated transport proteins (GO:0016192), and fucose metabolic pathway genes (GO:0006004) (Supplementary Table 5.9).

5.3.10 Further analysis of sRNA targeting using 5' rapid amplification of cDNA ends and quantitative PCR

5' RACE was used to find putative cleavage sites in the novel sRNA target gene *BnaA01g27570D*, which is an ethylene response factor. This gene was chosen as ethylene signalling has a well-documented role in plant immunity to pathogens [66] and, although ethylene response factors have a demonstrated role in response to *S. sclerotiorum* [67], there is little understanding of how they might be regulated by sRNAs. This novel siRNA expressed from chromosome A01 is 22 nt long, not conserved or characterised, is neither a pha-siRNA nor a miRNA. Fig 5.6A shows the 5'-RACE products of the predicted cleavage site from the infected sample. A T-plot of the cleavage signal is shown in Fig 5.6C. We also assessed the expression of this target gene during infection using RT-qPCR. The average $\log(2^{-\Delta C_t})$ value calculated for mock and infected sample was 0.76 and -0.09 respectively with a standard deviation of 0.35 and 0.53 Fig 5.6C. The combined degradome, 5'RACE, and qPCR results showed that a novel *B. napus* sRNA likely regulates an ethylene response factor gene during

S. sclerotiorum infection, leading to a decrease in its expression. Overall, the integrated degradome, 5'RACE, and RT-qPCR results showed the regulation of a plant immune response gene by a novel siRNA.

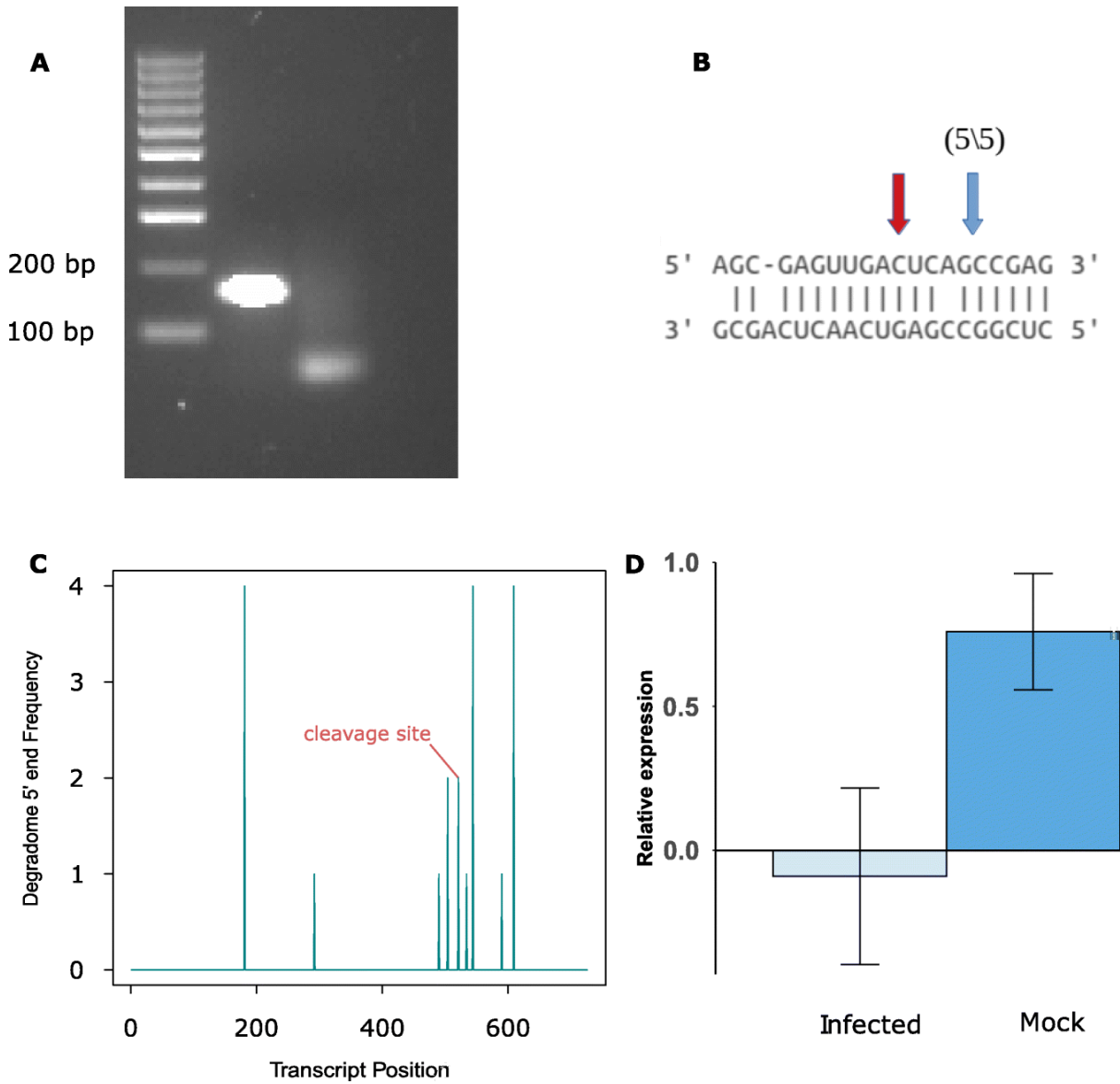


Fig. 5. 6. 5'Rapid amplification of cDNA ends (RACE), degradome result, and qPCR for an ethylene response factor gene putatively cleaved by a novel sRNA. (A) Gel electrophoresis of the 5'RACE result showing a band of the correct size; the second lane is a no template control. (B) Sequence complementarity of the sRNA and its target. The blue arrow shows the cleavage site identified from sequencing the 5'RACE product and the red arrow shows the cleavage site identified with degradome sequencing. (C) Target-plot (T-plot) of the degradome result of the 5'RACE validated gene showing the transcript position in the target gene (x-axis) and read 5' coverage at

each position (y-axis); the cleavage site predicted by PARESnip2 is labelled in red. **(D)** RT-qPCR result of target gene in mock and infected sample (x-axis) and relative expression of gene to the house keeping actin gene.

5.4 Discussion

Small RNA-omics studies have revealed the tight regulation of host immune pathways in plants [23, 68, 69]. In our study, we identified different classes of sRNAs in *B. napus* plants that responded to infection with *S. sclerotiorum* and showed how they may be involved in regulating different sets of genes using degradome sequencing.

We found evidence of sRNA-mediated regulation of an ethylene response factor gene (*BnaA01g27570D*), which we further investigated using RT-qPCR and 5'-RACE. The likely cleavage site identified through degradome sequencing and 5'-RACE and concomitant reduction in expression of this gene during infection suggest that it has a role in the plant response to pathogen attack. Enhanced ethylene production is an early response of plants to perception of pathogen attack, leading to induction of defence systems [70]. In *B. napus*, ethylene responsive element binding factors have been predicted to control biological processes related to defence signalling, secondary metabolite production and redox regulation. The closest homologue of *BnaA01g27570D* in *Arabidopsis thaliana*, RAP2.3, has been shown to co-localise in the nucleus with another ethylene response factor, ORA59, to mediate defence to the bacterial pathogen *Pectobacterium carotovorum* [71]. Since RAP2.3 has a known positive role in disease resistance and ethylene is generally found to be a positive contributor to defence, it is intriguing that our study showed a reduction in expression of *BnaA01g27570D* caused by potential sRNA-mediated regulation during infection. Several hypotheses could be put forward to explain this. For example, the pathogen could elicit responses in the plant that dampen ethylene-responsive immunity. Alternatively, some components of the ethylene response system could negatively regulate immunity, as shown in some other systems [66].

Further studies are required to determine the biological significance of sRNA cleavage of this particular ethylene response factor upon *S. sclerotiorum* infection of *B. napus*.

The distribution of size classes of total sRNAs was different between mock and infected samples. The size class of sRNAs gives insights into their biogenesis, for example 21 nt sRNAs are processed by DCL1 and DCL4, whereas 22 nt sRNAs are formed through the action of DCL2; 24 and 26 nt sRNAs are formed by DCL3 [72]. Similarly, sRNAs with 5' uracil, adenine or cytosine are loaded into AGO1, AGO2 and AGO4 or AGO5, respectively [73]. Furthermore, it has been reported previously that variation in sRNA lengths also has effects on downstream function of sRNAs [74]. In the mock sample, both redundant and non-redundant distributions peaked at 24 nt. However, in the infected sample, the redundant distribution had a major peak at 21 nt and the non-redundant distribution had peak at 24 nt. Overall, our data suggest that a different set of sRNA biogenesis pathways is initiated upon infection with *S. sclerotiorum*.

A total of 730 sRNA loci were upregulated in *B. napus* in response to *S. sclerotiorum* infection. The upregulated sRNAs mostly belonged to the size classes of 20 and 21 nt with a 5' bias of uracil, whereas downregulated sRNAs were overrepresented for 24 nt sequences with varied 5' biases. This suggests that upon pathogen infection, the host recruits different DCL and AGO proteins for subsequent gene silencing. The upregulated sRNAs likely silenced many genes related to stress signaling, as determined from the degradome data. This suggests that the enhancement of disease symptoms might be accompanied with the negative regulation of plant immune genes during *S. sclerotiorum* infection of *B. napus*.

The miRNAs are sRNAs which are produced from short hairpin precursors. Based on these criteria, thousands of miRNAs have been identified and deposited in miRBase from many plants. The majority of these miRNAs shows high conservation. However, only 92 mature

Brassica napus miRNAs have been discovered so far. These are less numerous than those of *A. thaliana* (428), *M. truncatula* (756), and *O. sativa* (738) revealing several miRNAs are yet to be discovered in this species.

Two previous studies have assessed *S. sclerotiorum* responsive *B. napus* miRNAs at time points 3, 12, and 48 HPI [36, 37]. A total of 227 [36] and 77 [37] conserved miRNA sequences were identified in these studies. Both of these studies were conducted with a single sRNA sequencing library per sample. The microarray miRNA expression analysis resulted in detection of 68 infection-responsive miRNAs, [36] and 10 of these were further analyzed with stem loop qPCR. Similarly, 10 miRNAs were found to be differentially expressed in Jian et al [37], while only one miR166 was found to be commonly differentially expressed with respect to Cao et al study [36]. Here, we found 529 mature conserved miRNAs belonging to 73 miRNA families, along with 20 infection-induced miRNAs based on a replicated differential expression analysis. The enhanced number of miRNAs in our study might be due to the replicated dataset and increased sequencing depth. We did not find any common differentially expressed miRNAs in comparison to the two previous studies [36, 37]. This might be due to the different time points, *B. napus* variety, tissue collection and *S. sclerotiorum* strains we used. Nevertheless, the degradome data suggested these miRNAs regulate expression of transcription factors related to development and defense responses, which corroborate previous findings.

Moreover, 135 novel miRNA loci were discovered with 67 loci that had both mature and star (passenger strand) read counts. Although this adds to the overall number of *B. napus* miRNAs identified, we were not able to identify any likely targets of these miRNAs. This suggests that they either do not have targets or that they regulate genes through a non-cleavage mechanism such as inhibition of translation [75].

Expression of a large number of transcription factors and auxin signaling pathway genes was likely regulated by these identified conserved miRNAs based on the infected sample degradome data. Some of these miRNAs had multiple targets in specific classes. We found 50 unique cleaved products from the infected sample that were not present in the mock samples. These genes were related to transcription factors, disease resistance proteins, and posttranscriptional silencing. For example, miR824, miR390/miR5083, miR403/miR838, miR5139, miR1885, cleaved transcripts of genes containing leucine rich repeats, zinc finger transcription factors, protein kinases, and disease resistance protein-encoding genes, respectively. Similarly, miR166 and miR858 cleaved transcripts of homeodomain and sant/myb domain-containing genes, respectively. MiR824 was shown previously in the Brassicaceae to have a role in the heat stress response [76]. It is worth mentioning here miR858 has been shown to negatively regulate MYB transcription factors, thereby controlling resistance to pathogen infection in *A. thaliana* [77]. Similarly, plant homeodomain proteins, which are potentially regulated by miR166 in this study, control transcriptional regulation of pathogen defense-related genes [78]. These findings suggest that miRNAs are involved in regulation of multiple aspects of the immune response of *B. napus* to *S. sclerotiorum*.

From our study, we found 20 infection-induced miRNAs. These miRNAs were previously shown to have roles in stress responses in different pathosystems [68, 79-81]. However, we only found degradome evidence of cleavage of the transcripts of five genes, *BnaA03g22590D*, *BnaA05g27620D*, *BnaCnng31260D*, *BnaCnng49390D*, and *BnaA02g31560D*, which were targeted by the upregulated miRNAs miR159_27 and miR5139_2 in the infected samples. Four of these five genes were potentially cleaved by miR159_27 and the remaining gene by miR5139_2. These genes contained InterPro domains such as SANT/myb domain, homeodomain, and zinc finger domain. Surprisingly, miR390_2, whose expression was ~69 %

($\log_2(0.31) = -1.7$) lower in the infected sample had three target genes, whose mRNAs were cleaved only in the infected samples (*BnaA06g09370D*, *BnaC05g49670D*, *BnaC08g16190D*). It suggests that these genes were not expressed in the mock sample.

On the other hand, we found 11 genes whose transcripts were cleaved in the mock sample from different isomiRs of upregulated miR395, while we did not find any of these cleaved genes in the infected samples. These data suggest that gene regulation by sRNAs is quite complex and the cleavage events are not always dependent on the expression pools and other variables might also come in play.

MiR390 has been shown to have a role in the formation of ta-siRNAs and regulate auxin response factor genes [82] while miR395 plays an important role in sulphur assimilation [83]. MiR159 is present in the majority of land plants where it regulates genes encoding R2R3 MYB domain transcription factors that transduce the gibberlin (GA) signal [84] and have roles in growth, development, and biotic stress responses. It was shown recently that cotton and *Arabidopsis* plants accumulate increased levels of miR159 in response to the fungus *Verticillium dahliae* [85]. miR159 was also reported to have a role in *Arabidopsis* root galls under the attack of root knot nematodes, since lines lacking the miR159-Gibberlin acid MYB pathway had better resistance to root knot nematode. Moreover, the miR159-GA MYB pathway has been shown in *A. thaliana* to promote the programmed cell death response [86]. Our finding of four MYB domain genes with transcripts cleaved by miR159 suggested that miR159 has a role in *B. napus* responses to *S. sclerotiorum*. However, further investigation will be needed to understand the precise role of this pathway.

In previous studies, miR5139 was shown to be regulated by ethylene in petal growth and was first detected in perennial herbs [87] but no specific functions were allocated to this miRNA.

Later it was also shown to have tissue-specific expression in wheat [88]. However, no homolog of miR5139 was reported in *B. napus* to date. Here, we found the expression level of miR5139 was nearly 7 times higher ($\log_2(6.96) = 2.8$) during infection and it was found to cleave the transcript of the gene *BnaA02g31560D* which encodes a zinc finger domain-containing gene. Zinc finger domains are reported to be present in plant resistance related proteins that are involved in effector triggered immune responses [89]. Our data shows that under the influence of *S. sclerotiorum* attack, *B. napus* deploys miR5139 to negatively regulate Zinc finger domain-containing genes as a defence response strategy [89].

Although we did not find any degradome cleavage signals for the other differentially expressed miRNAs, there were some interesting miRNAs identified here which were shown previously to have a role in plant responses against pathogens. miR164 was shown to manipulate programmed cell death in *A. thaliana*. Overexpression of miR164 target genes enhanced disease symptoms [90]. miR169 was found to negatively regulate rice immunity genes during infection by the rice blast fungus [68]. miR6300 and miR8175 were highly upregulated in *Alternaria*-treated tomato plants compared to the control with a $\log_2(\text{fold change})$ of 3.13 or 2.65. Here, we found a 3.5 and 4.7 fold increase of these two miRNAs after *S. sclerotiorum* infection [80].

With the aid of the degradome library we showed that miR1885 can trigger a disease resistance TAS gene which subsequently produces 10 ta-siRNAs for gene regulation. This particular miRNA was recently shown to directly silence the *B. napus* TIR-NBS-LRR resistance gene *BraTNL1* and the TAS gene *BraTIR1*, cleavage of which generates sRNAs that regulate the photosynthesis-related gene *BraCP24* [30]. By regulating both immunity and basal growth, this miRNA may be essential for optimizing resource allocation during development. This miRNA has been shown to be expressed at a low abundance under most conditions apart from flowering time, when the plant requires synergistic reductions in the levels of photosynthesis

and pathogen response. In our study, we did not observe a change in expression of miR1885 in response to *S. sclerotiorum* infection, although we observed potential cleavage of a galactose oxidase, beta propeller protein-encoding transcript (*BnaC05g49720D*) by one of the small RNAs derived from the TIR-NBS-LRR TAS gene it likely activates. However, such proteins have quite varied roles [91, 92], so it is not possible to come to any conclusions on the biological significance of this observation. Identification and elucidation of the regulatory network of pha-siRNAs is important so that the expression of these pha-siRNAs can be controlled by changing the expression of their miRNA triggers. This strategy could be useful to modulate the degree of silencing of endogenous and exogenous target genes.

In conclusion, our comprehensive data set allowed us to investigate overall pathogen-responsive RNA interference-based regulation of host transcripts and the actions of specific small RNA classes in response to pathogen attack. Our data suggest that targets of infection induced sRNAs including miRNAs may be associated with stress response genes. *B. napus* plants may differentially express both pha-siRNAs and conserved miRNAs when challenged with a necrotrophic pathogen. An integrated analysis revealed that a *B. napus* sRNA regulates an ethylene response factor gene during pathogen attack. Ethylene response factors regulate several jasmonate (JA) and (ET) pathways and are key players in plant innate immunity [50]. Our combined degradome, 5' RACE and qPCR results showed that the expression of one of the ethylene response genes is suppressed in *B. napus* after *S. sclerotiorum* infection. The silencing of this gene was mediated by a novel sRNA which was not characterized before.

5.5 Availability of data and materials

The smallRNA and degradome sequencing data has been deposited in GenBank under BioProject PRJNA678586.

5.6 References

1. Borges, F. and R.A. Martienssen, *The expanding world of small RNAs in plants*. Nature reviews Molecular Cell Biology, 2015. **16**(12): p. 727-741.
2. Ozata, D.M., et al., *PIWI-interacting RNAs: small RNAs with big functions*. Nature Reviews Genetics, 2019. **20**(2): p. 89-108.
3. Guleria, P., et al., *Plant small RNAs: biogenesis, mode of action and their roles in abiotic stresses*. Genomics, Proteomics & Bioinformatics, 2011. **9**(6): p. 183-199.
4. Pratt, A.J. and I.J. MacRae, *The RNA-induced Silencing Complex: A Versatile Gene-silencing Machine*. Journal of Biological Chemistry, 2009. **284**(27): p. 17897-17901.
5. Liu, W.-w., et al., *Characterization and function of MicroRNA *s in plants*. Frontiers in Plant Science, 2017. **8**: p. 2200.
6. Hackenberg, M., et al., *Characterization of phosphorus-regulated miR399 and miR827 and their isomirs in barley under phosphorus-sufficient and phosphorus-deficient conditions*. BMC Plant Biology, 2013. **13**(1): p. 1-17.
7. Trevisan, S., et al., *Expression and tissue-specific localization of nitrate-responsive miRNAs in roots of maize seedlings*. Plant, Cell & Environment, 2012. **35**(6): p. 1137-1155.
8. Zheng, Y., et al., *A dynamic evolutionary and functional landscape of plant phased small interfering RNAs*. BMC Biology, 2015. **13**: p. 15.

9. Xie, M. and B. Yu, *siRNA-directed DNA Methylation in Plants*. Current Genomics, 2015. **16**(1): p. 23-31.
10. Howell, M.D., et al., *Genome-wide analysis of the RNA-DEPENDENT RNA POLYMERASE6/DICER-LIKE4 pathway in Arabidopsis reveals dependency on miRNA-and tasiRNA-directed targeting*. The Plant Cell, 2007. **19**(3): p. 926-942.
11. Zheng, Y., et al., *A dynamic evolutionary and functional landscape of plant phased small interfering RNAs*. BMC Biology, 2015. **13**(1): p. 32.
12. de Felippes, F.F., et al., *A single miR390 targeting event is sufficient for triggering TAS3-tasiRNA biogenesis in Arabidopsis*. Nucleic Acids Research, 2017. **45**(9): p. 5539-5554.
13. Felippes, F.F. and D. Weigel, *Triggering the formation of tasiRNAs in Arabidopsis thaliana: the role of microRNA miR173*. EMBO Reports, 2009. **10**(3): p. 264-270.
14. Wu, F., et al., *Genome-wide identification and characterization of phased small interfering RNA genes in response to Botrytis cinerea infection in Solanum lycopersicum*. Scientific Reports, 2017. **7**(1): p. 1-10.
15. Zhang, C., et al., *Identification of trans-acting siRNAs and their regulatory cascades in grapevine*. Bioinformatics, 2012. **28**(20): p. 2561-2568.
16. Zhai, J., et al., *MicroRNAs as master regulators of the plant NB-LRR defense gene family via the production of phased, trans-acting siRNAs*. Genes & Development, 2011. **25**(23): p. 2540-2553.
17. Deng, P., et al., *Biogenesis and regulatory hierarchy of phased small interfering RNAs in plants*. Plant Biotechnology Journal, 2018. **16**(5): p. 965-975.
18. Jian, H., et al., *Identification of rapeseed microRNAs involved in early stage seed germination under salt and drought stresses*. Frontiers in Plant Science, 2016. **7**: p. 658.

19. Achard, P., et al., *Modulation of floral development by a gibberellin-regulated microRNA*. *Development*, 2004. **131**(14): p. 3357-3365.
20. Guo, H.-S., et al., *MicroRNA directs mRNA cleavage of the transcription factor NAC1 to downregulate auxin signals for Arabidopsis lateral root development*. *The Plant Cell*, 2005. **17**(5): p. 1376-1386.
21. Liu, Q., et al., *Expression analysis of phytohormone-regulated microRNAs in rice, implying their regulation roles in plant hormone signaling*. *FEBS Letters*, 2009. **583**(4): p. 723-728.
22. Navarro, L., et al., *A plant miRNA contributes to antibacterial resistance by repressing auxin signaling*. *Science*, 2006. **312**(5772): p. 436-439.
23. Du, J., et al., *Identification of microRNAs regulated by tobacco curly shoot virus co-infection with its betasatellite in Nicotiana benthamiana*. *Virology Journal*, 2019. **16**(1).
24. Islam, W., et al., *Plant microRNAs: Front line players against invading pathogens*. *Microbial Pathogenesis*, 2018. **118**: p. 9-17.
25. Gupta, O.P., et al., *MicroRNA regulated defense responses in Triticum aestivum L. during Puccinia graminis f. sp. tritici infection*. *Molecular Biology Reports*, 2012. **39**(2): p. 817-824.
26. Liang, G., Q. Ai, and D. Yu, *Uncovering miRNAs involved in crosstalk between nutrient deficiencies in Arabidopsis*. *Scientific Reports*, 2015. **5**: p. 11813.
27. Sánchez-Sanuy, F., et al., *Osa-miR7695 enhances transcriptional priming in defense responses against the rice blast fungus*. *BMC Plant Biology*, 2019. **19**(1): p. 1-16.
28. Peláez, P. and F. Sanchez, *Small RNAs in plant defense responses during viral and bacterial interactions: similarities and differences*. *Frontiers in Plant Science*, 2013. **4**: p. 343.

29. Shivaprasad, P.V., et al., *A microRNA superfamily regulates nucleotide binding site–leucine-rich repeats and other mRNAs*. *The Plant Cell*, 2012. **24**(3): p. 859-874.
30. Cui, C., et al., *A Brassica miRNA Regulates Plant Growth and Immunity through Distinct Modes of Action*. *Molecular Plant*, 2020. **13**(2): p. 231-245.
31. Carré, P. and A. Pouzet, *Rapeseed market, worldwide and in Europe*. *OCL*, 2014. **21**(1): p. 1-12.
32. Del Rio, L., et al., *Impact of Sclerotinia stem rot on yield of canola*. *Plant Disease*, 2007. **91**(2): p. 191-194.
33. Fu, Y., et al., *MicroRNA-mRNA expression profiles and their potential role in cadmium stress response in Brassica napus*. *BMC Plant Biology*, 2019. **19**(1): p. 1-20.
34. Wei, W., et al., *Small RNA and degradome profiling involved in seed development and oil synthesis of Brassica napus*. *PLoS One*, 2018. **13**(10): p. e0204998.
35. Shen, D., et al., *Identification and characterization of microRNAs 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*, 2014. **204**(3): p. 577-94.
36. Cao, J.Y., et al., *Tight regulation of the interaction between Brassica napus and Sclerotinia sclerotiorum at the microRNA level*. *Plant Molecular Biology*, 2016. **92**(1-2): p. 39-55.
37. Jian, H., et al., *Integrated mRNA, sRNA, and degradome sequencing reveal oilseed rape complex responses to Sclerotinia sclerotiorum (Lib.) infection*. *Scientific Reports*, 2018. **8**(1): p. 1-17.
38. Denton-Giles, M., et al., *Partial stem resistance in Brassica napus to highly aggressive and genetically diverse Sclerotinia sclerotiorum isolates from Australia*. *Canadian Journal of Plant Pathology*, 2018: p. 1-11.

39. Ma, Z., C. Coruh, and M.J. Axtell, *Arabidopsis lyrata* small RNAs: transient MIRNA and small interfering RNA loci within the *Arabidopsis* genus. *The Plant Cell*, 2010. **22**(4): p. 1090-1103.
40. Martin, M., *Cutadapt removes adapter sequences from high-throughput sequencing reads*. *EMBnet. Journal*, 2011. **17**(1): p. 10-12.
41. Andrews, S., *FastQC: a quality control tool for high throughput sequence data*. 2010, Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom.
42. Derbyshire, M., et al., *The Complete Genome Sequence of the Phytopathogenic Fungus Sclerotinia sclerotiorum Reveals Insights into the Genome Architecture of Broad Host Range Pathogens*. *Genome Biology and Evolution*, 2017. **9**(3): p. 593-618.
43. Chalhoub, B., *Early allopolyploid evolution in the post-Neolithic Brassica napus oilseed genome (vol 348, 1260782, 2014)*. *Science*, 2014. **345**(6202): p. 1255-1255.
44. Bushnell, B.B.A.F., *Accurate, Splice-Aware Aligner*. United States. , *BBMap: A Fast, Accurate, Splice-Aware Aligner*. United States. 2014.
45. Axtell, M.J., *ShortStack: comprehensive annotation and quantification of small RNA genes*. *RNA*, 2013. **19**(6): p. 740-751.
46. Mohorianu, I., et al., *The UEA small RNA workbench: a suite of computational tools for small RNA analysis*, in *MicroRNA Detection and Target Identification*. 2017, Springer. p. 193-224.
47. Friedländer, M.R., et al., *miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades*. *Nucleic Acids Research*, 2012. **40**(1): p. 37-52.
48. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2*. *Genome Biology*, 2014. **15**(12): p. 550.

49. Guo, Q., X. Qu, and W. Jin, *PhaseTank: genome-wide computational identification of phasiRNAs and their regulatory cascades*. *Bioinformatics*, 2015. **31**(2): p. 284-286.
50. Kalvari, I., et al., *Non-coding RNA analysis using the Rfam database*. *Current Protocols in Bioinformatics*, 2018. **62**(1): p. e51.
51. Nawrocki, E.P. and S.R. Eddy, *Infernal 1.1: 100-fold faster RNA homology searches*. *Bioinformatics*, 2013. **29**(22): p. 2933-2935.
52. Dai, X. and P.X. Zhao, *psRNATarget: a plant small RNA target analysis server*. *Nucleic Acids Research*, 2011. **39**(suppl_2): p. W155-W159.
53. Thody, J., et al., *PAREsnip2: a tool for high-throughput prediction of small RNA targets from degradome sequencing data using configurable targeting rules*. *Nucleic Acids Research*, 2018. **46**(17): p. 8730-8739.
54. Fahlgren, N. and J.C. Carrington, *miRNA target prediction in plants*, in *Plant MicroRNAs*. 2010, Springer. p. 51-57.
55. Alexa, A., *Rahnenfuhrer, J. topGO: enrichment analysis for gene ontology (R package version 2.40.0)*. Bioconductor. 2020.
56. Huang, P.-Y., J. Catinot, and L. Zimmerli, *Ethylene response factors in Arabidopsis immunity*. *Journal of Experimental Botany*, 2016. **67**(5): p. 1231-1241.
57. Omidvar, V., et al., *Identification of miRNAs with potential roles in regulation of anther development and male-sterility in 7B-1 male-sterile tomato mutant*. *Bmc Genomics*, 2015. **16**(1): p. 1-16.
58. Schwach, F., et al., *Deciphering the diversity of small RNAs in plants: the long and short of it*. *Briefings in Functional Genomics and Proteomics*, 2009. **8**(6): p. 472-481.
59. Ma, X., J. Wiedmer, and J. Palma-Guerrero, *Small RNA bidirectional crosstalk during the interaction between wheat and Zymoseptoria tritici*. *Frontiers in Plant Science*, 2020. **10**: p. 1669.

60. Luo, X., et al., *Identification of miRNAs and their target genes in peach (Prunus persica L.) using high-throughput sequencing and degradome analysis*. PLoS One, 2013. **8**(11): p. e79090.
61. Addo-Quaye, C., W. Miller, and M.J. Axtell, *CleaveLand: a pipeline for using degradome data to find cleaved small RNA targets*. Bioinformatics, 2008. **25**(1): p. 130-131.
62. Cui, J., et al., *Characterization of miRNA160/164 and their targets expression of beet (Beta vulgaris) seedlings under the salt tolerance*. Plant Molecular Biology Reporter, 2018. **36**(5-6): p. 790-799.
63. Wang, Y., et al., *MicroRNA167-directed regulation of the auxin response factors GmARF8a and GmARF8b is required for soybean nodulation and lateral root development*. Plant Physiology, 2015. **168**(3): p. 984-999.
64. Lu, Y., et al., *MiR393 and miR390 synergistically regulate lateral root growth in rice under different conditions*. BMC Plant Biology, 2018. **18**(1): p. 261.
65. Wu, G., et al., *The sequential action of miR156 and miR172 regulates developmental timing in Arabidopsis*. Cell, 2009. **138**(4): p. 750-759.
66. Pieterse, C.M., et al., *Hormonal modulation of plant immunity*. Annual Review of Cell and Developmental Biology, 2012. **28**: p. 489-521.
67. Girard, I.J., et al., *RNA sequencing of Brassica napus reveals cellular redox control of Sclerotinia infection*. Journal of Experimental Botany, 2017. **68**(18): p. 5079-5091.
68. Li, Y., et al., *Osa-miR169 negatively regulates rice immunity against the blast fungus Magnaporthe oryzae*. Frontiers in Plant Science, 2017. **8**: p. 2.
69. Soto-Suárez, M., et al., *The Arabidopsis miR396 mediates pathogen-associated molecular pattern-triggered immune responses against fungal pathogens*. Scientific Reports, 2017. **7**: p. 44898.

70. van Loon, L.C., B.P. Geraats, and H.J. Linthorst, *Ethylene as a modulator of disease resistance in plants*. Trends in Plant Science, 2006. **11**(4): p. 184-191.
71. Kim, N.Y., Y.J. Jang, and O.K. Park, *AP2/ERF family transcription factors ORA59 and RAP2. 3 interact in the nucleus and function together in ethylene responses*. Frontiers in Plant Science, 2018. **9**: p. 1675.
72. Ossowski, S., R. Schwab, and D. Weigel, *Gene silencing in plants using artificial microRNAs and other small RNAs*. The Plant Journal, 2008. **53**(4): p. 674-690.
73. Kim, V.N., *Sorting out small RNAs*. Cell, 2008. **133**(1): p. 25-26.
74. Wu, L., et al., *DNA methylation mediated by a microRNA pathway*. Molecular Cell, 2010. **38**(3): p. 465-75.
75. Li, S., et al., *MicroRNAs inhibit the translation of target mRNAs on the endoplasmic reticulum in Arabidopsis*. Cell, 2013. **153**(3): p. 562-574.
76. Szaker, H.M., et al., *miR824/AGAMOUS-LIKE16 Module Integrates Recurring Environmental Heat Stress Changes to Fine-Tune Poststress Development*. Frontiers in Plant Science, 2019. **10**: p. 1454.
77. Camargo-Ramírez, R., B. Val-Torregrosa, and B. San Segundo, *MiR858-mediated regulation of flavonoid-specific MYB transcription factor genes controls resistance to pathogen infection in Arabidopsis*. Plant and Cell Physiology, 2018. **59**(1): p. 190-204.
78. Korfhage, U., et al., *Plant homeodomain protein involved in transcriptional regulation of a pathogen defense-related gene*. The Plant Cell, 1994. **6**(5): p. 695-708.
79. Zhang, Q., et al., *Md-miR156ab and Md-miR395 target WRKY transcription factors to influence apple resistance to leaf spot disease*. Frontiers in Plant Science, 2017. **8**: p. 526.

80. Sarkar, D., et al., *Integrated miRNA and mRNA expression profiling reveals the response regulators of a susceptible tomato cultivar to early blight disease*. DNA Research, 2017. **24**(3): p. 235-250.
81. Gao, F., et al., *Identification of drought-responsive microRNAs and their targets in *Ammopiptanthus mongolicus* by using high-throughput sequencing*. Scientific Reports, 2016. **6**(1): p. 1-16.
82. Xia, R., J. Xu, and B.C. Meyers, *The emergence, evolution, and diversification of the miR390-TAS3-ARF pathway in land plants*. The Plant Cell, 2017. **29**(6): p. 1232-1247.
83. Matthewman, C.A., et al., *miR395 is a general component of the sulfate assimilation regulatory network in Arabidopsis*. FEBS Letters, 2012. **586**(19): p. 3242-3248.
84. Millar, A.A., A. Lohe, and G. Wong, *Biology and Function of miR159 in Plants*. Plants, 2019. **8**(8): p. 255.
85. Zhang, T., et al., *Cotton plants export microRNAs to inhibit virulence gene expression in a fungal pathogen*. Nature Plants, 2016. **2**(10): p. 1-6.
86. Alonso-Peral, M.M., et al., *The MicroRNA159-Regulated GAMYB-like Genes Inhibit Growth and Promote Programmed Cell Death*.
87. Yang, Y., et al., *Differential miRNA expression in *Rehmannia glutinosa* plants subjected to continuous cropping*. BMC Plant Biology, 2011. **11**(1): p. 1-11.
88. Pandey, R., et al., *A comprehensive genome-wide study on tissue-specific and abiotic stress-specific miRNAs in *Triticum aestivum**. PLoS One, 2014. **9**(4): p. e95800.
89. Cabot, C., et al., *A role for zinc in plant defense against pathogens and herbivores*. Frontiers in plant science, 2019. **10**: p. 1171.
90. Lee, M.H., et al., *An Arabidopsis NAC transcription factor NAC4 promotes pathogen-induced cell death under negative regulation by microRNA164*. New Phytologist, 2017. **214**(1): p. 343-360.

91. Chen, C.K.-M., N.-L. Chan, and A.H.-J. Wang, *The many blades of the β -propeller proteins: conserved but versatile*. Trends in Biochemical Sciences, 2011. **36**(10): p. 553-561.
92. Šola, K., et al., *RUBY, a putative galactose oxidase, influences pectin properties and promotes cell-to-cell adhesion in the seed coat epidermis of Arabidopsis*. The Plant Cell, 2019. **31**(4): p. 809-831.

Chapter 6: Generation of disruption mutants of RNAi machinery genes in *Sclerotinia sclerotiorum*

Abstract

RNA silencing is an integral phenomenon in eukaryotes that functions in gene regulation. RNAi has varied roles in the growth, development and pathogenicity of fungi. This study aimed to understand the roles of RNAi related genes in *Sclerotinia sclerotiorum* through generation of gene disruption mutants. Altogether, 24 RNAi related genes were found in the *S. sclerotiorum* genome. Disruption mutants for the three genes *Dicer1*, *Dicer2* and *RdRp* were generated. All disruption mutants showed normal mycelium growth compared to wild type strains. However, the number of sclerotia formed was reduced. *Brassica napus* leaf infection assays showed a reduction in growth for one of the independent *dicer1* strains, while two other *dicer1* strains exhibited a similar phenotype to the wild type. This study did not result in production of homokaryotic knockout strains for *Dicer2* or *RdRp*, although several heterokaryotic strains were developed. The findings of this study further expanded the knowledge on the RNAi machinery of *S. sclerotiorum* and their possible role in fungal growth and pathogenicity.

6.1 Introduction

RNA interference (RNAi) is a conserved mechanism facilitated by small RNAs which transcriptionally and post-transcriptionally regulate gene expression [1-3]. sRNAs are 20-30 nucleotide (nt) non-coding RNA molecules that trigger sequence-specific silencing of transcripts [4]. This phenomenon is conserved in the vast majority of eukaryotic organisms; however, specific RNA silencing pathways, the number and the nature of proteins involved in the process vary among different organisms [5]. The major components of the RNA silencing pathway comprise of Dicer and the slicer protein called Argonaute. Dicer is an RNase III

endonuclease that produces 20-30 nt sRNA duplexes from double-stranded RNA (dsRNA) precursors. Argonaute, the effector component of RNAi, binds to one of the strands of a sRNA duplex and forms the RNA-induced silencing complex (RISC). The siRNA guides the RISC to homologous mRNAs. In addition to this, RNA-dependent RNA polymerase (RdRp) induces the silencing process by generating secondary siRNAs or producing double stranded RNA (dsRNA) from single stranded RNA (ssRNA) [6, 7]. Secondary siRNAs, in contrast to primary siRNAs, are derived from the targeted mRNAs rather than the initial triggering molecules. Most fungal species investigated so far harbour two or more *Dicer*, *Argonaute* and *RdRp* genes. These genes are often called *Dicer*-like, *Argonaute*-like, and *RdRp*-like in fungi as their evolutionary relationship to animal and plant homologues is unclear; this thesis uses the terms *Dicer*, *Argonaute*, and *RdRp* for simplicity. Species such as *Schizosaccharomyces pombe* [8], *Cryptococcus neoformans* [9], and *Aspergillus nidulans* [10] only contain a single gene encoding each of the RNAi proteins. Other *Aspergillus* species possess at least two genes encoding RNAi components, suggesting partial or sporadic loss of the RNAi machinery genes in some fungal species [11]. Furthermore, in some fungi a large number of RNA silencing genes are present, probably due to gene or genome duplication. For example, the basidiomycetes *Coprinopsis cinerea* and *Phanerochaete chrysosporium* have seven *RdRp* genes each and eight and five *Dicer* genes, respectively [12]. On the other hand, RNAi machinery has been completely lost in species like *Saccharomyces cerevisiae* [13], *Ustilago maydis* [14], and *Cryptococcus deuterogattii* [15].

Bernstein et al. [16] investigated the role of *Dicer* in RNAi pathways for the first time in the model organism *Drosophila melanogaster*. *Dicer* is an endonuclease that belongs to the ribonuclease III (RNase III) family. Based on the RNase III domain architecture, *Dicer* proteins can be grouped into four major classes [17]. Class I *Dicers* contain double stranded RNA

binding domains (dsRBDs) and undergo homodimerization to bring two functional RNase III domains together and are present in bacteria, bacteriophages and some lower eukaryotes. In class II RNase III domains, an N-terminal domain is present which is required for homodimerization. This class is mainly involved in the processing of precursors of ribosomal RNA (rRNA), small nuclear RNA (snRNA), and small nucleolar RNA (snoRNA). Class III enzymes possess an additional RNase III domain along with a dsRNA binding domain (dsRBD), which enables them to function as monomers. The Drosha protein is a typical example of a class III Dicer which processes miRNA precursors in eukaryotic cells. The higher-class IV Dicercs comprise a helicase domain, two RNase III domains, two dsRBDs, and a Piwi/Ago/PAZ domain. Proteins in the Argonaute protein family comprise four conserved domains: an N-terminal domain, the Mid domain, Piwi and PAZ domain (RNA-binding domain) [18].

As previously mentioned, fungi may harbour different numbers of *Dicer* and *Argonaute* homologs [19, 20]. However, not all of them are involved in RNAi silencing pathways. In *Fusarium graminearum*, only one of two *Argonaute* genes is involved in silencing of viral nucleic acids [21]. In *Neurospora crassa*, QDE-2, a fungal Argonaute homolog is involved in silencing of repetitive elements such as transgenes [22]. Fungal RdRps have been proposed to generate long RNAs which can form Dicer substrates [23]. In plants, it is well known that RdRp recruits degraded mRNA to form secondary siRNAs [24]. However, a detailed role of RdRp in RNA silencing in fungi has not been investigated. One study demonstrated the function of RdRp in a *Mucor* species [4], which showed the involvement of two distinct RdRp enzymes in the RNAi process.

Initially, RNAi was identified to play a key role as a defensive mechanism against foreign nucleic acids in viruses, plants and fungi [21, 25, 26]. Later, its role in regulating various biological processes including the response to biotic and abiotic stress was reported [27-30]. The disruption mutants do not display lethal phenotypes [30-32]. Different studies conducted on non-pathogenic and pathogenic fungi have revealed varied sRNA biogenesis pathways [33]. While some of the RNA silencing genes could play dual roles [34], other genes can have redundant functions, such as defence against viruses, and processing of dsRNA or transgenes [21, 35].

Sclerotinia sclerotiorum is a broad host range pathogen causing diseases in more than 400 plant species [36]. It is considered one of the most economically damaging *Brassica napus* pathogens [37]. Recent observations have shown that the RNA silencing pathway can be used to control fungal pathogens. For example, the virulence of the fungal pathogen *Botrytis cinerea* has been suppressed by silencing of two *Dicer* genes [35]. A similar observation was also made in the reference isolate of *S. sclerotiorum* (1980) [38]. However, these studies did not investigate the role of RdRp. In this chapter, we demonstrated the presence of the RNAi machinery in the aggressive Australian *S. sclerotiorum* strain CU8.24. Furthermore, knockout strains lacking *Dicer* and *RdRp* genes were generated and growth *in vitro* and *in planta* investigated. In addition, an *Argonaute* knockout construct was generated for future investigation into *Argonaute* function in *S. sclerotiorum*.

6.2 Materials and Methods

6.2.1 Identification of RNAi machinery genes in the *S. sclerotiorum* CU8.24 genome

A translated nucleotide Basic Local Alignment Search Tool (tBLAST) search of RNAi-related genes in the reference strain of *S. sclerotiorum* 1980 [39] was conducted using the query sequences for *M. circinelloides* RdRp and *N. crassa* Dicer1, Dicer2 and Argonaute. The accessions for Dicer1, Dicer2, Argonaute, and RdRp were XP_961898.1, XP_963538.3, AF217760.1, and, XP_959047.1, respectively. Afterwards, two *Dicers* (sscle_10g079850 and sscle_16g109590), one *Argonaute* (sscle_03g027950) and one *RdRp* (sscle_07g057040) gene from 1980 were aligned against the CU8.24 genome [40] using the map to reference tool with Geneious Prime v13 and sequences from the CU8.24 loci to which they mapped were used to generate constructs for knockout experiments. *Dicer*, *Argonaute* and *RdRp* were chosen based on their major roles in sRNA biosynthesis pathways in previous reported literature [4, 38, 41]. Phylogenetic analysis of Dicer1, Dicer2, Argonaute and RdRp were constructed by extracting their amino acid sequences from funRNA for well characterized ascomycetes (database <http://funrna.riceblast.snu.ac.kr/index.php?a=view>).

6.2.1 Amplification of RNAi gene flanking sequences

At first, approximately 1 kb of both the 5' upstream and of 3' downstream sequences were amplified from each of the four genes, except 5' upstream for *Dicer2* which was 649 bp, using standard polymerase chain reaction (PCR). An overlapping 50 bp region was also added in the flanking sequence primers which facilitated the preparation of knockout constructs using Gibson assembly. The PCR was conducted using 2x Superfi enzyme (Invitrogen) in a 50 µL reaction. The PCR steps comprised of 98°C initial denaturation for 30 seconds, followed by 35 cycles of denaturation at 98°C for 10 seconds, annealing at 55°C for 10 seconds, and extension

at 72 °C for 1.5 min, followed by a final extension step at 72 °C for 3 min. The PCR reactions were conducted in an Applied Biosystems thermocycler (Life Technologies).

6.2.2 Assembly of knockout plasmid constructs

Gibson assembly was used for the construction of transformation plasmids [42]. The amplified RNAi gene flanking sequences along with 50 bp overlapping regions for Gibson assembly were gel purified with a GeneJet extraction kit (Thermo Fisher Scientific). JDC193 TAR-0 (hyg) plasmid was used as a transformation vector kindly provided by Johannes Debler (CCDM) (Fig. 6.1). The above transformation vector which carried a *hygromycin B phosphotransferase* gene (*hyg*) was digested with the *KpnI* restriction enzyme to generate two fragments, the vector backbone and the hygromycin B resistance cassette. The digested vector was run on an agarose gel by electrophoresis and the two fragments were excised and gel purified. These two gel purified fragments along with two flanking sequences were used for Gibson assembly. The concentration required for the assembly of each fragment was calculated using NEBioCalculator (<https://nebiocalculator.neb.com/#!/ligation>). The four fragments were assembled with NEBuilder HiFi DNA assembly Master Mix (BioLabs, New England, Australia) following the manufacturer's protocol and incubated at 50°C for one hour. The assembled products were run on an agarose gel by electrophoresis to see the presence of the assembled product. Once the expected assembled product size was confirmed, 2 µL of chilled product was used to transform NEB 5-alpha chemically competent cells (BioLab). The transformed competent cells were plated on lysogeny broth (LB) agar plates supplemented with kanamycin (30 µ/ml) and incubated overnight at 37°C.

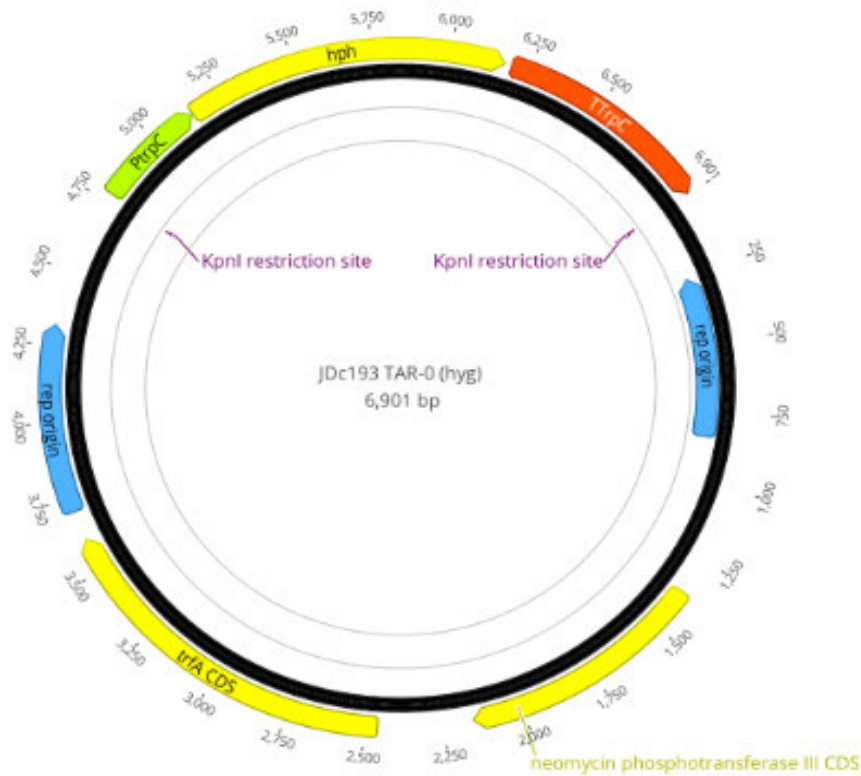


Fig. 6. 1. Plasmid used for the assembly of transformation constructs showing the restriction site, hygromycin cassette and background. The yellow coloured arrows represent hygromycin cassette for fungal selction, neomycin phosphotransferase III coding sequence, and trfA (trans-acting product essential for vegetative plasmid replication), the green coloured arrows represent P_{trpC}, is the *Aspergillus nidulans* tryptophan biosynthesis protein constitutive promoter to drive expression of the hph gene, the red coloured arrows represent T_{trpC}, a tryptophan terminator from *Aspergillus nidulans*, and blue the coloured arrows represent origin of replication.

6.2.3 Confirmation of correctly assembled constructs

The confirmation of correctly assembled constructs was conducted with colony PCR, restriction digestion, and Sanger sequencing. For colony PCR, 8-10 independent clones were used with the same flanking primers that were used for the amplification of flanking sequences of genes. Only those clones that amplified both 5' and 3' flanking regions were selected for plasmid isolation. Furthermore, the purified plasmids were used for restriction digest

diagnostic tests. Restriction digestion was carried out with different restriction enzymes: *KpnI* and *BspHI* based on the plasmid map in the Geneious Prime v13 Plasmids with expected band patterns following restriction digests were further confirmed with Sanger sequencing. The sequencing result was analysed with Geneious software by aligning with the original knockout construct sequence.

6.2.3 PEG-mediated protoplast transformation

The PEG-mediated transformation method was utilized to transfer the gene deletion cassettes into *S. sclerotiorum* CU8.24 protoplasts as described in Chapter 2.

6.2.4 Screening of transformants

S. sclerotiorum does not produce any conidia; therefore, repeated hyphal tip transfer onto selective media was conducted to purify transformants and generate homokaryotic strains. A maximum of ten independent transformants were selected for purification by hyphal tip transfer. Hyphal tip regeneration was performed for six rounds and sclerotia were allowed to develop; these were subsequently stored at 4°C for further use. We used standard PCR to confirm that homologous recombination had occurred with a combination of different primers (Supplementary Table 6.1).

6.2.5 Phenotypic analysis of transformants

Three independent strains each of *dicer1*, *dicer2* and *rdrp* were compared with the wild type CU8.24 for hyphal growth and sclerotia formation. For the *in vitro* phenotype tests, 6th generation sclerotia were cut into halves and germinated in PDA plates supplemented with hygromycin B for the knockout strains whereas the wild type strain was grown in only PDA

plates without hygromycin B. Once the sclerotia germinated, three mm mycelial plugs were used to subculture mycelia onto fresh PDA plates for all strains and incubated at 22°C. The growth of mycelia was checked from these plates after 24 hours, by taking two perpendicular measurements of the colony diameter with a ruler and averaging them. For sclerotia counting, the mycelia were grown with a similar method and the plates were incubated for two weeks and sclerotia number from each plate was counted. For each *in vitro* experiment, there were three replicates and data were analysed using analysis of variance followed by Tukey's honest significant difference (HSD) *post hoc* test using R studio version 3.6.

For *in planta* infection assays, either first or second true leaves of six-week-old *B. napus* (cultivar Av Garnet) plants were inoculated with a two-day-old mycelial plug taken from the leading edge of colonies on PDA plates. After infection, plants were kept in a large plastic box and covered with another plastic box to enhance humidity, facilitating the infection process. Altogether, 9-10 independent plants were infected for each strain. The lesions were measured across their vertical axes at 24 and 48 HPI with a ruler and the difference between wild type and knockout strains was assessed using an analysis of variance followed by Tukey's HSD.

6.3 Results

6.3.1 Presence of RNAi machinery genes in the *Sclerotinia sclerotiorum* genome

Altogether, 24 proteins in the *S. sclerotiorum* reference strain 1980 were found to have significant sequence similarity to characterised RNAi proteins from *M. circinelloides* RdRp and *N. crassa* Dicer1, Dicer2 and Argonaute (. We focused on four genes, two *Dicer*, one *Argonaute*, and one *RdRp* for knockout experiments in the *S. sclerotiorum* CU8.24 strain (Supplementary Fig 6.1) as the translated amino acids of these gene predictions had the highest sequence similarity with the BLAST query sequences (Supplementary Table Table 6.2). The

Dicers and Argonaute were chosen as BLAST queries as they have been experimentally characterised in *N. crassa* and *M. circinelloides* [4, 41].

A translated nucleotide BLAST search confirmed the presence of a protein with homology to QDE2, an Argonaute protein of *N. crassa* (AF217760.1). This protein (sscle_03g027950) encodes an Argonaute protein in the reference strain of *S. sclerotiorum* 1980. A homology search using the nucleotide sequence for this gene model revealed the presence of this gene in scaffold 18 of the CU8.24 strain. The Dicer1 of *N. crassa* (XP_961898.1) was homologous to sscle_10g079850 in *S. sclerotiorum* 1980 and was found to be present in scaffold 77 of CU8.24. Similarly, Dicer2 (XP_963538.3) in *N. crassa* had homology to sscle_16g109590 in the *S. sclerotiorum* reference strain and was found to be present in scaffold 194 in CU8.24. Furthermore, the RdRp protein from *N. crassa* (XP_959047.1) had homology to sscle_07g057040 in the reference strain and was present in scaffold 50 in CU8.24. The percentage identity between the nucleotide sequences of two strains was approximately 96 % for Dicers and Argonaute, and 100 % for RdRp. Subsequently, the expression of these genes using data from the study of Seifbarghi et al [40] across seven time points during infection of *B. napus* was assessed. The data indicated that the genes were expressed with read counts ranging from three to 385 counts per million.

We further compared the RNAi machinery in several well characterized ascomycetes such as *N. crassa*, *F. oxysporum*, *B. cinerea*, *A. versicolor*, *Schizosaccharomyces* and conducted phylogenetic analysis to see which Sclerotinia genes correspond to which genes in other species. Altogether 38 protein sequences were retrieved from the funRNA database. Interestingly, similar molecules were grouped together as evident from Fig. 6.2. *B. cinerea* is closely match to *S. sclerotiorum* as expected.

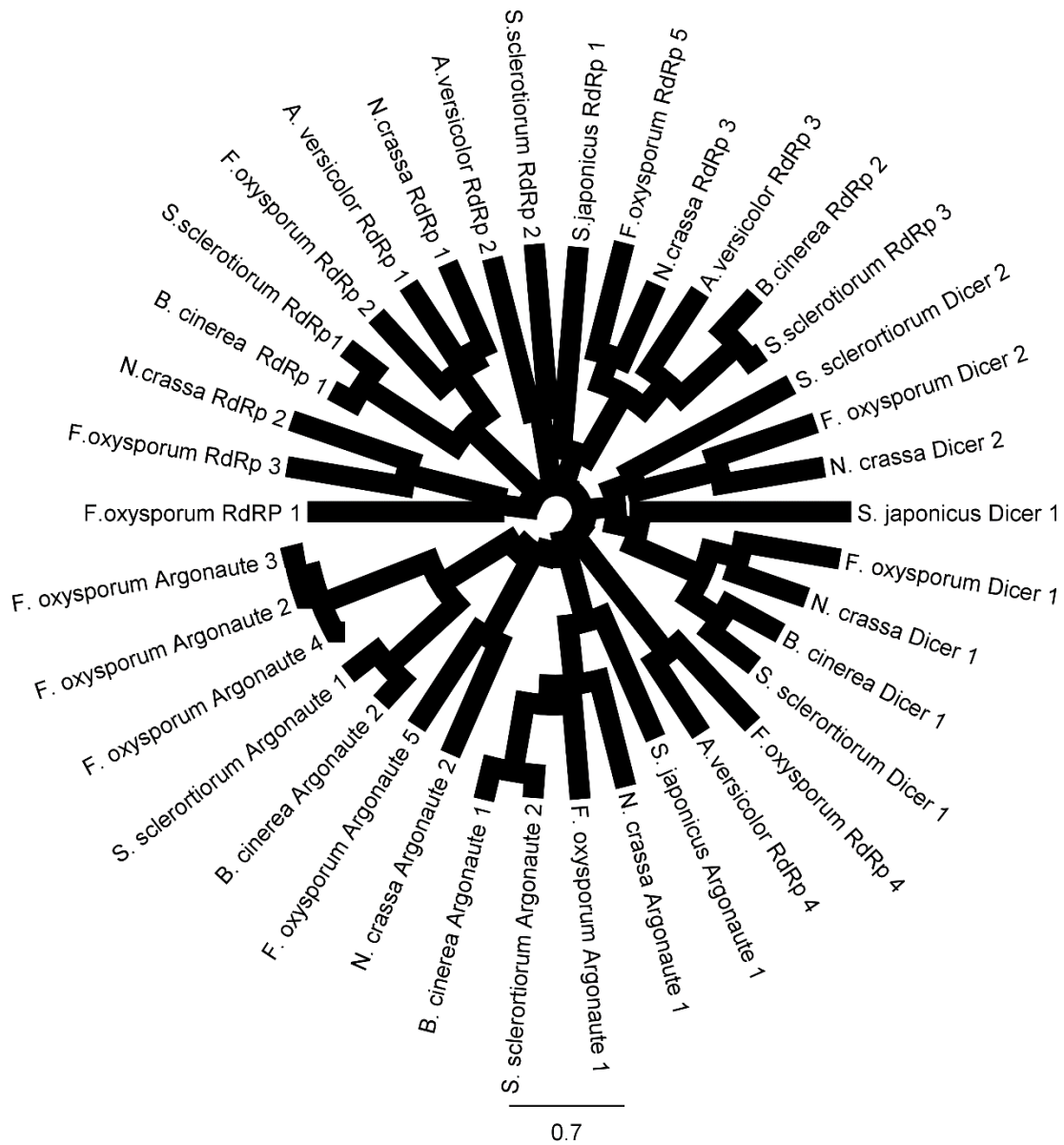


Fig. 6. 2. Neighbor-joining consensus tree based on the alignment of RNAi protein sequences of 5 filamentous fungi. The protein sequences were retrieved from funRNA database.

6.3.2 Confirmation of knockout plasmid constructs

We used the Gibson assembly method for the assembly of plasmid constructs. Successful assembly, therefore, should give the product size equivalent to the sum of all assembled products. Supplementary Fig. 6.1 shows gel images acquired in the generation of the *Dicer1*

knockout construct. For example, the assembled products were examined by gel electrophoresis in order to check for successful assembly. As expected, the assembled product size was visualized. The 5' and 3' flanking sequences were 1,050 bp for *RdRp*, the hygromycin B resistance cassette and the vector backbone had sizes of 2,125 bp and 4,784 bp, respectively. The correctly assembled *Dicer1* knockout construct, therefore, should have the size of 9,009 bp. Similar results were observed for assembly of the other knockout constructs. We also conducted colony PCR and restriction digestion for the confirmation of putative true constructs. A colony PCR resulted in eight colonies carrying the *Dicer1* knockout construct, where only two colonies amplified both 5' and 3' flanking sequences. These colonies were, therefore, selected for restriction digest analysis. Restriction digestion tests were conducted for positive colony PCR clones using *KpnI* and *BspHI* restriction enzymes; the product of the correctly assembled construct had two fragments of size 4,784 bp and 4,135 bp after *KpnI* digestion and three fragments of size of 6,339 bp, 2,414 bp and 170 bp after *BspHI* digestion, respectively (Supplementary Fig. 6.2). Furthermore, Sanger sequencing also validated the correct assembly product (Fig. 6.3). Together, these techniques confirmed the correctly assembled transformation constructs.

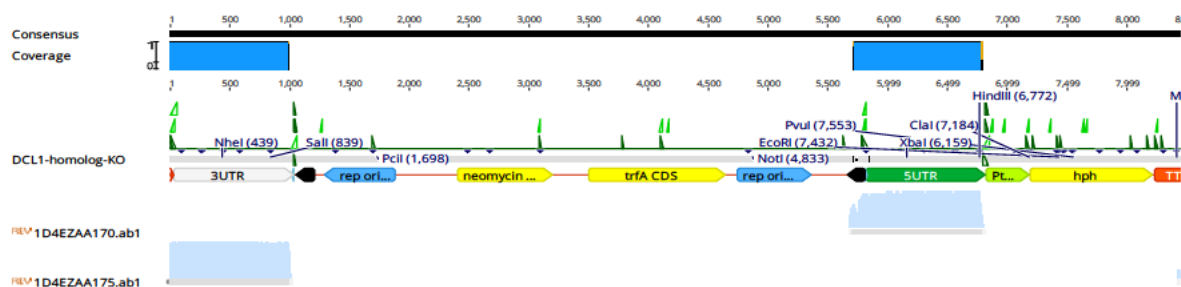


Fig. 6. 3. Confirmation of true construct with Sanger sequencing. Consensus is the original construct and coverage is 5' left and 3' right flank mapped with the tool map to reference in Geneious prime. The aligned file is sequenced file obtained from the sanger sequencing.

6.3.3. Disruption mutants of *Dicer* and *RdRp* genes

Although knockout constructs were generated for four aforementioned genes in this thesis, the knockout experiments were only conducted for *Dicer1*, *Dicer2*, and *RdRp*. Knockout of the *Argonaute* gene was not conducted as during cross-kingdom RNAi, sRNAs are loaded into the plant Argonaute. However, deletion of the *Argonaute* gene is in process at the CCDM to investigate its role in endogenous gene regulation.

From the screening of ten transformants, four independent homokaryotic knockout strains were identified for the *Dicer1* gene target, and no homokaryotic knockout strains were identified for the *Dicer2* and *RdRp* gene targets, although several heterokaryons that harboured a mix of wild-type and transformed nuclei were identified (Supplementary Fig 6.3). Purification of strains by hyphal tip transfer did not produce pure homokaryotic strains with *RdRp* and *Dicer2* genes deleted. However, this project has developed several heterokaryotic knockout strains which could be a useful resource for future work to develop pure gene deletion strains.

We used several primer combinations for the confirmation of knockout strains. Figure 6.6 shows the primer pairs used for the confirmation of *Dicer1* gene knockout and electrophoresis gel image confirmed *Dicer1* gene disruption using Dicer_1_5'F and Dicer_1_3'R. The wild type gave an amplicon size of 8,680 bp while the deletion strains had a size of 4,570 bp. Primer pairs Dicer 1_5'F, Dicer_1_5'R and Dicer 1_3'F, Dicer 1_3'R were also used for the amplification of the left and right insertion flanks, respectively. A similar technique was used for the confirmation of knock out events for *Dicer2* and *RdRp* genes.

Furthermore, primers specific for the wild type locus and primers specific to the wild type outside of the inserted flank (Dicer 1_5'F) and the primer specific to hygromycin insert (Dicer

1_5R for left flank and Dicer 1_3'F and Dicer 1_3R for the right flank) in the knockout strains were also used to determine the presence of the wild type gene and knockout insertion. The expected bands for left and right flanks were 2,360 bp and 2,721 bp. All of the Dicer 2 and RdRps showed the presence of some wild type gene along with the knockout flanks suggesting the heterokaryotic nature of these transformants while Dicer1 strain showed no amplification of wildtype gene and amplification of insertion flanks revealing the homokaryotic nature.

6.3.4 RNAi gene mutants of *S. sclerotiorum* do not exhibit altered mycelium growth but do have reduced sclerotia formation

After 24 hours, all mutant strains grew well on 90 mm PDA plates with no lag in growth compared to the wild type strain. The average hyphal diameter after 48 h for *dicer1*, and the heterokaryotic *dicer2* and *rdrp* strains were 35.33, 35.67, and 35.22 mm respectively (Fig. 6.4A). Similar to mutant strains, the wild type strain had a mean hyphal diameter of 36.11 mm. There is no significant difference between wild type vs *dicer1* ($p_{adj}=0.80$), wild type vs *dicer2* ($p_{adj}=0.96$), and wild type vs *rdrp* strains ($p_{adj}=0.74$).

However, the average number of sclerotia from the wild type was significantly higher than the mutant strains (Fig. 6.4B). The average number of sclerotia formed by *dicer1*, and the heterokaryotic *dicer2*, and *rdrp* strains was 12, 11, and respectively. However, for the wild type strain the mean number of sclerotia formed was 20. There is significant difference between the wild type and mutant strains with p_{adj} value of 0.00 for each wild type vs *dicer1*, wild type vs *dicer2*, and wild type vs *rdrp*. This suggests these genes might play role in sclerotia formation. However, further studies need to be conducted to confirm this phenomenon.

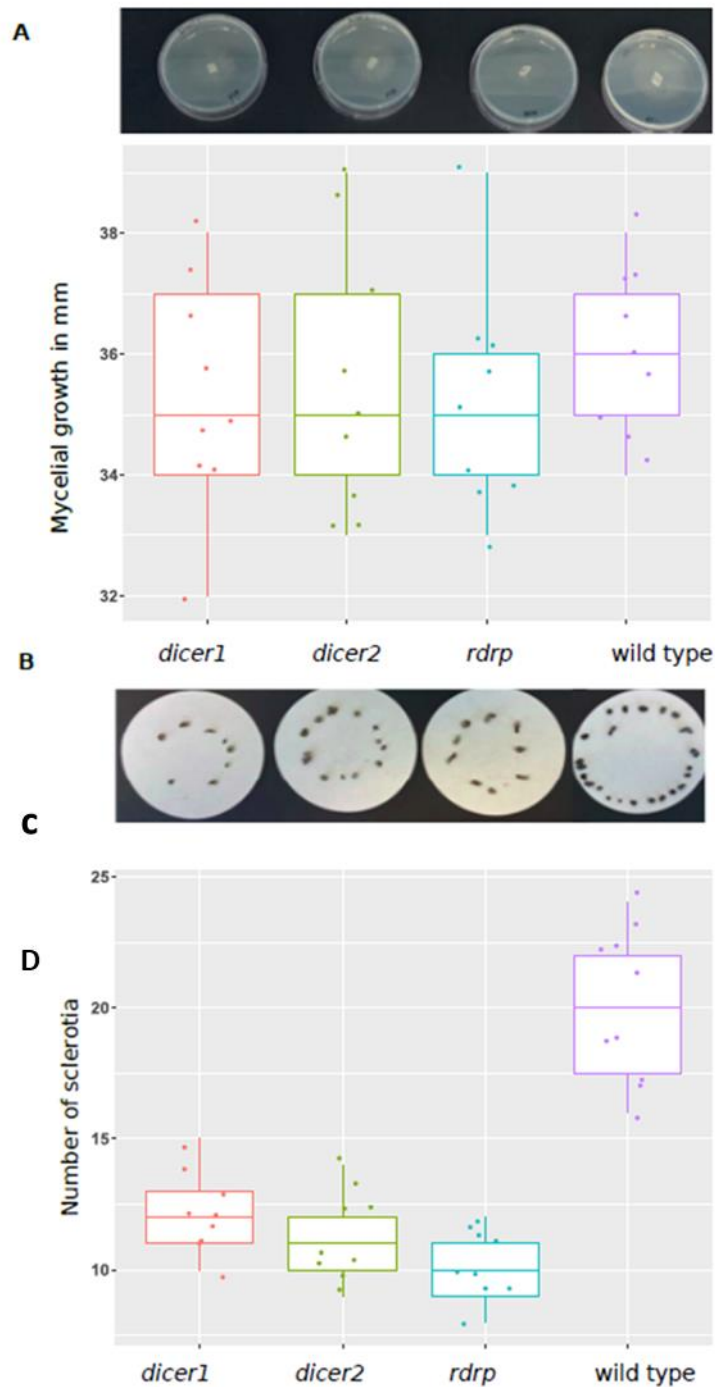


Fig. 6. 4 Comparison of *Sclerotinia sclerotiorum* mycelium morphology in wild type (WT) and gene disruption mutants. A. Photographs of mycelium length taken after 24 hours of incubation. B. Box plot showing mycelium measurement after 24 hours of incubation. Each coloured dot represents data recorded for each of the replicates.

C. Photographs of sclerotia formation after two weeks of incubation. D. Box plot showing sclerotia number. Each coloured dot represents the data recorded for each of the replicates and upper, middle and lower quartiles.

6.3.5 *Dicer1* disruption mutants do not exhibit reduced pathogenicity

The virulence of *dicer1* was assessed by inoculating attached *B. napus* leaves of six-week-old plants of the cultivar AV Garnet with mycelial agar plugs. The lesion length was recorded at 24 and 48 HPI and analysis showed that there was no difference in the average lesion length of two of the independent transformants. However, no lesion was produced by one of the *dicer1* (*dicer1_9*) independent transformants at 24 HPI (Fig. 6.5A) ($p < 0.05$). The average lesion length induced by *dicer1_1*, *dicer1_8*, and *dicer1_9* at 24 HPI were 8.4 mm, 8.3 mm, and 0 mm, respectively, while for wild type the average lesion length was 6.5 mm. There is no significant difference between wild type vs *dicer1_1*, and wild type vs *dicer1_8* with p adj value of 0.46 and 0.47. However, p adj value for wild type vs *dicer1_9* was 0.00008 revealing the significant decrease in symptoms in this strain compared to wild type strain. At 48 HPI, the average lesion lengths induced by *dicer1_1*, *dicer1_8*, and *dicer1_9* were 23.4 mm, 25.75 mm, and 14.65 mm, respectively (Fig. 6.5B). The wild type caused an average lesion length of 25.8 mm. The p adj values for wild type vs *dicer1_1* and *dicer1_8* were 0.8 and 0.99 with no significant difference. However, for wild type vs *dicer1_9* was 0.00066 suggesting a significant difference in the lesion length induced by this strain.

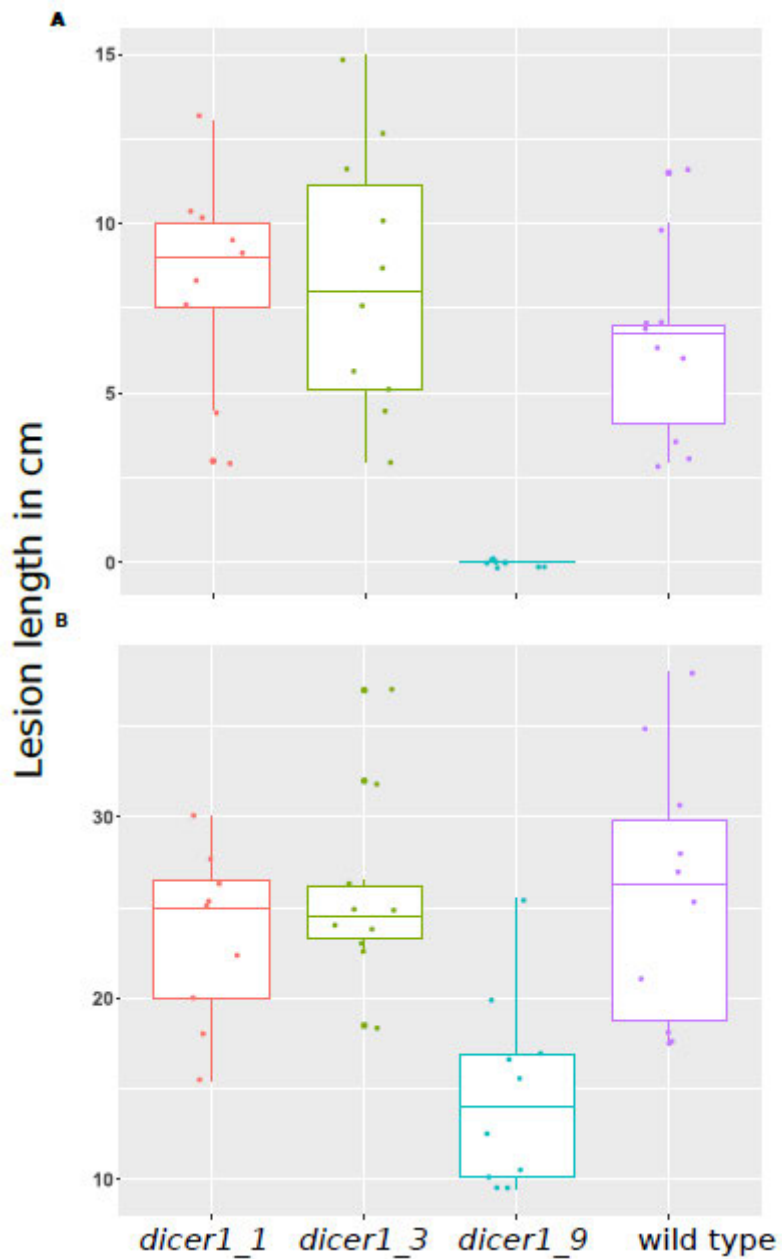


Fig. 6. 5. Effects of Dicer1 gene disruption on *Sclerotinia sclerotiorum* CU8.24 pathogenicity in *Brassica napus* plants. Lesion length at 24 HPI (A) and 48 HPI (B).

6.4 Discussion

Several studies conducted on a number of fungal pathogens have unveiled RNAi silencing mechanisms with varied numbers of RNAi components [30, 43, 44]. Similarly, this study demonstrated the presence of RNAi genes in *S. sclerotiorum* isolate CU8.24. The genome-wide

investigation identified 24 RNAi-related genes in the *S. sclerotiorum* genome with multiple copies of *Dicer-like*, *Argonaute* and *RdRp*-like genes. The gene disruption mutants of *Dicer1* and heterokaryons of *Dicer2* and *RdRp* displayed a non-lethal, wild-type *in vitro* growth phenotype, with no significant difference in mycelium length. Interestingly, the number of sclerotia formed was reduced in the three gene disruption mutants. One of the functions of endogenous fungal genes is to maintain proper fungal morphology and development [45]. Moreover, fungal sRNAs have been demonstrated to be important for sclerotia formation in *S. sclerotiorum* [46]. The decreased number of sclerotia formed by *dicer1*, the heterokaryotic *dicer2* and *rdrp* mutants compared to wild type in our study might be due to the role of these genes in the growth and development of sclerotia. It might be possible that the phenotypes of heterokaryons and *dicer1* knockouts could be due to ectopic insertions as we have not tested the expression of these genes,

If sRNA generation is dependent on the activity of Dicer then *dicer* mutant strains might have an altered phenotype for traits involving sRNA-mediated modulation of gene expression. In this study, one of the strains showed reduced lesion length at both 24 and 48 h post inoculation while no significant difference in pathogenicity was observed in the other two *dicer1* strains. Dicer1 and Dicer2 are mostly functionally redundant in *S. sclerotiorum* and *B. cinerea* [35, 38]. But our observations suggest they may not be functionally redundant for sclerotia formation.

Even in a double knockout of Dicer1 and Dicer2 in the *S. sclerotiorum* 1980 strain, the endogenous sRNA population did not alter, suggesting alternative pathways of sRNA generation in *S. sclerotiorum* [38]. As we found almost 24 RNAi-related genes in CU8.24 strain, we infer that *S. sclerotiorum* might possess Dicer-independent sRNA generation pathways. In addition, RdRp is a required component of the RNA silencing pathway and

conserved across fungi and plants such as *C. elegans*, *N. crassa* and *A. thaliana* [47]. RdRp mostly produces dsRNAs from the primary template or from secondary mRNAs and plays an important role in production of secondary siRNAs such as pha-siRNAs. In Chapter 3 we found evidence to suggest that *S. sclerotiorum* regulates gene expression through the production of pha-siRNAs. In this study, we generated gene disruption strains of one of the predicted *RdRp* genes which has not been investigated before. Our findings are inconclusive for the function of RdRp as pure homokaryotic gene deletion strains were not produced. In *M. circinelloides*, two distinct RdRp proteins were required for initiation and amplification of RNA silencing [4]. A clear picture would only be obtained if all *RdRp* genes were knocked out or knocked down in double or triple mutants.

Protoplast-mediated transformation involves the direct delivery of DNA to individual fungal protoplasts using polyethylene glycol [48]. *S. sclerotiorum* cells are multinucleate and transformation generally results in the transformation of some but not all nuclei within individual protoplasts. The resulting colonies are heterokaryotic with a mixture of transformed and wild type nuclei. Therefore, several generations of hyphal tip transfer on selective media were carried out in an attempt to purify the strains to a homokaryotic state. We could only confirm four independent pure knockout strains for *Dicer1*. However, this study generated several heterokaryotic strains which could be used in future with further purification.

Previous knockout studies in *S. sclerotiorum* have involved screening of 15 or more initial transformants to identify homokaryons [49, 50] Here, we used ten initial transformants for screening of knockout events. While we were able to get four homokaryotic knockout strains out of ten screened strains for *Dicer1*, we were not able to produce any homokaryotic strains lacking *RdRp* and *Dicer2*. It is possible that screening a larger number of strains would result

in identification of homokaryons for these genes if they have a lower transformation efficiency. . In conclusion, this chapter provides insights on knockout experiments including designing of gene deletion constructs and use of the homologous recombination method of gene knockout.

6.4 References

1. Fire, A., et al., *Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans**. *Nature*, 1998. **391**(6669): p. 806-811.
2. Mello, C.C. and D. Conte, *Revealing the world of RNA interference*. *Nature*, 2004. **431**(7006): p. 338-342.
3. Baulcombe, D., *Small RNA—The secret of noble rot*. *Science*, 2013. **342**(6154): p. 45-46.
4. Calo, S., et al., *Two distinct RNA-dependent RNA polymerases are required for initiation and amplification of RNA silencing in the basal fungus *Mucor circinelloides**. *Molecular Microbiology*, 2012. **83**(2): p. 379-394.
5. Cerutti, H. and J.A. Casas-Mollano, *On the origin and functions of RNA-mediated silencing: from protists to man*. *Current Genetics*, 2006. **50**(2): p. 81-99.
6. Sijen, T., et al., *On the role of RNA amplification in dsRNA-triggered gene silencing*. *Cell*, 2001. **107**(4): p. 465-476.
7. Vaistij, F.E., L. Jones, and D.C. Baulcombe, *Spreading of RNA targeting and DNA methylation in RNA silencing requires transcription of the target gene and a putative RNA-dependent RNA polymerase*. *The Plant Cell*, 2002. **14**(4): p. 857-867.

8. Martienssen, R.A., M. Zaratiegui, and D.B. Goto, *RNA interference and heterochromatin in the fission yeast Schizosaccharomyces pombe*. Trends in Genetics, 2005. **21**(8): p. 450-456.
9. Liu, H., et al., *RNA interference in the pathogenic fungus Cryptococcus neoformans*. Genetics, 2002. **160**(2): p. 463-470.
10. Mouyna, I., et al., *Gene silencing with RNA interference in the human pathogenic fungus Aspergillus fumigatus*. FEMS Microbiology Letters, 2004. **237**(2): p. 317-324.
11. Nakayashiki, H. and Q.B. Nguyen, *RNA interference: roles in fungal biology*. Current Opinion in Microbiology, 2008. **11**(6): p. 494-502.
12. Nakayashiki, H., et al., *RNA silencing as a tool for exploring gene function in ascomycete fungi*. Fungal Genetics and Biology, 2005. **42**(4): p. 275-283.
13. Drinnenberg, I.A., et al., *RNAi in budding yeast*. Science, 2009. **326**(5952): p. 544-550.
14. Laurie, J.D., et al., *Genome comparison of barley and maize smut fungi reveals targeted loss of RNA silencing components and species-specific presence of transposable elements*. The Plant Cell, 2012. **24**(5): p. 1733-1745.
15. Feretzaki, M., et al., *Gene network polymorphism illuminates loss and retention of novel RNAi silencing components in the Cryptococcus pathogenic species complex*. PLoS Genetics, 2016. **12**(3): p. e1005868.
16. Bernstein, E., et al., *Role for a bidentate ribonuclease in the initiation step of RNA interference*. Nature, 2001. **409**(6818): p. 363.

17. Doyle, M., L. Jaskiewicz, and W. Filipowicz, *Dicer proteins and their role in gene silencing pathways*, in *The enzymes*. 2012, Elsevier. p. 1-35.
18. Poulsen, C., H. Vaucheret, and P. Brodersen, *Lessons on RNA silencing mechanisms in plants from eukaryotic argonaute structures*. *The Plant Cell*, 2013. **25**(1): p. 22-37.
19. Sun, Q., G.H. Choi, and D.L. Nuss, *A single Argonaute gene is required for induction of RNA silencing antiviral defense and promotes viral RNA recombination*. *Proceedings of the National Academy of Sciences*, 2009. **106**(42): p. 17927-17932.
20. Campo, S., K.B. Gilbert, and J.C. Carrington, *Small RNA-based antiviral defense in the phytopathogenic fungus *Colletotrichum higginsianum**. *PLoS Pathogens*, 2016. **12**(6): p. e1005640.
21. Yu, J., et al., *Differential contribution of RNA interference components in response to distinct *Fusarium graminearum* virus infections*. *Journal of Virology*, 2018. **92**(9): p. e01756-17.
22. Fulci, V. and G. Macino, *Quelling: post-transcriptional gene silencing guided by small RNAs in *Neurospora crassa**. *Current Opinion in Microbiology*, 2007. **10**(2): p. 199-203.
23. Pinzón, N., et al., *Functional lability of RNA-dependent RNA polymerases in animals*. *PLoS Genetics*, 2019. **15**(2): p. e1007915.
24. Yu, L., et al., *Identification of novel phasiRNAs loci on long non-coding RNAs in *Arabidopsis thaliana**. *Genomics*, 2019. **111**(6): p. 1668-1675.
25. Waterhouse, P.M., M.-B. Wang, and T. Lough, *Gene silencing as an adaptive defence against viruses*. *Nature*, 2001. **411**(6839): p. 834-842.

26. Zhang, D.-X., M.J. Spiering, and D.L. Nuss, *Characterizing the roles of Cryphonectria parasitica RNA-dependent RNA polymerase-like genes in antiviral defense, viral recombination and transposon transcript accumulation*. PLoS One, 2014. **9**(9): p. e108653.
27. Weiberg, A., et al., *Fungal small RNAs suppress plant immunity by hijacking host RNA interference pathways*. Science, 2013. 342(6154): p. 118-123.
28. Katiyar-Agarwal, S. and H. Jin, *Role of small RNAs in host-microbe interactions*. Annual Review of Phytopathology, 2010. **48**: p. 225-246.
29. Cui, J., et al., *Characterization of miRNA160/164 and their targets expression of beet (Beta vulgaris) seedlings under the salt tolerance*. Plant Molecular Biology Reporter, 2018. **36**(5-6): p. 790-799.
30. Chen, Y., et al., *Characterization of RNA silencing components in the plant pathogenic fungus Fusarium graminearum*. Scientific Reports, 2015. **5**: p. 12500.
31. Zeng, W., et al., *Dicer-like proteins regulate sexual development via the biogenesis of perithecium-specific microRNAs in a plant pathogenic fungus Fusarium graminearum*. Frontiers in Microbiology, 2018. **9**: p. 818.
32. Dahlmann, T.A. and U. Kück, *Dicer-dependent biogenesis of small RNAs and evidence for microRNA-like RNAs in the penicillin producing fungus Penicillium chrysogenum*. PLoS One, 2015. **10**(5): p. e0125989.
33. Lax, C., et al., *The evolutionary significance of rnai in the fungal kingdom*. International Journal of Molecular Sciences, 2020. **21**(24): p. 9348.

34. Hua, C., J.-H. Zhao, and H.-S. Guo, *Trans-kingdom RNA silencing in plant–fungal pathogen interactions*. *Molecular Plant*, 2018. **11**(2): p. 235-244.
35. Wang, M., et al., *Bidirectional cross-kingdom RNAi and fungal uptake of external RNAs confer plant protection*. *Nature Plants*, 2016. **2**(10): p. 16151.
36. Boland, G.J. and R. Hall, *Index of plant hosts of Sclerotinia sclerotiorum*. *Canadian Journal of Plant Pathology*, 1994. **16**(2): p. 93-108.
37. Derbyshire, M.C. and M. Denton-Giles, *The control of sclerotinia stem rot on oilseed rape (Brassica napus): current practices and future opportunities*. *Plant Pathology*, 2016. **65**(6): p. 859-877.
38. Mochama, P., et al., *Mycoviruses as triggers and targets of RNA silencing in white mold fungus Sclerotinia sclerotiorum*. *Viruses*, 2018. **10**(4): p. 214.
39. Derbyshire, M., et al., *The complete genome sequence of the phytopathogenic fungus Sclerotinia sclerotiorum reveals insights into the genome architecture of broad host range pathogens*. *Genome Biology and Evolution*, 2017. **9**(3): p. 593-618.
40. Derbyshire, M.C., et al., *A whole genome scan of SNP data suggests a lack of abundant hard selective sweeps in the genome of the broad host range plant pathogenic fungus Sclerotinia sclerotiorum*. *PLoS One*, 2019. **14**(3): p. e0214201.
41. Neupane, A., et al., *Roles of argonautes and dicers on Sclerotinia sclerotiorum antiviral RNA silencing*. *Frontiers in Plant Science*, 2019. **10**.
42. Thomas, S., N.D. Maynard, and J. Gill, *DNA library construction using Gibson Assembly®*. *Nature Methods*, 2015. **12**(11): p. i-ii.

43. Fagard, M., et al., *AGO1, QDE-2, and RDE-1 are related proteins required for post-transcriptional gene silencing in plants, quelling in fungi, and RNA interference in animals*. Proceedings of the National Academy of Sciences, 2000. **97**(21): p. 11650-11654.
44. Wang, M., et al., *Bidirectional cross-kingdom RNAi and fungal uptake of external RNAs confer plant protection*. Nature Plants, 2016. **2**(10): p. 1-10.
45. Dang, Y., et al., *RNA interference in fungi: pathways, functions, and applications*. Eukaryotic Cell, 2011. **10**(9): p. 1148-1155.
46. Xia, Z., et al., *Characterization of microRNA-like RNAs associated with sclerotial development in Sclerotinia sclerotiorum*. Fungal Genetics and Biology, 2020. **144**: p. 1034
47. Lipardi, C. and B.M. Paterson, *Identification of an RNA-dependent RNA polymerase in Drosophila establishes a common theme in RNA silencing*. Fly, 2010. **4**(1): p. 30-35.
48. Li, D., et al., *Methods for genetic transformation of filamentous fungi*. Microbial Cell Factories, 2017. **16**(1): p. 1-13.
49. Liang, X., et al., *Oxaloacetate acetylhydrolase gene mutants of Sclerotinia sclerotiorum do not accumulate oxalic acid, but do produce limited lesions on host plants*. Molecular Plant Pathology, 2015. **16**(6): p. 559-571.
50. Sang, H., H.-X. Chang, and M.I. Chilvers, *A Sclerotinia sclerotiorum transcription factor involved in sclerotial development and virulence on pea*. MSphere, 2019. **4**(1): p. e00615-18.

Chapter 7: General discussion and conclusions

7.1 General discussion and conclusions

The last decade has seen a significant increase in our understanding of the role of non-coding RNA, especially small RNAs [1-4]. A single sRNA can regulate hundreds of mRNAs [5], and its repression can influence growth and development [6-8], as well as responses to biotic [9, 10] and abiotic stresses [11, 12]. Several studies have investigated the role of sRNAs in the regulation of endogenous [13] and foreign targets [14, 15] in different phytopathogens to unveil interesting biological processes regulated by these sRNAs. However, the roles of sRNAs in the interaction between *B. napus* and *S. sclerotiorum* are unknown and have been the focus of this PhD research. The results from this thesis further improve on miRNA identification in *B. napus* and their mRNA regulatory network. There are still a large number of sRNAs including miRNAs to be identified from *S. sclerotiorum* and *B. napus*. This thesis aimed to fill this gap by exploring the potential involvement of sRNAs during the *S. sclerotiorum* and *B. napus* interaction. The principle strategy adopted for this project involved using next generation sequencing technology (sRNA and degradome sequencing), bioinformatics and wet lab approaches. This thesis identified different classes of sRNAs from *S. sclerotiorum* and *B. napus*. At first, the endogenous gene regulation by *S. sclerotiorum* sRNAs was investigated (Chapter 3). This chapter identified micro-RNA-like structures from *S. sclerotiorum* and their targets and further improved the understanding on the miRNA-mRNA regulatory network in *S. sclerotiorum*. Secondly, this thesis provided direct evidence that *S. sclerotiorum* sRNAs can act as effectors by silencing plant immune response-related genes in *B. napus* (Chapter 4). A third study on sRNAs focussed on the endogenous regulation of *B. napus* sRNAs during *S. sclerotiorum* infection (Chapter 5). Two previous studies have been conducted on the role of *B. napus* sRNAs during interaction with *S. sclerotiorum* [16, 17]. Chapters 3, 4, and 5

investigated these interactions in detail and identified additional sRNAs of both the host and pathogen compared to previous studies.

Recently, mRNA based studies of pathogens and hosts has enabled a better understanding of the patterns of gene expression and regulation during infection [18, 19]. This technique is highly resourceful in the identification of the key protein-coding genes involved in different biological processes. However, to ascertain a clearer picture of gene regulation, investigation of sRNAs is required, as they regulate several biological processes including development and growth, maintenance of genome integrity, and response to biotic and abiotic stresses [20, 21]. Chapter 3 focused on the identification of sRNAs of *S. sclerotiorum* *in vitro* and during infection of *B. napus* by peeling off the mycelium mat from infected *B. napus* leaves. Furthermore, it identified endogenous targets of *S. sclerotiorum*. Moreover, it was unknown prior to this thesis if expression of *S. sclerotiorum* sRNAs is altered during infection of *B. napus* although there were a few studies which characterized sRNAs from *in vitro* mycelium [22], sclerotia [13] and during infection of *A. thaliana* and common bean [23]. Most of these studies were conducted with a single sRNA sequencing library which limits the detailed understanding of differentially expressed sRNAs and still there are lots of miRNAs and their regulatory networks that are yet to be discovered from this pathogen. In addition, identification of conserved, novel miRNAs, and phasiRNAs, which are involved in the regulation of endogenous mRNA transcripts is important for enhancement of the understanding on the role of these sRNAs in fungal growth and development. Chapter 3 provided insights on identification of different classes of sRNAs and their target genes. This study has made significant progress in the field by identifying different classes of sRNAs and their endogenous targets. Almost 90 % of endogenous targets of *S. sclerotiorum* sRNAs were related to repeat elements. In addition, targets with domains like binding factor, exchange factor and

phosphatase acetyltransferase were identified from the degradome datasets. These domains are known to be important in host and pathogen interactions [24], mRNA stability [25], and as catalytic factors [26]. The growth and development of phytopathogens is regulated by epigenetic processes accompanied with gene silencing and transposon regulation [27, 28]. *S. sclerotiorum* miRNAs are involved in the regulation of transcription factors related to histone modification during sclerotia formation [13]. The results of this thesis showed that a phosphatase acetyltransferase-related gene is potentially regulated by miRNAs. Acetyltransferase enzymes transfer an acetyl group from acetyl CoA to a lysine residue. A fungal specific acetyltransferase has been reported to regulate fungal development and virulence in *Aspergillus fumigatus* [29]. This study suggested that endogenous sRNA regulation in the fungus *S. sclerotiorum* may be important both for interactions with its hosts and maintenance of important cellular functions. Whether phasiRNAs are associated with the RNAi pathway in *S. sclerotiorum* warrants further investigation. PhasiRNAs are characteristic plant sRNAs [30] and only a single study so far has investigated phasiRNAs in *S. sclerotiorum* [31]. This thesis echoes this finding and further expanded the knowledge of phasiRNAs in the fungal pathogen *S. sclerotiorum*. The presence of this class of sRNAs in *S. sclerotiorum* suggests that some sRNAs in this pathogen might trigger the formation of phasiRNAs thereby regulating mRNA transcripts. Although miRNAs are conserved within the same kingdom, whether they are conserved across kingdoms has not been reported so far. This thesis has been able to identify some *S. sclerotiorum* sRNA homologues in other species. Targets of these homologous sRNAs are related to nuclear and exchange factor, which is similar to the functions of targets identified in *S. sclerotiorum*.

Recently, some studies revealed that fungi are capable of secreting sRNAs into plants and the secreted sRNAs can contribute to pathogenesis [15, 32]. This phenomenon is called cross-

kingdom silencing. While some *in silico* studies identified sRNA complementary to mRNAs from *B. napus* [33], *A. thaliana* and common bean, suggesting the probable target genes of these sRNAs [31], there has been no direct experimental evidence. This thesis (Chapter 4) provided experimental evidence of cross-kingdom silencing during *S. sclerotiorum* and *B. napus* interaction. The high throughput method of sRNA target validation termed degradome sequencing was used and predicted hundreds of *B. napus* mRNAs to be cleaved by *S. sclerotiorum* sRNAs, suggesting the presence of cross-kingdom silencing in this pathosystem. Regulation of transcription factors is a highly enriched biological process among the *S. sclerotiorum* sRNA target genes. The role of sRNAs/miRNAs is known to be important in modulating the expression of transcription factors [34]. Enrichment of transcription factors among *B. napus* sRNA targets suggests it might be possible that fungal sRNAs have a broad impact on the regulation of gene expression in *B. napus* which may be important for facilitating infection.

This thesis showed that *S. sclerotiorum* sRNA effectors target plant genes involved in the immune response against fungal pathogens including genes with domains “AP2/ERF”, “MORN motif”, “xylan-binding”, “protein kinase”, “zinc finger domain” motifs. Apart from the involvement of pathogen sRNAs during infection, plants can also employ sRNAs as a defence response against pathogens. These plant sRNAs might regulate auxin-related genes, transcription factors, and some disease resistance genes in the host [35-38]. AP2/ERF transcription factors are only present in plants and are involved in broad biological processes including regulation of disease resistance pathways. ERF genes have been demonstrated to respond to disease-related stimuli, such as ethylene, jasmonic acid, salicylic acid, and infection by pathogens. One of the key features of the plant immune system is the proper subcellular localization of defence factors. Kinases and xylan and have been shown to be involved in plant

defence directly or indirectly [39, 40]. MORN motif containing genes are found in phosphatidylinositol monophosphate kinase (PIP2K), which is a key enzyme in PI- signalling pathways [41]. PI- signalling is one of the components of a plant defense response against pathogens [42].

Cell wall associated kinases are receptor like kinases that possess a cytoplasmic protein kinase domain and their expression is affected by numerous stimuli including pathogen attack [43]. Xylan maintains the integrity of the plant cell wall and enhances cell wall recalcitrance to enzymatic digestion thereby enhancing plant defence against fungal pathogens [44]. Zinc finger proteins are abundant groups of proteins with a wide range of molecular and cellular functions in eukaryotes [45]. These proteins have been shown to have a role in different diseases in humans like tumorigenesis, cancer progression, skin diseases and diabetes [45]. In plants, zinc finger protein transcription factors have key roles in different biological processes like phytohormone response, growth and development, transcriptional regulation, binding of RNA and protein-protein interaction, and stress-tolerance [46].

To date, only two studies have been conducted to investigate endogenous gene regulation mediated by *B. napus* sRNAs and thus, far fewer miRNA families have been identified in *B. napus* compared with model plants *A. thaliana* and *Medicago truncatula*. This could suggest that a larger number of miRNAs are yet to be discovered in *B. napus*. Interestingly, in chapter 5, 73 *B. napus* miRNA families containing 528 mature miRNA sequences were identified.

In addition to these conserved miRNA sequences an additional 135 novel *B. napus* miRNAs were identified. The composition of infected sRNAs were quite different from the mock sRNAs expressed suggesting the sRNA biogenesis pathway in *B. napus* changes upon *S. sclerotiorum* infection. A total of 50 predicted target genes were unique to infected samples. Altogether 172

InterPro domains were related to these genes which have functions such as transcriptional regulation, disease resistance and post transcriptional gene silencing. Moreover, 64 target genes were predicted from upregulated sRNAs during infection which were related to “ABC transcription factor”, “heat shock response”, “disease resistance protein-like”, “zinc finger domain”, and “leucine zipper domain”. The expression analysis of conserved miRNAs unveiled that miR5139 is likely expressed in infected *B. napus* which regulates a zinc finger domain-containing gene. Zinc finger domains are reported to be present in plant resistance related proteins that are involved in the effector-triggered immune response [46]. One of the ethylene response factor genes was also validated to be cleaved by a novel miRNA. Ethylene responsive binding factors have been implicated in regulating biological processes including defence signalling, secondary metabolite production and redox regulation [47].

Chapters 2 and 6 shifted focus to protoplast-mediated transformation of *S. sclerotiorum*. The focus of these chapters was genetic transformation of *S. sclerotiorum* to introduce GFP (Chapter 2) and knock out genes involved in the RNA interference machinery (Chapter 6). The green fluorescent protein facilitates the study of host and pathogen interaction when a fungus is stably transformed with this protein. This thesis developed a *S. sclerotiorum* GFP strain of an aggressive Australian isolate (CU8.24). This GFP-transformed strain can be used to understand the interactions between *S. sclerotiorum* and *B. napus* (and other hosts) under different conditions. Two major phenomena were observed during GFP transformation of *S. sclerotiorum*. The intensity of fluorescence improved with subsequent hyphal tip regeneration and within the stable transformants intensity of GFP fluorescence varied in different hyphae. Such phenomena might be due to the multinucleate nature of *S. sclerotiorum*. Nevertheless, the stably transformed GFP-strains showed no significant differences in growth *in vitro* or in pathogenicity when infecting *B. napus* compared to the untransformed wild type strains. The

isolates can thus be used for future microscopy studies investigating *S. sclerotiorum*-host interactions.

A number of fungal pathogens have been shown to have different numbers of RNAi silencing components [48]. In *S. sclerotiorum* altogether 24 genes were identified with homology to *Dicer*, *Argonaute* and *RdRP* gene sequences from *N. crassa*. The closest homologues of *Dicer1*, *Dicer2*, *Argonaute* and *RdRP* were selected for targeted gene knockout. Protoplast-mediated transformation includes the direct transfer of DNA to individual fungal protoplasts using polyethylene glycol [49]. Due to multinucleate nature of *S. sclerotiorum* cells, the transformants might have a mixture of transformed and untransformed nuclei within an individual protoplast. Therefore, multiple hyphal tip transfer is conducted over several generations to purify the strains to a homokaryotic state. This thesis could only produce homokaryotic strains for *Dicer1* gene knockouts. Different numbers of initial transformants were screened to select transformants for purification by hyphal tip transfer on selective media. If PEG-mediated transformation is time and resource consuming for such genes it could be viable to adopt a siRNA hairpin-mediated silencing technique to silence these genes in *S. sclerotiorum*.

7.2 Final comments and future direction

This thesis has taken advantage of new generation sequencing techniques to unravel the role of sRNAs in regulating their target transcripts. Combined sRNA and degradome sequencing in single experiments were complicated by handling of large datasets, bioinformatics challenges, and understanding of the biases and artefacts which are associated with the generated sequencing datasets. In the underlying aim of this project to better understand the functionality of sRNAs, the molecular technique such as 5'RACE and expression analysis through

quantitative qPCR was adopted. Important conclusions were made on the role of sRNAs in regulating endogenous targets and trans-species target during *S. sclerotiorum* infection of *B. napus* plants. To conclude, I will emphasize that repeat elements are hot spots for sRNA generation in *S. sclerotiorum*. Among hundreds of endogenous targets, phosphate acetyltransferase, RNA-binding factor and exchange factor were found in the degradome dataset. I again accentuate here that *S. sclerotiorum*, along with siRNAs and miRNAs, can also produce secondary siRNAs called phasiRNAs suggesting that the pathogen might produce sRNAs that function similarly to miRNAs in higher eukaryotes. Furthermore, *S. sclerotiorum* sRNAs effectors can silence *B. napus* transcripts. These transcripts are mostly enriched for “transcription factors”. The degradome dataset further identified that *S. sclerotiorum* sRNAs target *B. napus* genes involved in the plant immune response including disease resistance genes and genes with the following domains: AP2/ERF, zinc finger domain, protein kinase and MORN motif. I further emphasise the fact that 5' RACE, *in silico* target prediction like psRNA and degradome sequencing all provide insights on the sRNA- targets. *In silico* target prediction might result in false positive transcripts which might not be produced due to true cleavage events. Therefore, integrated analysis and orthogonal experiments are required. Finally, due to the multinucleate nature of *S. sclerotiorum*, protoplast mediated transformation might not result in pure homokaryotic strains. For this reason, siRNA mediated silencing method may be a more efficient alternative.

Canola growers and farmers are facing serious production constraints due to Sclerotinia stem rot disease. Crop protection strategies mostly rely on the widespread use of chemical pesticides. RNAi approaches can be utilised as a novel and sustainable method for disease management. However, it demands extensive research in sRNAs and their regulatory networks since RNAi based technology relies on the identification of gene functions and targets related to

pathogenesis. Although there is growing research on the molecular processes of RNAi in pathogens and plants, still a lot remains to be explored. Particularly, from the pathogen side, whether it has a similar sRNA biogenesis pathway like plants or different. Whether the binding site from fungal pathogen is similar like in plants. Does the broad host pathogen like *S. sclerotiorum* sRNAs evolve similar way like other pathogens? At first, the function of all RNAi related genes needs to be illustrated better with the reverse genetics or silencing techniques. Do fungal sRNAs trigger the formation of phasiRNAs which silence several targets at the same time rather than silencing mediated by single a sRNA? More attention should be focused on the regulation of phasiRNAs in fungal pathogen and their regulatory networks during host interactions.

In conclusion, the results of this thesis have advanced the understanding of the *S. sclerotiorum* -*B. napus* interaction at the sRNA level. The prediction of *S. sclerotiorum* targets in *B. napus* might assist in the identification of plant genes involved in disease resistance. Moreover, the findings could facilitate the use of RNAi to silence host susceptibility genes and pathogen virulence genes.

7.3 References

1. Betti, F., et al., *Exogenous miRNAs induce post-transcriptional gene silencing in plants*. Nature Plants, 2021. 7(10): p. 1379-1388.
2. Jin, Y., et al., *Evaluating the microRNA targeting sites by luciferase reporter gene assay*, in *MicroRNA Protocols*. 2013, Springer. p. 117-127.
3. Mineno, J., et al., *The expression profile of microRNAs in mouse embryos*. Nucleic Acids Research, 2006. 34(6): p. 1765-1771.

4. Huang, J., M. Yang, and X. Zhang, *The function of small RNAs in plant biotic stress response*. Journal of Integrative Plant Biology, 2016. **58**(4): p. 312-327.
5. Shkumatava, A., et al., *Coherent but overlapping expression of microRNAs and their targets during vertebrate development*. Genes & Development, 2009. **23**(4): p. 466-481.
6. Mraheil, M.A., et al., *The intracellular sRNA transcriptome of Listeria monocytogenes during growth in macrophages*. Nucleic Acids Research, 2011. **39**(10): p. 4235-4248.
7. Raman, V., et al., *Small RNA functions are required for growth and development of Magnaporthe oryzae*. Molecular Plant-Microbe Interactions, 2017. **30**(7): p. 517-530.
8. Wang, L., et al., *Integrated microRNA and mRNA analysis in the pathogenic filamentous fungus Trichophyton rubrum*. BMC Genomics, 2018. **19**(1): p. 933.
9. Cao, J.-Y., et al., *Tight regulation of the interaction between Brassica napus and Sclerotinia sclerotiorum at the microRNA level*. Plant Molecular Biology, 2016. **92**(1): p. 39-55.
10. Fei, S., et al., *Small RNA profiling of Cavendish banana roots inoculated with Fusarium oxysporum f. sp. cubense race 1 and tropical race 4*. Phytopathology Research, 2019. **1**(1): p. 22.
11. Huang, S.Q., et al., *A set of miRNAs from Brassica napus in response to sulphate deficiency and cadmium stress*. Plant Biotechnology Journal, 2010. **8**(8): p. 887-899.
12. Cui, J., et al., *Characterization of miRNA160/164 and their targets expression of beet (Beta vulgaris) seedlings under the salt tolerance*. Plant Molecular Biology Reporter, 2018. **36**(5-6): p. 790-799.
13. Xia, Z., et al., *Characterization of microRNA-like RNAs associated with sclerotial development in Sclerotinia sclerotiorum*. Fungal Genetics and Biology, 2020. **144**: p. 103471.

14. Dieuez, M.J., et al., *Physical and genetic mapping of amplified fragment length polymorphisms and the leaf rust resistance Lr3 gene on chromosome 6BL of wheat*. 2006. 112(2):p.251-257
15. Wang, M., et al., *Botrytis small RNA Bc-siR37 suppresses plant defense genes by cross-kingdom RNAi*. RNA biology, 2017. **14**(4): p. 421-428.
16. Cao, J.Y., et al., *Tight regulation of the interaction between Brassica napus and Sclerotinia sclerotiorum at the microRNA level*. Plant Molecular Biology, 2016. **92**(1-2): p. 39-55.
17. Jian, H., et al., *Integrated mRNA, sRNA, and degradome sequencing reveal oilseed rape complex responses to Sclerotinia sclerotiorum (Lib.) infection*. Scientific Reports, 2018. **8**(1): p. 1-17.
18. Westermann, A.J., S.A. Gorski, and J. Vogel, *Dual RNA-seq of pathogen and host*. Nature Reviews Microbiology, 2012. **10**(9): p. 618-630.
19. Dewage, C.S.K., et al., *Host-pathogen interactions in relation to management of light leaf spot disease (caused by Pyrenopeziza brassicae) on Brassica species*. Crop and Pasture Science, 2018. **69**(1): p. 9-19.
20. Kettles, G.J., et al., *sRNA profiling combined with gene function analysis reveals a lack of evidence for cross-kingdom RNAi in the wheat–Zymoseptoria tritici pathosystem*. Frontiers in Plant Science, 2019. **10**: p. 892.
21. Du, J., et al., *Identification of microRNAs regulated by tobacco curly shoot virus co-infection with its betasatellite in Nicotiana benthamiana*. Virology Journal, 2019. **16**(1).
22. Zhou, J., et al., *Identification of microRNA-like RNAs in a plant pathogenic fungus Sclerotinia sclerotiorum by high-throughput sequencing*. Molecular Genetics and Genomics, 2012. **287**(4): p. 275-282.

23. Derbyshire, M., et al., *Small RNAs from the plant pathogenic fungus Sclerotinia sclerotiorum highlight host candidate genes associated with quantitative disease resistance*. *Molecular Plant Pathology*, 2019. **20**(9): p. 1279-1297.
24. Fawke, S., et al., *Glycerol-3-phosphate acyltransferase 6 controls filamentous pathogen interactions and cell wall properties of the tomato and Nicotiana benthamiana leaf epidermis*. *New Phytologist*, 2019. **223**(3): p. 1547-1559.
25. Brockmann, C., et al., *Structural basis for polyadenosine-RNA binding by Nab2 Zn fingers and its function in mRNA nuclear export*. *Structure*, 2012. **20**(6): p. 1007-1018.
26. Bravo-Plaza, I., et al., *Identification of the guanine nucleotide exchange factor for SARI in the filamentous fungal model Aspergillus nidulans*. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, 2019. **1866**(12): p. 118551.
27. Martienssen, R.A. and V. Colot, *DNA methylation and epigenetic inheritance in plants and filamentous fungi*. *Science*, 2001. **293**(5532): p. 1070-1074.
28. Sarikaya-Bayram, Ö., et al., *Membrane-bound methyltransferase complex VapA-VipC-VapB guides epigenetic control of fungal development*. *Developmental Cell*, 2014. **29**(4): p. 406-420.
29. Zhang, Y., et al., *The fungal-specific histone acetyltransferase Rtt109 regulates development, DNA damage response, and virulence in Aspergillus fumigatus*. *Molecular Microbiology*, 2021. **115**(6): p. 1191-1206.
30. Felippes, F.F. and D. Weigel, *Triggering the formation of tasiRNAs in Arabidopsis thaliana: the role of microRNA miR173*. *EMBO Reports*, 2009. **10**(3): p. 264-270.
31. Lee Marzano, S.-Y., A. Neupane, and L. Domier, *Transcriptional and small RNA responses of the white mold fungus Sclerotinia sclerotiorum to infection by a virulence-attenuating hypovirus*. *Viruses*, 2018. **10**(12): p. 713.

32. da Silva, R.P., et al., *Extracellular vesicle-mediated export of fungal RNA*. Scientific Reports, 2015. **5**(1): p. 1-12.
33. Mochama, P., et al., *Mycoviruses as triggers and targets of RNA silencing in white mold fungus Sclerotinia sclerotiorum*. Viruses, 2018. **10**(4): p. 214.
34. Chen, S., et al., *Genome-Wide Analysis of Coding and Non-coding RNA Reveals a Conserved miR164–NAC–mRNA Regulatory Pathway for Disease Defense in Populus*. Frontiers in Genetics, 2021. **12**:p.1-12
35. Achard, P., et al., *The plant stress hormone ethylene controls floral transition via DELLA-dependent regulation of floral meristem-identity genes*. Proceedings of the National Academy of Sciences, 2007. **104**(15): p. 6484-6489.
36. Chen, Z., et al., *Pseudomonas syringae type III effector AvrRpt2 alters Arabidopsis thaliana auxin physiology*. Proceedings of the National Academy of Sciences, 2007. **104**(50): p. 20131-20136.
37. Navarro, L., et al., *Suppression of the microRNA pathway by bacterial effector proteins*. Science, 2008. **321**(5891): p. 964-967.
38. Samad, A.F., et al., *MicroRNA and transcription factor: key players in plant regulatory network*. Frontiers in Plant Science, 2017. **8**: p. 565.
39. Cecchini, N.M., et al., *Kinases and protein motifs required for AZI1 plastid localization and trafficking during plant defense induction*. The Plant Journal, 2021. **105**(6): p. 1615-1629.
40. Malinovsky, F.G., J.U. Fangel, and W.G. Willats, *The role of the cell wall in plant immunity*. Frontiers in Plant Science, 2014. **5**: p. 178.
41. Ma, H., et al., *MORN motifs in plant PIPKs are involved in the regulation of subcellular localization and phospholipid binding*. Cell Research, 2006. **16**(5): p. 466-478.

42. Hung, C.-Y., et al., *Phosphoinositide-signaling is one component of a robust plant defense response*. *Frontiers in Plant Science*, 2014. **5**: p. 267.
43. Kohorn, B.D. and S.L. Kohorn, *The cell wall-associated kinases, WAKs, as pectin receptors*. *Frontiers in Plant Science*, 2012. **3**: p. 88.
44. Wierzbicki, M.P., et al., *Xylan in the middle: understanding xylan biosynthesis and its metabolic dependencies toward improving wood fiber for industrial processing*. *Frontiers in Plant Science*, 2019. **10**: p. 176.
45. Cassandri, M., et al., *Zinc-finger proteins in health and disease*. *Cell Death Discovery*, 2017. **3**(1): p. 1-12.
46. Cabot, C., et al., *A role for zinc in plant defense against pathogens and herbivores*. *Frontiers in Plant Science*, 2019. **10**: p. 1171.
47. van Loon, L.C., B.P. Geraats, and H.J. Linthorst, *Ethylene as a modulator of disease resistance in plants*. *Trends in Plant Science*, 2006. **11**(4): p. 184-191.
48. Chen, Y., et al., *Characterization of RNA silencing components in the plant pathogenic fungus *Fusarium graminearum**. *Scientific Reports*, 2015. **5**: p. 12500.
49. Li, D., et al., *Methods for genetic transformation of filamentous fungi*. *Microbial Cell Factories*, 2017. **16**(1): p. 1-13.

APPENDIX

This section contains Supplementary Figures, Tables and Files for Chapters 3, 4, 5 and 6 of this thesis. This supplementary information can be accessed from the following URL:

<https://figshare.com/s/8a3250270b294579f50e>

To login to the site used the following details

E-mail address: roshan.regmi@postgrad.curtin.edu.au

Password: RRThesis2022@

List of Supplementary Data for the Thesis

Supplementary File 3.1 Summary of 12 endogenous targets predicted from 165 highly expressed *Sclerotinia sclerotiorum* miRNAs.

Supplementary Fig. 4.1 Gel image showing the amplification of correct RACE amplified products of four genes.

Supplementary File 4.1 680 miRNAs identified from miRDeep2.

Supplementary File 4.2 80 miRNAs identified from miRCat.

Supplementary Table 4.1 Primer used for 5' Rapid Amplification of cDNA ends and quantitative real time polymerase chain reaction.

Supplementary Table 4.2 87 common *Sclerotinia sclerotiorum* miRNAs loci identified by Shortstack and miRDeep2.

Supplementary Table 4.3 psRNA targets of *Sclerotinia sclerotiorum* small RNAs in *Brassica napus*.

Supplementary Table 4.4 Degradome validated targets of *Sclerotinia sclerotiorum* small RNAs in *Brassica napus*.

Supplementary Fig. 5.1 Size distribution of unique sRNAs.

Supplementary File 5.1 PHAS loci identified from PhaseTank

Supplementary Table 5.1 Primer used for qPCR and RACE experiments

Supplementary Table 5.2 All small RNA biogenesis loci identified using the Shortstack program.

Supplementary Table 5.3 Cleaved products of the 73 miRNA families. All targets identified using degradome sequencing in infected and mock samples are included.

Supplementary Table 5.4 Targets of conserved microRNAs based on degradome sequencing data from the infected sample. Table includes genes targeted by multiple miRNAs and multiple genes targeted by the same miRNA. Targets only found in the infected sample are included.

Supplementary Table 5.5 Genes targeted by conserved miRNAs only in the infected sample.

Supplementary Table 5.6 Conserved micro RNA sequences with evidence of differential expression during infection.

Supplementary Table 5.7 Known and novel miRNAs identified in this study.

Supplementary Table 5.8 GO term enrichment analysis of 5,918 transcripts possibly regulated by pha-siRNAs based on psRNA target analysis.

Supplementart Table 5.9 GO term enrichment analysis of psRNA target-predicted targets of 1,601 targets of miR1885-triggered ta-siRNAs.

Supplementary Fig. 6.1 Amplification of RNAi related genome flanks for four genes *Dicer1*, *Dicer2*, *Argonaute*, and *RdRp*.

Supplementary Fig 6.2 Gel images showing the confirmation of correct assembly involving the evidence of undigested full plasmid construct from Gibson assembly and restriction digest tests.

Supplementary Fig 6.3 Gel images showing the confirmation of homokaryotic strains of *dicer1* and heterokaryotic strains from *dicer2* and *rdrp* strains.

Supplementary Table 6.1 Primer used to generate the knockout constructs and validation of the correct insertions of the plasmids as well as the validation of homo and heterokaryons.

Supplementary Table 6.2 24 RNAi related genes in the genome sequences of CU8.24 *Sclerotinia sclerotiorum* through translated amino acid blast search from NCBI