Project # CGDP SCDC 2011-8
Date Rec'd December 1st 2015
Researcher Gopalan Selvaraj
Reviewed by M. Devine
Date Reviewed Jan 4/16

SaskCanola

1. Project title

Building durable clubroot resistance in Canola: Identification of multiple clubroot resistance genes from *Brassica napus* and *B. rapa* for marker-assisted gene stacking in canola breeding

2. Name of the Principal Investigator and contact information.

Gopalan Selvaraj, National Research Council of Canada, 110 Gymnasium Place, Saskatoon SK S7N 0W9
Tel 1-306-975-5577; 1-306-251-2640; gopalan.selvaraj@nrc.ca (or gopalan.selvaraj@canada.ca AFTER email migration)

3. Name of the collaborators and contact information.

Dr Manoj Kulkarni, Research Associate, National Research Council of Canada, 110 Gymnasium Place, Saskatoon SK S7N 0W9; manoj.kulkarni@nrc.ca (employed on this project)
Dr Habibur Rahman, University of Alberta, Agriculture, Food and Nutritional Science Edmonton, ABT6G 2P5 Edmonton habibur.rahman@ualberta.ca (Collaboration as noted in the text)

4. Abstract/ Summary: *This must include project objectives, results, and conclusions for use in publications and in the Ministry database. Maximum of 300 words in lay language.*

Clubroot is a serious soil-borne disease of canola caused by *Plasmodiophora brassicae*. The spores survive in soil for many years and thus the disease is always a threat. Our objective was to find the genes that can be associated with disease resistance so that resistance genes can be used judiciously to build durable resistance. Of the many different pathotypes of *P. brassicae,* we focused on P3 that is the most significant pathotype in western Canada. Our study is focused on *Brassica napus* and not *B. rapa*. We have mapped a major disease resistance locus in *B. napus* and identified a tightly linked DNA marker for resistance. Given that other minor genes contribute to disease resistance in plants and that these collectively provide durable resistance, we undertook a comprehensive study of genes whose activities can be associated with clubroot disease progression in canola (*B. napus*). This study provided first of its kind analysis covering all three important disease development stages (primary, secondary and maturity phase). Gene expression control is exquisite in higher organisms like plants. Small RNAs that by themselves are too small to encode proteins control protein-coding gene functions. We completed small RNA analysis for the primary, secondary and maturity phases of disease progression in susceptible and resistant canola and identified differentially expressed small RNA were identified. This provides yet another potential arsenal for manipulating disease control in canola.

5. Introduction: *Brief project background and rationale.*

Clubroot is a serious soil-borne disease of Brassica crops caused by *Plasmodiophora brassicae*. The pathogen's hardy resting spores that survive for ~15 years make the disease spread a serious threat. Resistant phenotypes have been identified and some commerical varieties have been released but their genetic make-up is unknown. There are only few genetic markers that can be used for introducing resistance trait into cultvars/advanced germplasm of choice. Nor is there a way to do informed rotations or gene stacking.

Clubroot on canola in Canada was confirmed in 1997 in Quebec, in Alberta in 2003, and was found in west central Saskatchewan in 2008. It has been declared a pest by the Saskatchewan Ministry of Agriculture in 2009. In Alberta, since its discovery near Edmonton, clubroot has spread to 18 municipalities (Strelkov et al., 2007, J Phytopath.155:

706; www1.agric.gov.ab.ca/$department/deptdocs.nsf/all/agdex11519). Clubroot is a concern to Saskatchewan canola production; no chemical treatments are available, and good farming practices are required (Dokken–Bouchard and Vakulabharanam, 2010). In addition, genetic solutions are necessary to avert disease outbreaks that can cause economic disasters for producers and allied industries. Varieties with some degree of resistance are now available but these are not adequate. The resistance afforded by R genes must be durable so that pathogens cannot outpace the host and overcome the resistance. Thus, duplication of the same R genes in new varieties is unproductive, and multiple genes that can be clearly identified and stacked or rotated are required for long-term effectiveness of clubroot management toward protecting canola yield.

Our objective is to map and characterize CR genes so that the gene identities can be elucidated. This information will be useful to canola breeders to introduce CR genes into modern cultivars and these cultivars may then be used in rotations as a proactive measure to mitigate breakdown of resistance. Gene identification will also help in stacking appropriate CR genes in the event of a need. Additonally, building durable resistance necessarily requires finding a collection of genes that contributes to resistance even at modest level when considerered individually. The literature on disease resistance states that assembling a suite of minor genes will contribute to durable resistance than the use of one or two major genes.

Breeding CR gene requires markers. Disease resistance is generally difficult to track by its phenotype during breeding. With a disease such as clubroot, it is even more difficult. With the CR genes from various sources, genic markers can be developed for use in marker-assisted breeding.

6. Methodology: *Include approaches, experimental design, methodology, materials, sites, etc.*

Plant materials: Progenies from a genetic cross of 'Mendel' with two susceptible spring *B. napus* canola lines A04-71NA and A04-75NA served as the source. The University of Alberta (H Rahman lab) characterized one population and we characterized the other population. Thus, the two populations complement each other since both have one common resistant parent. Nine-four double haploid (DH) lines from each cross comprised a given population. All our in-depth studies were conducted in this population.

In the first year, screening of DH mapping population included two parents (clubroot resistant and susceptible) and 12 DH lines each from resistant and susceptible category (Fig.1). These root samples were used for RNAseq experiment after processing. This experiment was repeated again to phenotype the disease reaction using root dip protocol as described by Rahman et al (2011) to screen a set of 94 DH mapping population. The aim of this second batch of screening was to reconfirm the disease reaction of the doubled haploid population to pathotype 3 before using segregating population for genetic linkage mapping. This screening was done in two batches at the University of Alberta by Dr. H. Rahman group (NRC does not have a containment facility and therefore we did not perform disease assays in Saskatoon). This clearly established both type of individuals (resistant and susceptible types) which were scored for disease reaction. Leaf samples collected from this trial were further used for further genetic mapping purpose.

Genetic linkage mapping: These 94 lines were used for genetic linkage mapping. Initially Dr. Rahman's group provided SSR markers from Bulked Segregant Analysis (BSA) studies which were genotyped and additional markers

were added to prepare the first genetic map for clubroot resistance in 2013. Later on this genetic map was further populated with initial GBS analysis using *B. rapa* as reference genome. This genetic map was further strengthened by re-doing GBS analysis with *B. napus* genome (available in August 2014). In the end, we had successfully developed genetic and physical map for Clubroot resistance with tightly linked marker for clubroot resistance.

RNASeq transcriptome analysis: We have sequenced more than 2.5 billion transcriptome reads to cover the three major stages of disease development with pathotype 3. RNASeq analysis was conducted to identify differentially expressed genes. Tuxedo suite pipeline was used to identify differentially expressed genes using initially *B. rapa* and later *B. napus*. This analysis provided differentially expressed genes at primary and secondary infection stages and at disease maturity in resistant and susceptibility responses. The data were further annotated and clustered using MapMan and PageMan software for predicting functions. This analysis also provided a comprehensive set of multiple genes which are consistently expressed at higher level in the resistant parent. This can be useful in building durable clubroot resistance.

Small RNA sequencing and analysis:
48 small RNA libraries were prepared using Illumina TrueSeq Small RNA libraries and sequenced using Illumina HiSeq2500 v4 chemistry (University of Toronto). We annotated more than 400 million small RNA reads using CLC Bio software workbench 7 and mapped them using *B. napus* microRNA database miRBase Release 21. Small RNAs were annotated, mapped and merged using CLC Bio pipeline and mapped to *B. napus* v.1 genome to get genome wide abundance of small RNA targeting the genes. Differentially expressed and unique to resistant and susceptible reaction were identified specifically at the secondary infection stage that target the genomic regions on chromosome N3 (this is the chromosome where Clubroot resistance gene for P3 pathotype derived from Mendel is mapped).

Fig. 1. Clubroot resistance screening using single spore isolate of Pathotype 3 for the parents and 12 individual double haploids representing resistant and susceptible population.

Seeding plans

First Seeding & inoculation = Rep I and 3 time points (seeding in all trays in same order)

	Rep I, 10 DAI							Rep I, 21-23 DAI							Rep I, 41-43 DAI						
Plant 1	R1	S3	R4	S6	C-R	S10	R11	R1	S3	R4	S6	C-R	S10	R11	R1	S3	R4	S6	C-R	S10	R11
Plant 1	S1	P1	S5	R6	S8	R9	S11	S1	P1	S5	R6	S8	R9	S11	S1	P1	S5	R6	S8	R9	S11
Plant 1	R2	S4	R5	S7	R8	P2	R12	R2	S4	R5	S7	R8	P2	R12	R2	S4	R5	S7	R8	P2	R12
Plant 1	S2	R3	C-S	R7	S9	R10	S12	S2	R3	C-S	R7	S9	R10	S12	S2	R3	C-S	R7	S9	R10	S12
Plant 2	R1	S3	R4	S6	C-R	S10	R11	R1	S3	R4	S6	C-R	S10	R11	R1	S3	R4	S6	C-R	S10	R11
Plant 2	S1	P1	S5	R6	S8	R9	S11	S1	P1	S5	R6	S8	R9	S11	S1	P1	S5	R6	S8	R9	S11
Plant 2	R2	S4	R5	S7	R8	P2	R12	R2	S4	R5	S7	R8	P2	R12	R2	S4	R5	S7	R8	P2	R12
Plant 2	S2	R3	C-S	R7	S9	R10	S12	S2	R3	C-S	R7	S9	R10	S12	S2	R3	C-S	R7	S9	R10	S12
Plant 3	R1	S3	R4	S6	C-R	S10	R11	R1	S3	R4	S6	C-R	S10	R11	R1	S3	R4	S6	C-R	S10	R11
Plant 3	S1	P1	S5	R6	S8	R9	S11	S1	P1	S5	R6	S8	R9	S11	S1	P1	S5	R6	S8	R9	S11
Plant 3	R2	S4	R5	S7	R8	P2	R12	R2	S4	R5	S7	R8	P2	R12	R2	S4	R5	S7	R8	P2	R12
Plant 3	S2	R3	C-S	R7	S9	R10	S12	S2	R3	C-S	R7	S9	R10	S12	S2	R3	C-S	R7	S9	R10	S12

Fisrt seeding - untreated patents (control) = Rep I

	Rep I, 10 day		Rep I, 21-23 d		Rep I, 41-43	
Plant 1	P1	P2	P1	P2	P1	P2
Plant 2	P1	P2	P1	P2	P1	P2
Plant 3	P1	P2	P1	P2	P1	P2
Plant 4	P1	P2	P1	P2	P1	P2

Second Seeding & inoculation = Rep II and 3 time points (seeding in all trays in same order)

	Rep II, 10 DAI							Rep II, 21-23 DAI							Rep II, 41-43 DAI						
Plant 1	R10	S6	R3	S9	R4	S2	R8	R10	S6	R3	S9	R4	S2	R8	R10	S6	R3	S9	R4	S2	R8
Plant 1	C-S	R7	P2	R11	S1	R12	S10	C-S	R7	P2	R11	S1	R12	S10	C-S	R7	P2	R11	S1	R12	S10
Plant 1	R6	S5	R9	S4	P1	S8	C-R	R6	S5	R9	S4	P1	S8	C-R	R6	S5	R9	S4	P1	S8	C-R
Plant 1	S3	R5	S12	R2	S11	R1	S7	S3	R5	S12	R2	S11	R1	S7	S3	R5	S12	R2	S11	R1	S7
Plant 2	R10	S6	R3	S9	R4	S2	R8	R10	S6	R3	S9	R4	S2	R8	R10	S6	R3	S9	R4	S2	R8
Plant 2	C-S	R7	P2	R11	S1	R12	S10	C-S	R7	P2	R11	S1	R12	S10	C-S	R7	P2	R11	S1	R12	S10
Plant 2	R6	S5	R9	S4	P1	S8	C-R	R6	S5	R9	S4	P1	S8	C-R	R6	S5	R9	S4	P1	S8	C-R
Plant 2	S3	R5	S12	R2	S11	R1	S7	S3	R5	S12	R2	S11	R1	S7	S3	R5	S12	R2	S11	R1	S7
Plant 3	R10	S6	R3	S9	R4	S2	R8	R10	S6	R3	S9	R4	S2	R8	R10	S6	R3	S9	R4	S2	R8
Plant 3	C-S	R7	P2	R11	S1	R12	S10	C-S	R7	P2	R11	S1	R12	S10	C-S	R7	P2	R11	S1	R12	S10
Plant 3	R6	S5	R9	S4	P1	S8	C-R	R6	S5	R9	S4	P1	S8	C-R	R6	S5	R9	S4	P1	S8	C-R
Plant 3	S3	R5	S12	R2	S11	R1	S7	S3	R5	S12	R2	S11	R1	S7	S3	R5	S12	R2	S11	R1	S7

Second seeding - untreated patents (control) = Rep

	Rep II, 10 day		Rep II, 21-23 d		Rep II, 41-43	
Plant 1	P1	P2	P1	P2	P1	P2
Plant 2	P1	P2	P1	P2	P1	P2
Plant 3	P1	P2	P1	P2	P1	P2
Plant 4	P1	P2	P1	P2	P1	P2

Third Seeding & inoculation = Rep III and 3 time points (seeding in all trays in same order)

	Rep III, 10 DAI							Rep III, 21-23 DAI							Rep III, 41-43 DAI						
Plant 1	P1	S1	R10	S10	R5	C-S	R2	P1	S1	R10	S10	R5	C-S	R2	P1	S1	R10	S10	R5	C-S	R2
Plant 1	S5	R6	S11	R1	S2	R8	S8	S5	R6	S11	R1	S2	R8	S8	S5	R6	S11	R1	S2	R8	S8
Plant 1	R9	S7	R7	P2	R12	S9	R3	R9	S7	R7	P2	R12	S9	R3	R9	S7	R7	P2	R12	S9	R3
Plant 1	S6	C-R	S4	R4	S12	R11	S3	S6	C-R	S4	R4	S12	R11	S3	S6	C-R	S4	R4	S12	R11	S3
Plant 2	P1	S1	R10	S10	R5	C-S	R2	P1	S1	R10	S10	R5	C-S	R2	P1	S1	R10	S10	R5	C-S	R2
Plant 2	S5	R6	S11	R1	S2	R8	S8	S5	R6	S11	R1	S2	R8	S8	S5	R6	S11	R1	S2	R8	S8
Plant 2	R9	S7	R7	P2	R12	S9	R3	R9	S7	R7	P2	R12	S9	R3	R9	S7	R7	P2	R12	S9	R3
Plant 2	S6	C-R	S4	R4	S12	R11	S3	S6	C-R	S4	R4	S12	R11	S3	S6	C-R	S4	R4	S12	R11	S3
Plant 3	P1	S1	R10	S10	R5	C-S	R2	P1	S1	R10	S10	R5	C-S	R2	P1	S1	R10	S10	R5	C-S	R2
Plant 3	S5	R6	S11	R1	S2	R8	S8	S5	R6	S11	R1	S2	R8	S8	S5	R6	S11	R1	S2	R8	S8
Plant 3	R9	S7	R7	P2	R12	S9	R3	R9	S7	R7	P2	R12	S9	R3	R9	S7	R7	P2	R12	S9	R3
Plant 3	S6	C-R	S4	R4	S12	R11	S3	S6	C-R	S4	R4	S12	R11	S3	S6	C-R	S4	R4	S12	R11	S3

Third seeding - untreated patents (control) = Rep III

	Rep III, 10 day		Rep III, 21-23 d		Rep III, 41-43	
Plant 1	P1	P2	P1	P2	P1	P2
Plant 2	P1	P2	P1	P2	P1	P2
Plant 3	P1	P2	P1	P2	P1	P2
Plant 4	P1	P2	P1	P2	P1	P2

7. Research accomplishments: *(Describe progress towards meeting objectives. Please use revised objectives if Ministry-approved revisions have been made to original objectives.)*

The annual reports provided the progress up to the period of February 2015 contained the details on the basis of Objectives. The final report consolidates these details and organizes the objectives as approved in the proposal and the progress.

Identification of a spring canola (*B. napus*) population for in-depth investigations:

Because clubroot is an old European disease and it is relatively a new disease in Canada, an oilseed *B. napus* source of R genes is limited to a European winter line (Mendel). Other sources include *B. oleraceae*, rutabaga, Chinese cabbage, and oilseed *B. rapa*, *B. juncea* and *B. carinata*. As noted below, we have focused on *B. napus* and oilseed *B. rapa* because it will be a relatively shorter pathway from gene discovery to introgression of the trait.

In the early part of this work, we attempted to find more resistance genotypes for the molecular work. *B. rapa* Purple top forage turnip (PTFT-2) is a vernalization-requiring, clubroot resistant line. We obtained from Drs. Gary Peng and Kevin Falk of Agriculture and Agri-Food Canada (AAFC) under MTA: Forty (40) seeds each of (a) *B. rapa* Purple top forage turnip (PTFT-2); (b) B. *rapa* L. ACDC 10; F1 from ACDC10 X PTFT-2. *B. rapa* is generally self-incompatible, and therefore obtaining homozygous pure lines requires artificial methods such as hand-pollination at green bud stage prior to manifestation of self-incompatibility or pollination following treatment with salt solution to reduce self-incompatibility. ACDC 10 is a self-compatible version of *B. rapa* developed at AAFC. Our goal was to transfer the disease resistance gene of PTFT-2 (self-incompatible) to ACDC 10 (self-compatible) to derive F_2 progenies that would be capable of flowering without a vernalization requirement and without the impediment of self-incompatibility. We grew PTFT-2 with a vernalization treatment recommended by AAFC. We needed a larger number of F_2 seeds for F_2:F_3 mapping, but found that this line was recalcitrant to setting adequate seeds. Concurrent to the above, *B. rapa* ssp. *perkinensis* Bilko has also been found to be resistant to Pathotype 3. This is also a self-incompatible line. In preliminary experiments, F_2:F_3 screening has identified resistant, homozygous lines (H Rahman). Since we required a pool of sensitive segregants as well these efforts had to be continued outside of our labs and did not appear that these would become available for further work within the timeframe of this project. *B. napus* var. *napobrassica* resistant to Pathotype 3 has also been found (H Rahman) but required more time for developing segregating population than the time frame of the project. Therefore, we focused on the spring canola population identified in Methods. All our further work on Brassica is on this population.

We worked with Dr. Habibur Rahman at the University of Alberta. In his laboratory, there is a dedicated plant growth facility that is used for clubroot screening (NRC does not have a containment facility and therefore we did not perform disease assays in Saskatoon). We have screened and confirmed that *B. napus* cv. Mendel is resistant to Pathotype 3, which is of importance to western Canada. An example of the disease resistant (left) and sensitive phenotype (right) phenotype in *B. napus* is shown in Figure 2 below.

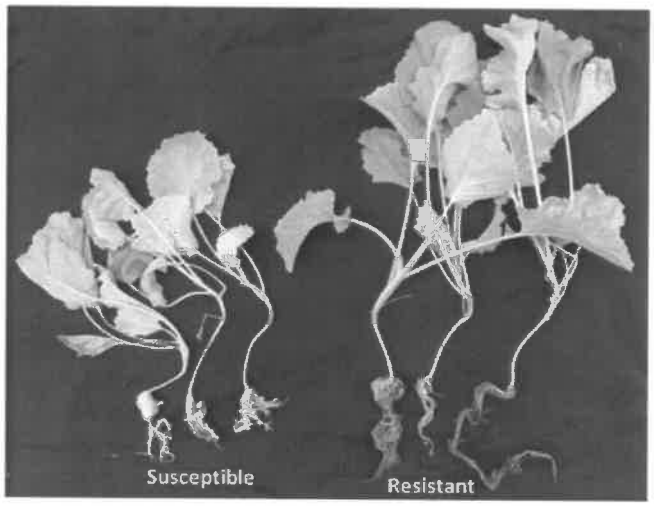


Fig 2. Phenotyping of *B. napus*. Phenotype rating is based on the visual appearance of the gall (Kuginuki et al. 1999).

Confirmation of disease phenotype:

A disease screening trial that included two parents (resistant parent and susceptible parent of *B. napus*); two additional accessions (one resistant and one susceptible line other than resistant and susceptible parent); 12 doubled haploid individuals each from the resistant and susceptible categories was undertaken. These 12 individuals were selected based on their earlier stable reaction for Pathotype 3 from at least 3 earlier screenings. There were three individual plants in each replication and each entry was replicated three times. The experimental plan is shown in Fig. 1. Root dip inoculation method was followed for screening as described by Rahman et al (2011). Three time points were used for sample collection – 10, 22 and 42 DAI (days after inoculation) which represent the stages of primary infection, secondary infection and late infection stage at maturity, respectively.

This screening trial confirmed earlier results and a clear resistant and susceptible reaction were observed. As shown in Fig. 3, the green highlight indicates clubroot resistant material including the resistant parent, the additional resistant accession as well as 12 individual double haploids which showed consistent score of '0' indicating no gall formation. Susceptible parent, another susceptible accession that was included in the study, and 12 individual doubled haploid lines exhibited a clear susceptible reaction by formation of large galls and these are highlighted in yellow in Fig 3.

From this screening we concluded that that we could use the 12 resistant types and 12 sensitive types among in the DH population for further in-depth investigations.

New Reg Number	P1	P2	P3	P4	P5	P6	P7	P8	Mean	R/S	P1	P2	P3	P4	P5	P6	P7	P8	Mean	R/S	P1	P2	P3	P4	P5	P6	P7	P8	Mean	R/S
1RA1333.005-A1201	0	0	0	0	0	-	-	-	0.00	R	0	0	0	0	0	0	0	0	0.00	R	0	0	0	0	0	0	-	-	0.00	R
1RA1333.006-A1201	0	0	0	0	0	0	0	0	0.00	R	0	0	0	0	0	0	0	-	0.00	R	0	0	0	0	0	0	0	0	0.00	R
1RA1333.008-A1201	0	0	0	0	0	0	0	-	0.00	R	0	0	0	0	0	0	0	-	0.00	R	0	0	0	0	0	0	0	0	0.00	R
1RA1333.017-A1201	0	0	0	0	0	0	-	-	0.00	R	0	0	0	0	0	0	0	-	0.00	R	0	0	0	0	0	0	0	0	0.00	R
1RA1333.019-A1201	0	0	0	0	0	0	-	-	0.00	R	0	0	0	0	0	0	0	-	0.00	R	0	0	0	0	0	0	0	0	0.00	R
1RA1333.022-A1201	0	0	0	0	0	0	-	-	0.00	R	0	0	0	0	0	0	0	-	0.00	R	0	0	0	0	0	0	0	0	0.00	R
1RA1333.028-A1201	0	0	0	0	0	0	-	-	0.00	R	0	0	0	0	0	0	0	0	0.00	R	0	0	0	0	0	0	0	0	0.00	R
1RA1333.031-A1201	0	0	0	0	0	0	0	0	0.00	R	0	0	0	0	0	0	0	0	0.00	R	0	0	0	0	0	0	-	-	0.00	R
1RA1333.038-A1201	0	0	0	0	0	0	0	0	0.00	R	0	0	0	0	0	0	-	-	0.00	R	0	0	0	0	0	0	0	0	0.00	R
1RA1333.047-A1201	0	0	0	0	0	0	-	-	0.00	R	0	0	0	0	0	0	-	-	0.00	R	0	0	0	0	0	0	0	0	0.00	R
1RA1333.050-A1201	0	0	0	0	0	0	-	-	0.00	R	0	0	0	0	0	-	-	-	0.00	R	0	0	0	0	0	0	-	-	0.00	R
1RA1333.051-A1201	0	0	0	0	0	-	-	-	0.00	R	0	0	0	0	0	0	-	-	0.00	R	0	0	0	0	0	0	0	0	0.00	R
1RA1333.052-A1201	0	0	0	0	0	0	-	-	0.00	R	0	0	0	0	0	0	-	-	0.00	R	0	0	0	0	0	0	0	0	0.00	R
1RA1333.060-A1201	0	0	0	0	0	0	0	-	0.00	R	0	0	0	0	0	0	-	-	0.00	R	0	0	0	0	0	0	0	0	0.00	R
1RA1333.061-A1201	0	0	0	0	0	0	0	-	0.00	R	0	0	0	0	0	0	-	-	0.00	R	0	0	0	0	0	0	0	0	0.00	R
1RA1333.063-A1201	0	0	0	0	0	0	-	-	0.00	R	0	0	0	0	0	0	-	-	0.00	R	0	0	0	0	0	0	0	0	0.00	R
1RA1333.010-A1201	3	3	3	3	3	3	-	-	3.00	S	3	3	3	3	3	3	-	-	3.00	S	3	3	3	3	3	3	3	3	3.00	S
1RA1333.013-A1201	3	3	3	3	3	-	-	-	3.00	S	3	3	3	3	3	3	-	-	3.00	S	3	3	3	3	3	3	3	3	3.00	S
1RA1333.018-A1201	3	3	3	3	3	-	-	-	3.00	S	3	3	3	3	3	3	-	-	3.00	S	3	3	3	3	3	3	3	3	3.00	S
1RA1333.020-A1201	3	3	3	3	3	3	-	-	3.00	S	3	3	3	3	3	-	-	-	3.00	S	3	3	3	3	3	3	3	3	3.00	S
1RA1333.029-A1201	3	3	3	3	-	-	-	-	3.00	S	3	3	3	3	3	3	3	-	3.00	S	3	3	3	3	3	-	-	-	3.00	S
1RA1333.033-A1201	3	3	3	3	3	3	3	-	3.00	S	3	3	3	3	3	3	3	3	3.00	S	3	3	3	3	3	3	3	3	3.00	S
1RA1333.034-A1201	3	3	3	3	3	-	-	-	3.00	S	3	3	3	3	3	3	3	3	3.00	S	3	3	3	3	-	-	-	-	3.00	S
1RA1333.035-A1201	3	3	3	3	3	3	3	-	3.00	S	3	3	3	3	3	3	3	3	3.00	S	3	3	3	3	3	3	-	-	3.00	S
1RA1333.036-A1201	3	3	3	3	3	3	3	3	3.00	S	3	3	3	3	3	3	3	-	3.00	S	3	3	3	3	3	3	3	3	3.00	S
1RA1333.040-A1201	3	3	3	3	3	-	-	-	3.00	S	3	3	3	3	3	3	-	-	3.00	S	3	3	3	3	3	3	3	3	3.00	S
1RA1333.041-A1201	3	3	3	3	3	-	-	-	3.00	S	3	3	3	3	3	3	3	3	3.00	S	3	3	3	3	-	-	-	-	3.00	S
1RA1333.048-A1201	3	3	3	3	3	3	-	-	3.00	S	3	3	3	3	3	3	3	-	3.00	S	3	3	3	3	3	3	3	3	3.00	S
1RA1333.054-A1201	3	3	3	3	3	-	-	-	3.00	S	3	3	3	3	3	3	-	-	3.00	S	3	3	3	3	3	3	-	-	3.00	S
1RA1333.055-A1201	3	3	3	3	3	3	-	-	3.00	S	3	3	3	3	3	3	-	-	3.00	S	3	3	3	3	3	3	3	3	3.00	S
1RA1333.058-A1201	3	3	3	3	3	3	3	-	3.00	S	3	3	3	3	3	3	-	-	3.00	S	3	3	3	3	3	3	3	3	3.00	S
1RA1333.059-A1201	3	3	3	3	3	3	-	-	3.00	S	3	3	3	3	3	3	-	-	3.00	S	3	3	3	3	3	3	3	3	3.00	S

Fig. 3. Rating of clubroot disease reaction after inoculation with pathotype-3. The rating is based on Kuginuki et al. (1999) where 0= no galls; 1= a few small galls; 2= moderate galls and 3= severe galls.

Reference Transcriptomics:

After completing the disease phenotyping to obtain 12 DH lines each for resistant and susceptible phenotypes, we performed RNAseq. We standardized the protocol for RNA isolation from roots. Using this protocol we have prepared RNA from infected roots at three stages of disease cycle (primary infection; secondary infection; mature stage). Clubroot infection was also confirmed by using pathotype specific primers reported by Sundelin et al. (2010). These are ITS primers specific to *P. brassicae*. Single 139 bp band was amplified to confirm presence of clubroot pathogen (Fig.4). These bands were sequenced and confirmed to match with clubroot pathogen ITS sequence.

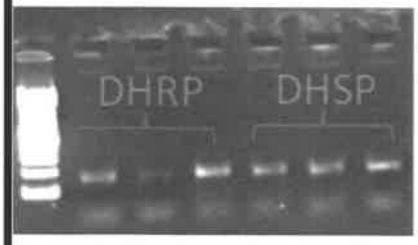


Fig. 4. Amplification of the ITS region of *P. brassicae* from clubroot-challenged roots. DHRP-Doubled haploid resistant pool; DHSP-Doubled haploid susceptible pool; Marker 100 bp ladder. Three biological replicates are shown under the braces in the figure.

P. brassicae inoculated and control (mock inoculated) root samples were harvested at 10 DAI (primary infection); 22 DAI (Secondary infection) and 42 DAI (maturation phase) and used for RNA isolation and transcriptome sequencing. Details of experimental plan are given in Fig 5. Disease scoring clearly gave disease response contrast with disease resistant parent as well as derived doubled haploid individuals scoring "0" (indicating least effect of pathogen) whereas susceptible parent as well as derived doubled haploid individuals scoring "3" (indicating highest severity of clubbed roots) as per Kuginuki et al. (1999).

mRNA from the roots of the parents and pools of resistant and susceptible DH lines were isolated and RNASeq data were generated. As shown in Fig.5 data analysis comprised of comparison between clubroot resistant and susceptible parent, doubled haploid pools representing resistant and susceptible recombinants in doubled haploid mapping population. One set of control resistant and susceptible accession was also used as independent accessions. This RNAseq data was analyzed using Tuxedo suite pipeline (Trapnell et al., 2012). It was also analyzed using Venny software to identify unique transcripts expressed in resistant and susceptible condition.

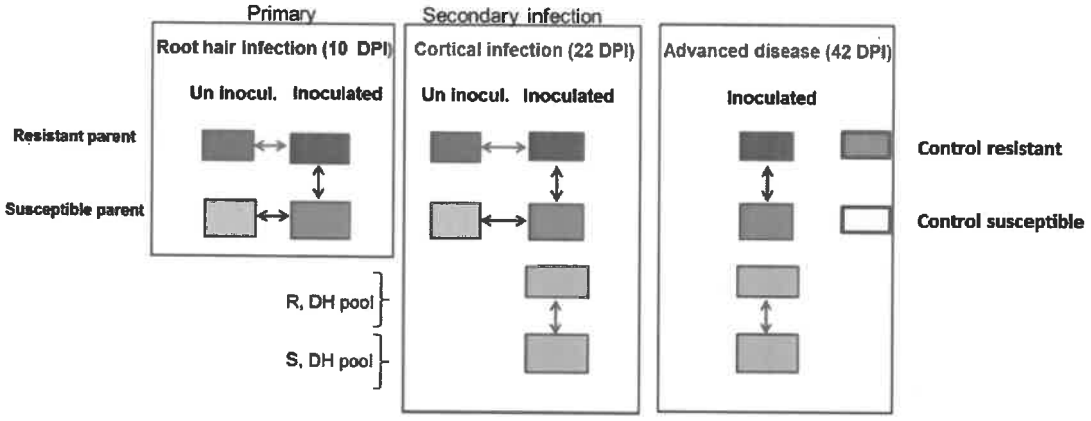


Fig. 5 Experimental design for *B. napus* root transcriptome sequencing to assess clubroot response to pathotype 3 resistance and susceptible reactions. DPI, days post inoculation.

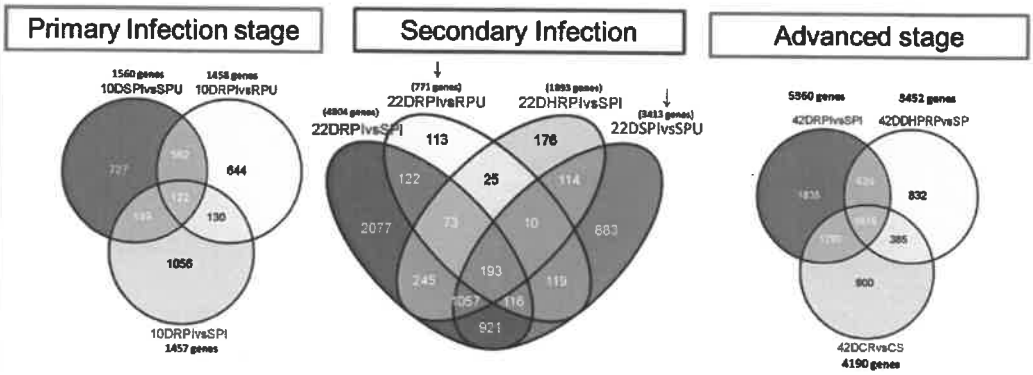


Fig. 6 Venn diagram of B. napus root transcriptome data indicating unique gene expressed in resistant as well as susceptible accessions at three disease developmental stages i.e. Primary and secondary infection stage as well as

advanced disease developmental stage. SPU-susceptible parent un-inoculated (control); SPI-susceptible parent inoculated; RPU- resistant parent un-inoculated (control); RPI-resistant parent inoculated; DHPR- doubled haploid pool resistant; DHPS- doubled haploid pool susceptible; 10D- 10 days post inoculation representing primary infection stage ; 22D-22 days post inoculation representing secondary infection stage; 42D- 42 days post inoculation representing advanced disease developmental stage.

Statistical analysis revealed significantly up and down regulated genes expression change (at least 4-fold; FDR > 0.05) in each comparison. Out of the 76,946 transcripts predicted from the *B. napus* genome, around 1.9% genes were differentially expressed in the resistant parent at primary infection stage as compared to the susceptible parent. The differences were much more contrasting at the secondary infection stage. Around 1.0% were differentially expressed in resistant parent as compared to control whereas around 4.4% genes are differentially expressed in susceptible accession as compared to the control. It confirms the inferences from pathology that the secondary stage of infection could be a better indicator of differences between resistant and susceptible reactions. Venn diagram comparisons indicated 644 unique specific differentially expressed transcripts in resistant parent ("referred to as resistant transcripts" and 727 such transcripts in the susceptible parent at the primary infection stage. There was a greater contrast between the "resistant transcripts" of 113 and 883 "susceptible transcripts", indicating very strong compatible reaction at the secondary infection stage. The number of differentially expressed genes were highest at maturation phase (5360) as compared to secondary infection stage (Fig.7). The data were also analyzed to identify the up- or down-regulated genes in the case of resistant compared with susceptible case for each dataset. There were 2719 genes up regulated in the resistant parent and 2641 genes were up-regulated in the susceptible parent at 42 DPI stage. This contrast reduced in doubled haploid pools indicating the advantage of individual lines from the mapping population to find specific differences in disease resistance transcriptome studies (Fig.7).

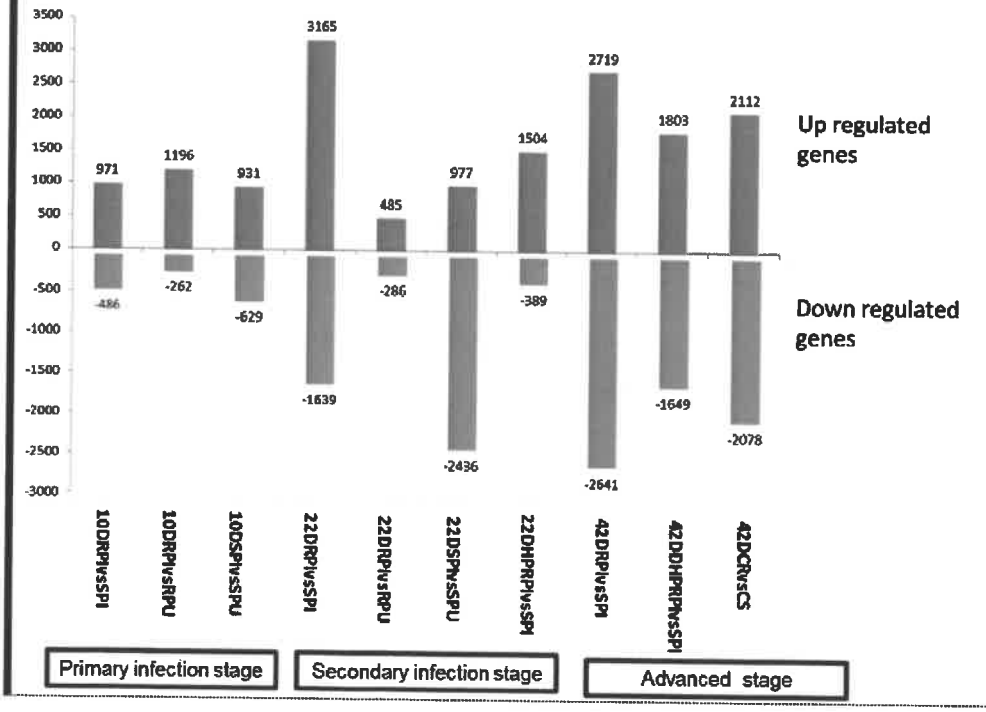


Fig.7. Differentially up and down regulated genes in *B. napus* root transcriptome data at three disease developmental stages i.e. Primary and Secondary infection as well as advanced disease developmental stage.

MapMan and PageMan analysis for functional annotation of key disease resistance pathways:

We performed MAPMAN analysis to identify the key metabolic pathways in disease resistance reaction as well as the pathways that associate with a strong compatible reaction. MapMan was also used for cluster analysis at the primary, secondary infection (reported earlier) and maturation stages; the data were further processed using PageMan software for generating heat maps of co-regulated genes.

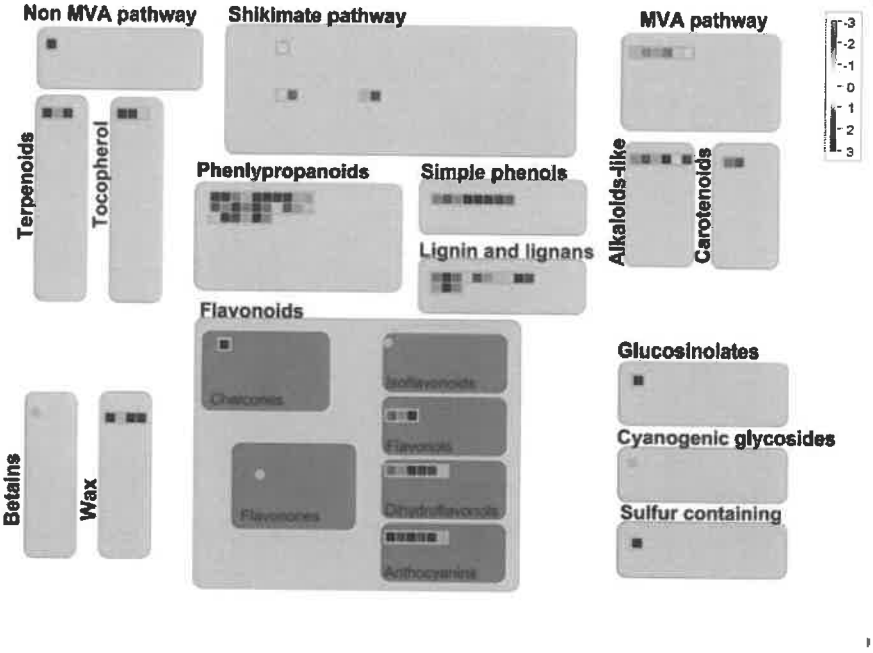


Fig.9. MapMan output indicating higher expression of phenylpropanoid and lignin biosynthesis genes in resistant parent at advanced disease development stage (42DPI). Red color indicates higher expression in resistant than susceptible.

MapMan analysis indicated (Fig. 8 and 9) that phenylpropanoid and lignin biosynthesis genes were up-regulated in the resistant parent and pointed potentially defective cell wall lignification in the susceptible parent. Similar results were obtained from 42 DPI stage transcriptome dataset and it agreed with our earlier biochemical analysis for 42 DPI stage (data not shown). This gives interesting results and confirmed importance of cell wall lignification for clubroot resistance mechanism in B. napus. A recent review indicated significant roles of lignin and phenylpropanoid genes in plant disease resistance (Miedes et al. 2014) .

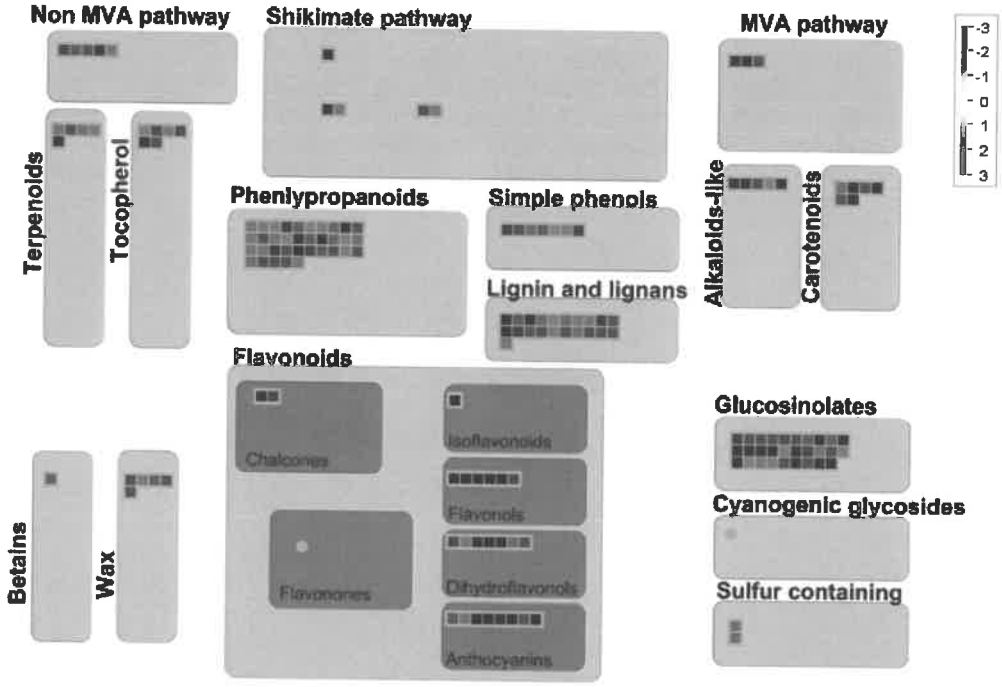


Fig.9. MapMan output indicating higher expression of phenylpropanoid and lignin biosynthesis genes in resistant parent at advanced disease development stage (42DPI). Red color indicates higher expression in resistant than susceptible.

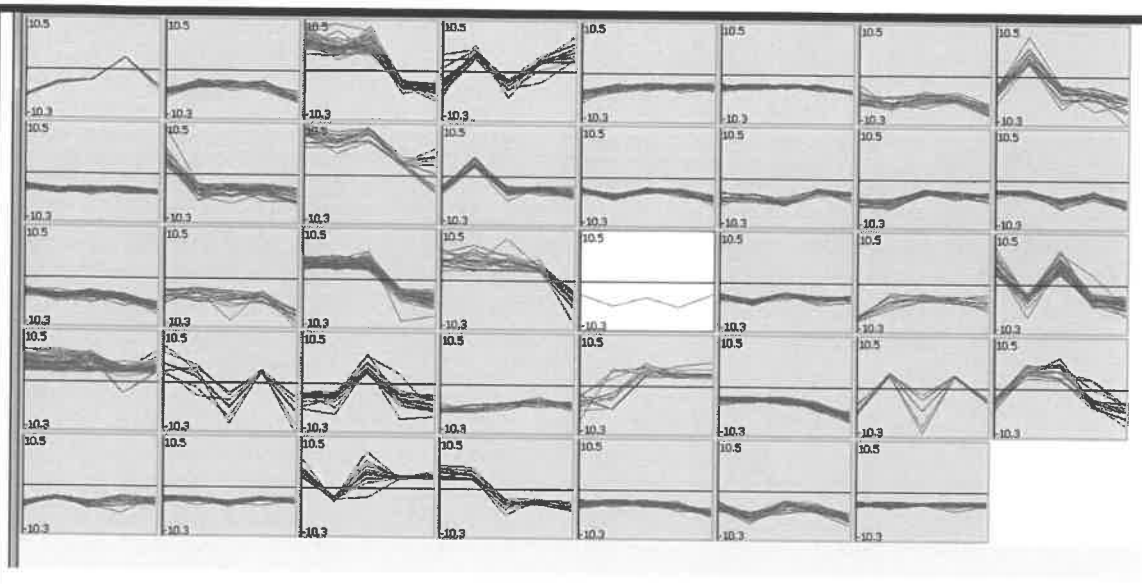


Fig.10 Cluster analysis combined for 22DPI and 42DPI stage for clubroot disease reaction indicating co-expressed genes. Data used for analysis includes 5 datasets (42DPI CR vs CS; 42DPI DHPRP vs DHPSP; 42 DPI RP vs SP; 22DPI DHPRP vs DHPSP; 22DPI RP vs SP). Genes in blue rectangle indicate higher level of expression in all resistant samples as compared to susceptible samples. Genes in purple rectangle indicate higher level of expression in all susceptible samples as compared to resistant samples. Clusters are as per standard MAPMAN Bin codes.

Cluster analysis of Differentially expressed genes:

Differentially expressed genes were clustered using MapMan software (Fig. 10) and functional categorization was done using Pageman software. Wilcoxon Benjamini Hochberg multiple testing correction was applied and genes possibly playing important roles in clubroot disease response were identified.

To get an overview of the transcriptional changes as a portrayal of the resistant as well as susceptible responses, the data were subjected for pathway specific visualization using MapMan (Thimm et al., 2004). The dataset for secondary and advanced disease development stages were further clustered using MapMan to get co-expression network of both stages together. The data from 42 DPI stage were further subjected to PageMan (Usadal et al. 2005) for obtaining a comprehensive heatmap exhibiting co-regulated genes. The genes indicated in red in the figure indicate up-regulation in resistant as compared to susceptible reaction whereas and those in blue were up regulated in susceptible as compared to resistant genotype (Fig.11).

The primary infection stage clearly indicated higher level expression of phenylpropanoid and lignin biosynthesis in the resistant parent. This suggested a stronger resistant response even at the primary infection stage. Higher levels of expression of cellular biosynthesis, auxin metabolism and calcium signaling pathway in the susceptible parent indicated a weaker disease resistance to pathogen attack (Fig.11).

The secondary infection stage cluster analysis clearly indicated a stronger sink development caused by clubroot pathogen infection, and contrasting differences between resistant and susceptible reaction (Fig.12).
The key pathway genes up-regulated in disease resistant parent includes phenylpropanoid biosynthesis, salicylic

acid (SA) metabolism, pathogenesis response (PR) proteins, ubiquitin degradation, nitrate transport, receptor like kinase (RLK) genes. The notable pathways that were highly expressed specifically in the susceptible parent were carbohydrate metabolism, lipid biosynthesis, cell wall biosynthesis, and auxin metabolism, indicating cell proliferation (clubroot development) and establishment of nutrition source for clubroot pathogen. A representative MapMan output indicating higher levels of expression of phenylpropanoid and lignin biosynthesis genes in the resistant parent are shown in Fig. 8 and 9.

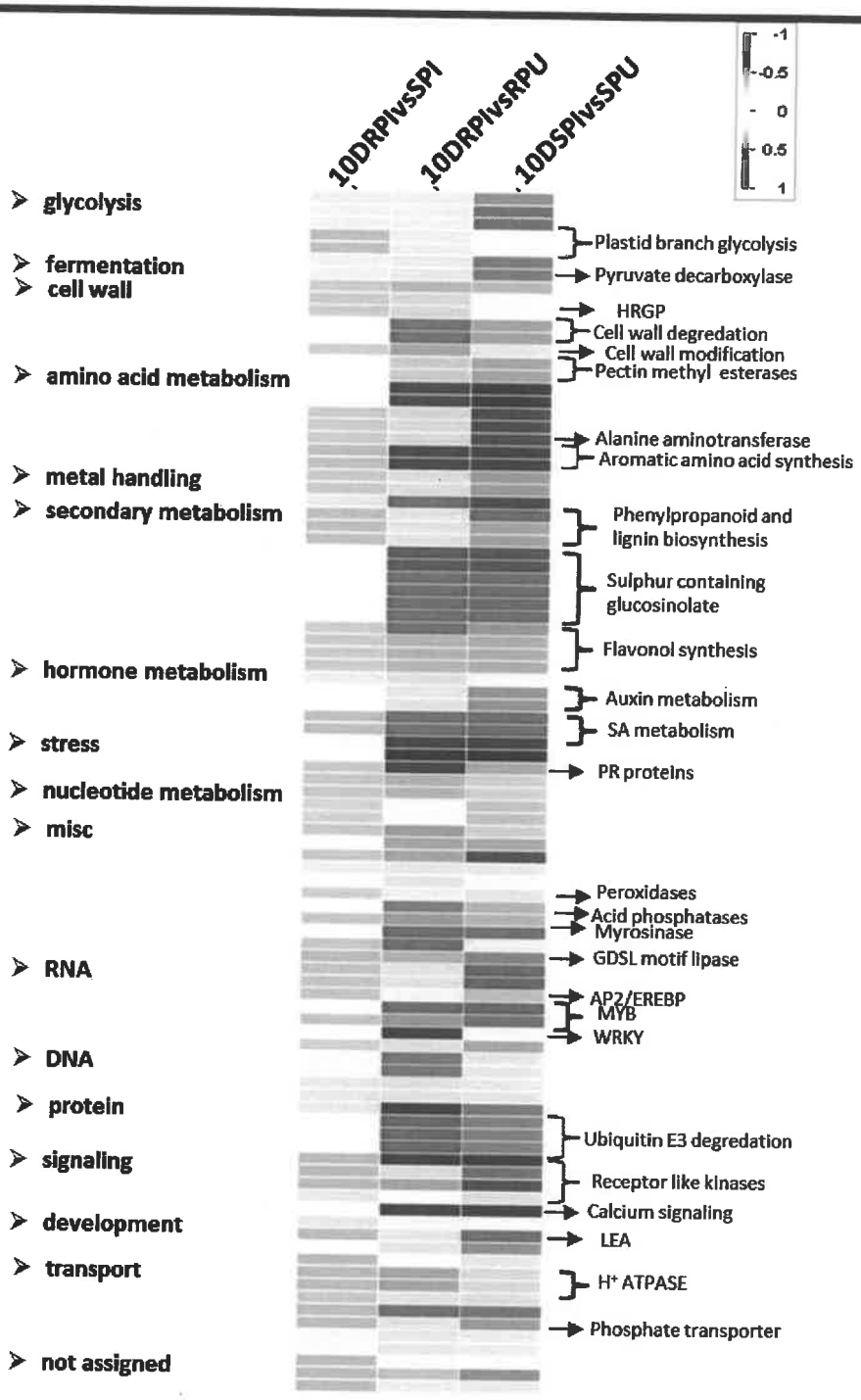


Fig. 11. Heat map generated from clustered genes differentially regulated between resistant and susceptible accessions at primary infection stage challenged with pathotype 3 of *P. brassicae*

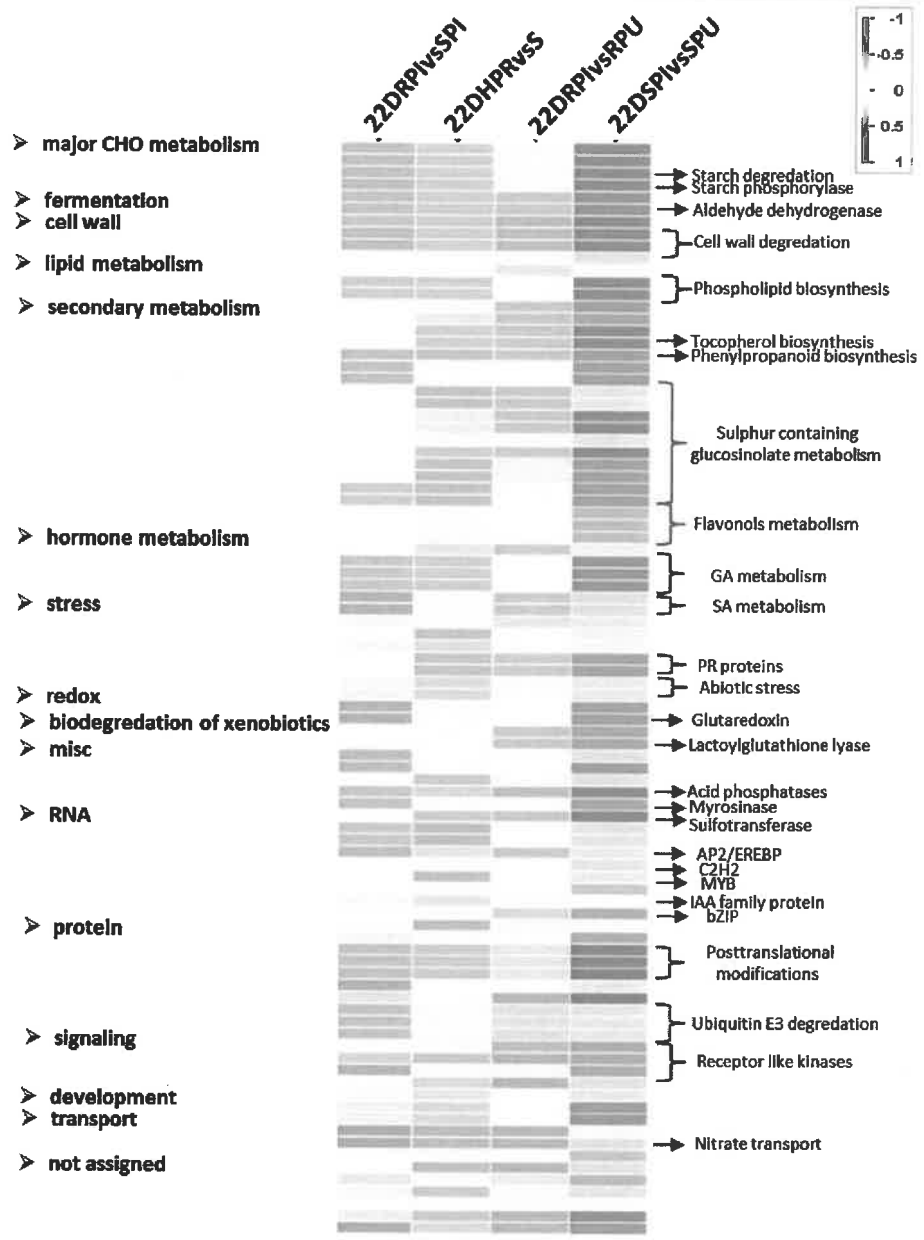


Fig .12. Heat map generated from clustered genes differentially regulated between resistant and susceptible accessions at secondary infection stage challenged with pathotype 3 of *P. brassicae*

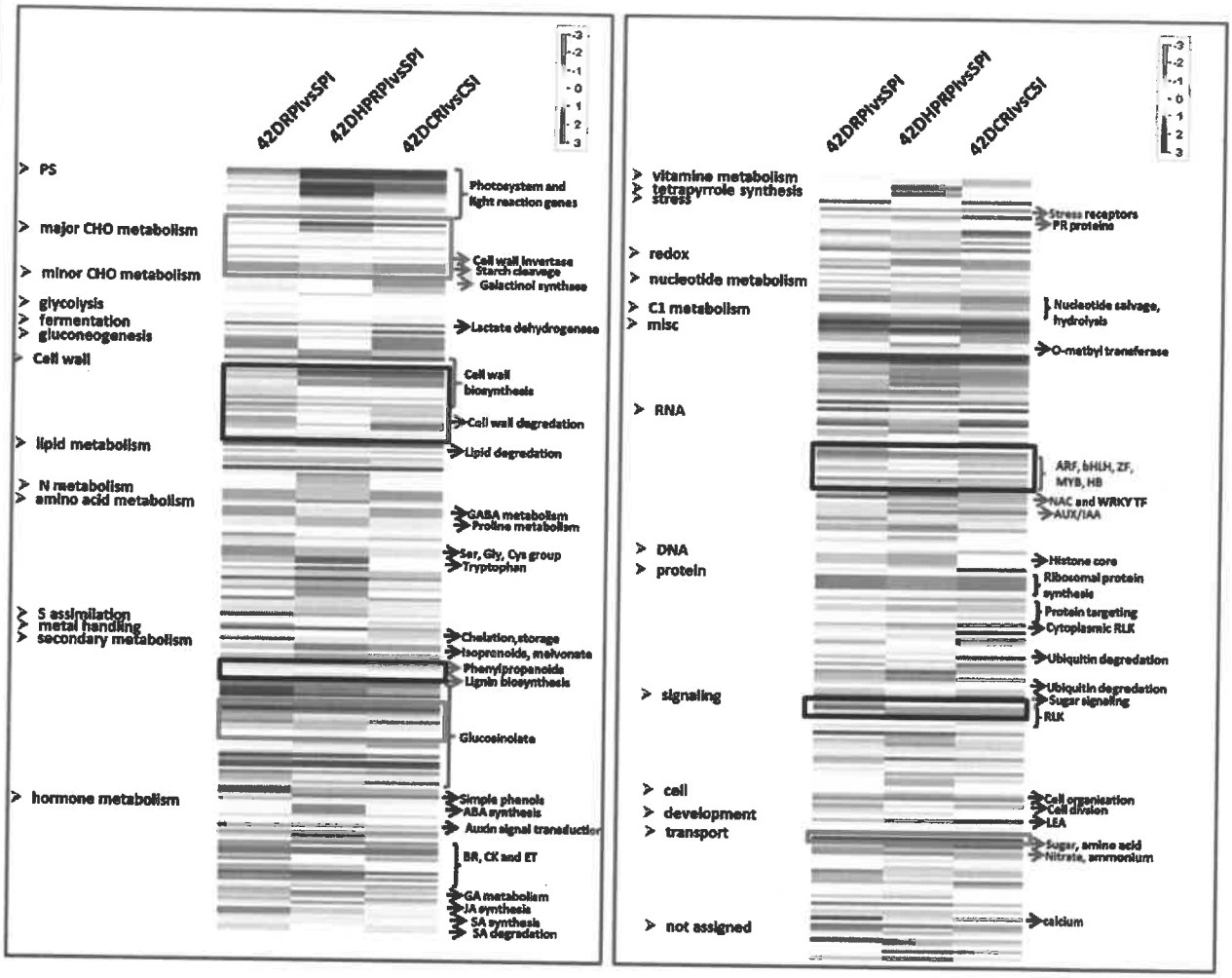
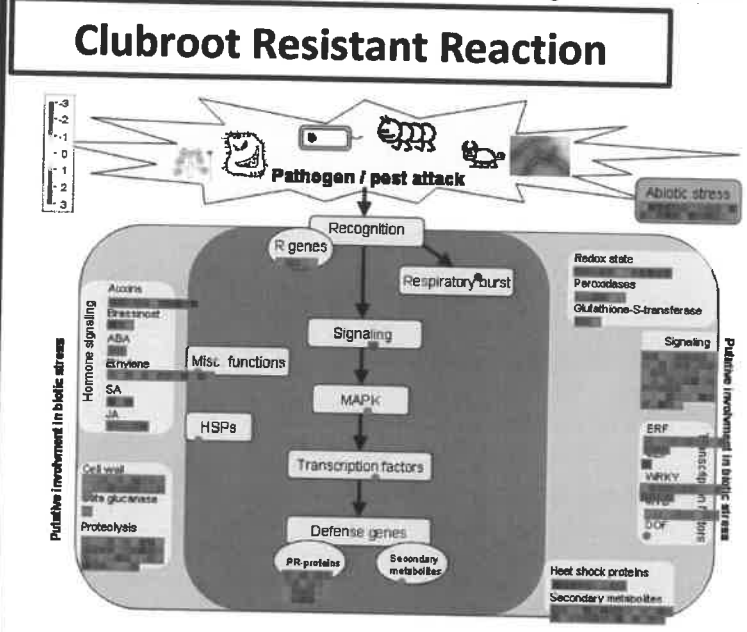


Fig.13. Heat map generated from clustered genes differentially regulated between resistant and susceptible accessions at advanced disease developmental stage (42DPI) challenged with pathotype 3 of *P. brassicae*. Resistance enhancing cascades are indicated by black rectangle and susceptibility enhancing cascades are indicated by green rectangles.

Clubroot infection resulted in higher metabolic activities, such as major carbohydrate synthesis (starch, sucrose and glycolysis), TCA cycle and lipid metabolism. With respect to metabolic changes, the major differences could be observed between resistant and susceptible genotype consistently at advanced disease development stage. Mapman clustering also indicated induction of metabolic pathways, especially secondary metabolites, N and S metabolism and disease response signaling cascades.

With regard to secondary metabolite synthesis , the differential regulation of phenylpropanoids, lignin biosynthesis were consistently up- regulated in resistant parent (at primary, secondary infection as well as

advanced disease development stage . It was also interesting to note that, glucosinolate biosynthesis pathway genes were suppressed in resistant parent specifically at secondary infection stage. We have earlier reported higher glucosinolate levels (4-5) times higher in susceptible parent as a result of clubroot infection using biotechnical methods (last year's progress report). Plant defense-associated hormones such as SA, JA and ET were differentially regulated at primary and secondary infection level. Overall it was imperative from cluster analysis that, JA, ET and SA synthesis genes were up-regulated in resistant parent indicating strong resistance response.



Auxin and cytokinins (plant growth –promoting hormones) are important for clubroot development. Increased expression of Auxin-related genes like GH3; IAA-amino acid conjugate hydrolases, nitrilases were up-regulated strongly in susceptible parent as compared to resistant parent.

In all, PR proteins, SA, ET and JA hormones, phenylpropanoids and signaling plays major role in clubroot disease resistance as shown in summary picture of disease resitant parent and susceptible parent at secondary infection stage. Similar results were confirmed at advanced disease development stage.

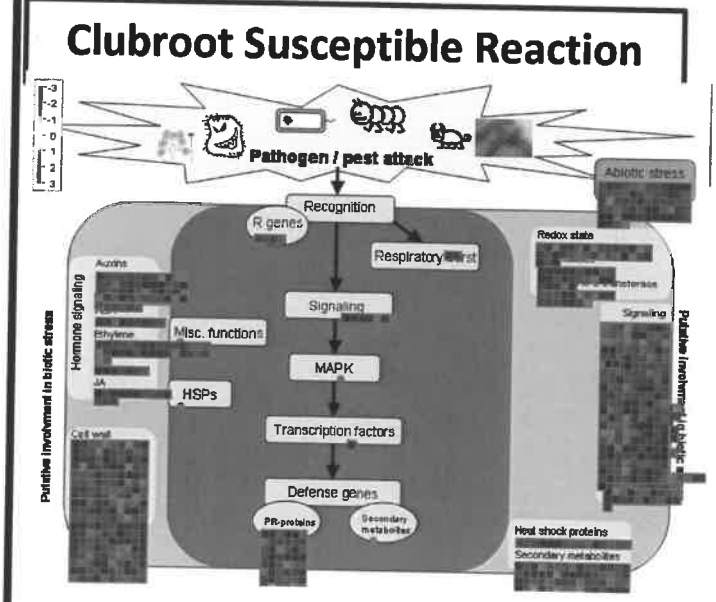


Fig.14. Biotic stress related gene expression profile of clubroot resistant and susceptible parents at secondary infection stage (22 DAI stage). Each square symbolizes a differentially expressed gene related to specific biosynthesis pathway. Numbers +3 to -3 on the color scale represent log2 value of the fold change between inoculated and non-inoculated root transcriptome. Red color indicates enhanced expression of genes whereas blue color indicates repressed genes.

Genetic mapping:

In collaboration with the laboratory of Dr. H. Rahman at the University of Alberta, initially we developed genetic map using SSR markers to pinpoint clubroot resistance loci using a DH population derived from a cross between A07-26NR (susceptible parent) x 1CA0591.263 (Resistant Parent derived from cv. Mendel). More than 800 SSR markers were initially screened using bulked segregant analysis (BSA) to identify linked markers and these markers were run on a DH mapping population of 94. This identified a tightly linked marker BnGMS291 (later named CR1) with ~ 2% recombination for resistance locus. As shown in Fig.14 , most of the markers for their order aligned well to *B. rapa* chromosome A3 indicating potential origin of this clubroot resistance loci in *B. napus* from *B. rapa*. *B. rapa* is one of the most important clubroot resistance gene sources as reported in our recent review (Rahman et al., 2013). Further fine mapping to dissect this genomic region to identify candidate genes for clubroot resistance was undertaken.

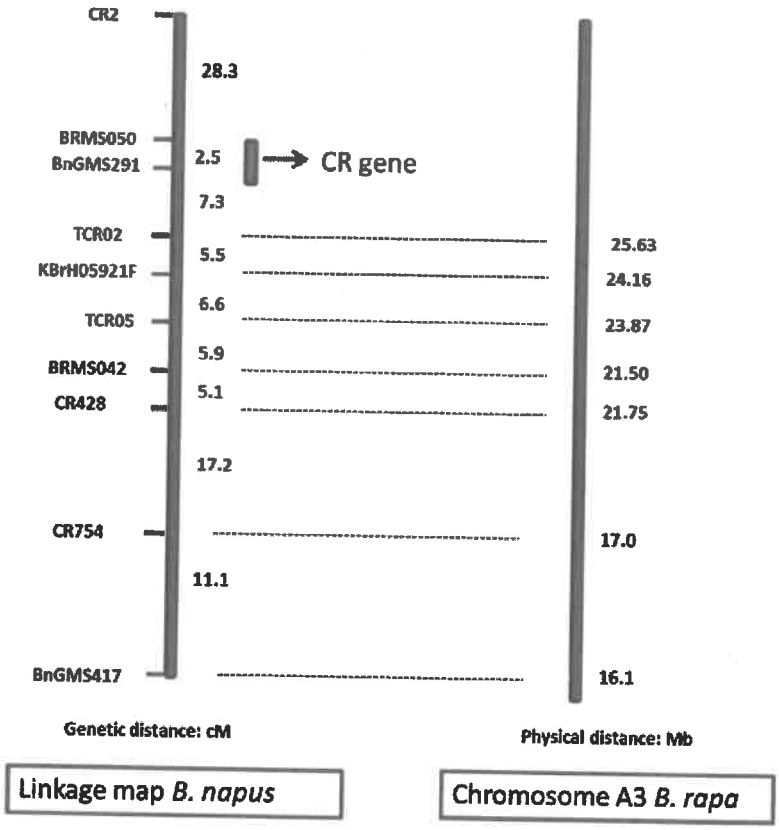


Fig .14. Genetic and physical map of clubroot resistance locus in *B. napus* under present investigation. Physical map reference is *B. rapa* v1.2 genome (2014 update)

This initial map was further updated with SNP mapping using *B. napus* as reference genome (Fig.15) We also performed Composite Interval Mapping (with 500 permutations tested at $p < 0.05$ indicated a LOD value of 13.0 at position 121.3 cM indicating tight linkage with clubroot resistance genomic region and further future use for clubroot resistance mapping (Fig.16). This marker can explain 51.0% (i.e. R^2=0.51) genetic variation about clubroot resistance response. On a 4 point scale for clubroot resistance (1 being resistant and 4 being susceptible presence of this loci could increase the resistance by around 50%. It also indicates that there is a need to stack more resistance loci or resistance supporting genes with major clubroot resistance locus to build better durable resistance. Earlier research in Arabidopsis reported markers with a marker with LOD 14.9 explaining 22.9% genetic variation (Jubault et al., 2008). In another QTL mapping work, genetic mapping of *Brassica rapa* for clubroot resistance identified a QTL region for Pb7 isolate (China) on A08 chromosome (PbBa8.1) with a LOD of 8.5 and could explain 35.2% genetic variation (Chen et al.,2013). With these efforts, we have successfully mapped clubroot resistance loci for Mendel derived resistance. This marker can be used in Marker Assisted Selection (MAS) of clubroot resistance. This will be a component in stacking multiple clubroot resistance loci to breed durable clubroot resistance.

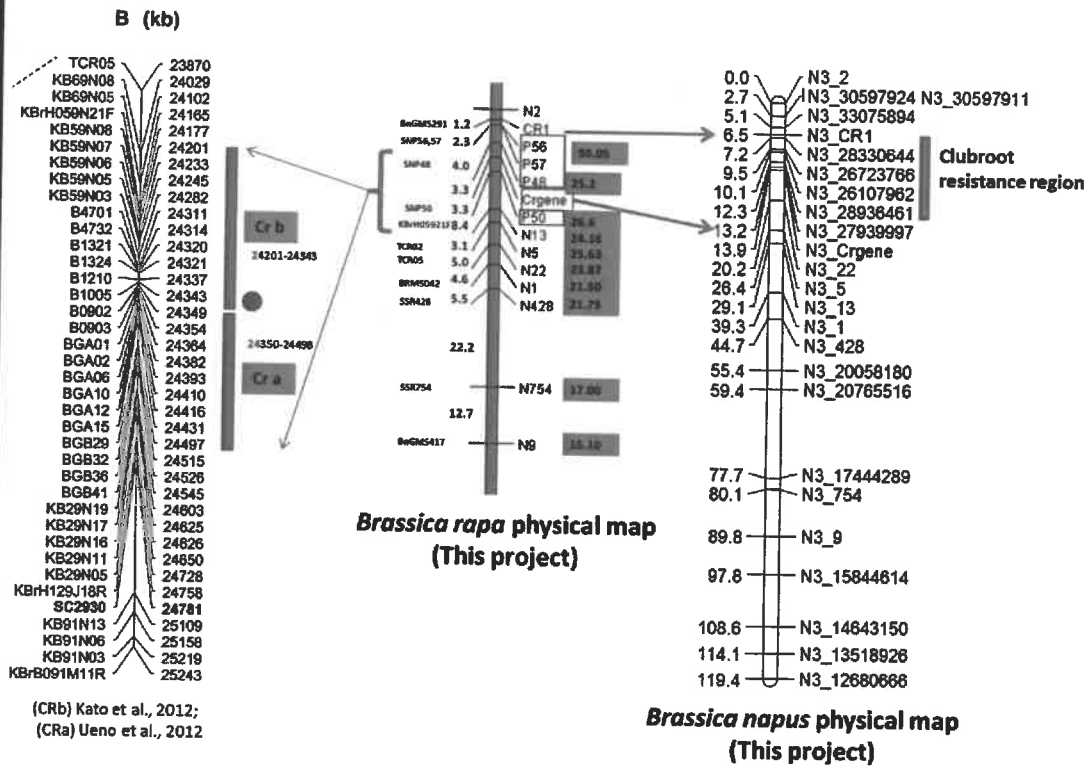


Fig. 15. An updated genetic and physical map for clubroot resistance region that combines SSR and SNP markers together, compared between earlier *Brassica rapa* physical map and *B. napus* physical map. Physical regions from earlier published work for clubroot resistance loci from Kato et al., 2012 (CRb) and Uneo et al., 2012 (CRa) are also shown. Number on the right indicates physical location of SNP markers on *B. napus* genome (v3). Clubroot

resistance region is indicated in red bar (from CR1 to N3_27939997). (2015 final update)

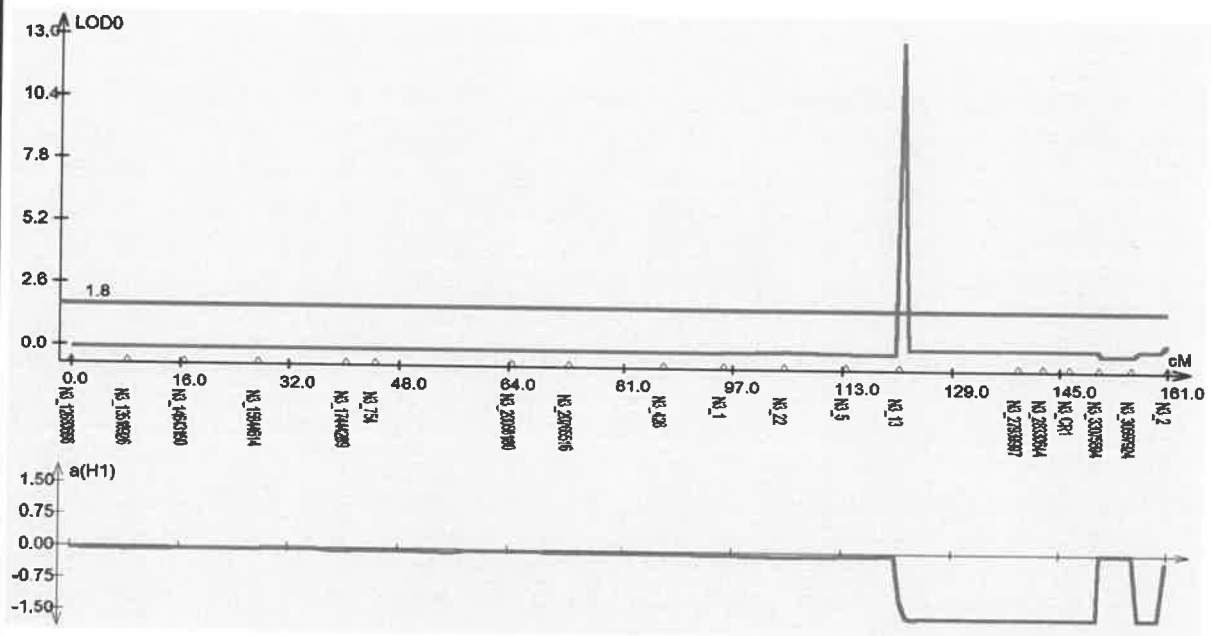


Fig. 16. Composite interval mapping for identification of major QTL region for clubroot resistance derived from Mendel for P3 isolate from Alberta and Saskatchewan region.

Small RNA sequencing and analysis in relation to clubroot resistant and susceptible reaction

48 small RNA libraries were prepared using Illumina TrueSeq Small RNA libraries and sequenced using Illumina HiSeq2500 v4 chemistry (Donnelly Sequencing Centre, University of Toronto). We annotated more than 400 million small RNA reads using CLC Bio software workbench 7 and mapped them using *B. napus* microRNA database miRBase Release 21. Small RNAs were annotated, mapped and merged using CLC Bio pipeline and mapped to *B. napus* v.1 genome to get genome wide abundance of small RNA targeting *B. napus* genes. Differentially expressed and unique to resistant and susceptible reaction were identified specifically at secondary infection stage targeting genomic regions on chromosome N3 (this is the chromosome where Clubroot resistance gene for P3 pathotype derived from Mendel is mapped). Unique regions specific to resistant and susceptible reactions were identified (Fig.17). We also did cluster analysis to identify small RNA which are specifically expressed at higher level in resistant reaction (Fig.18). These include MIR 397a,b; MIR156e; MIR396a; MIR167c, MIR156d,f,a; MIR168a and MIR6031. Interestingly these small RNA were expressed significantly higher level in secondary as well as maturation phase resistant parent. This indicates possible important role of small RNA in

resistant reaction. Sequence and category of small RNA identified in this study are given in Table1.

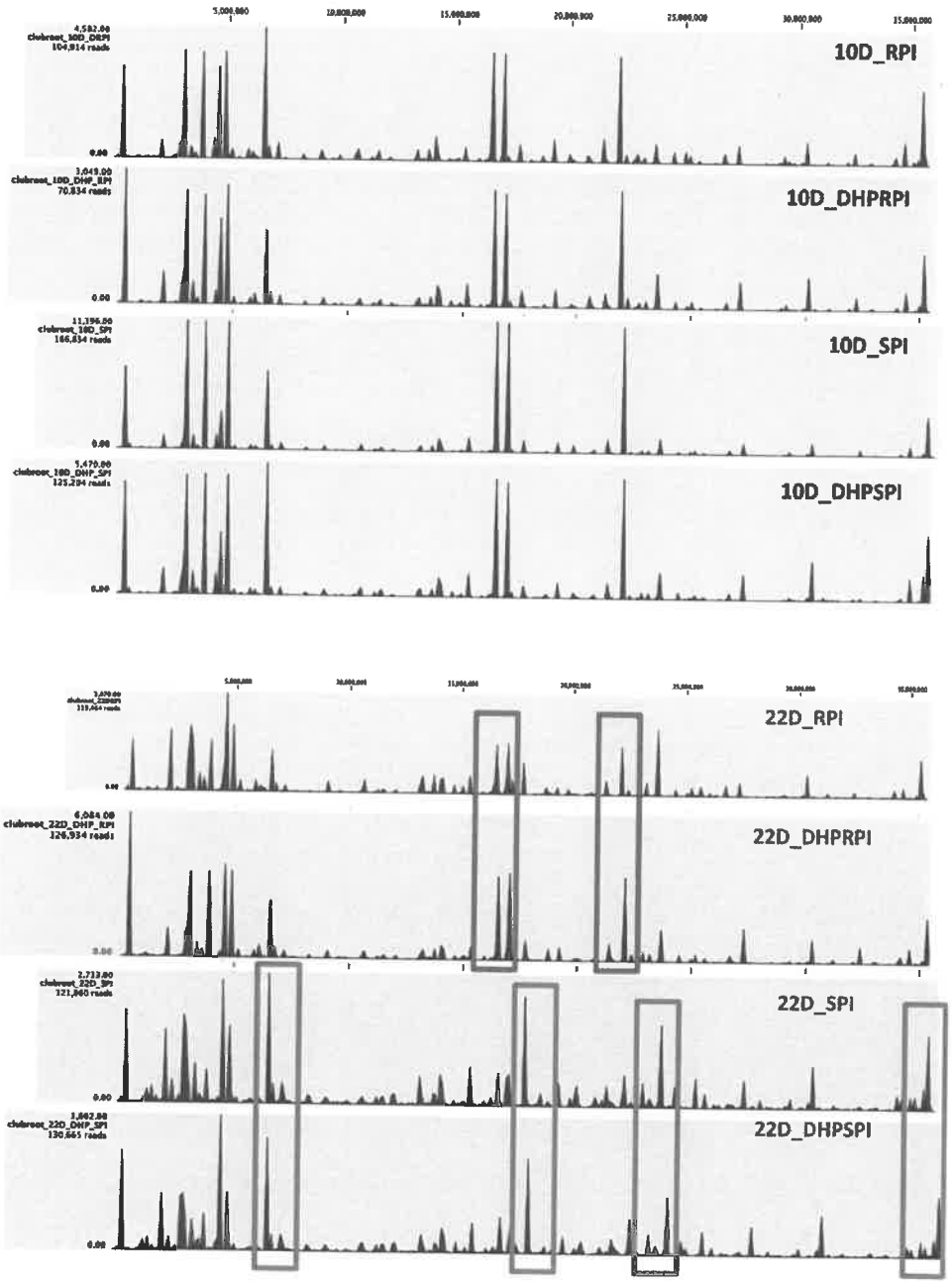


Fig.17. small RNA mapping of clubroot infected samples to N3 chromosome of *B. napus* genome. Samples include 10 D resistant Parent Inoculated (10D_RPI); 10 D doubled haploid pool of resistant population inoculated (10D_DHPRPI); 10 D susceptible parent Inoculated (10D_SPI); 10 D doubled haploid pool of susceptible population inoculated (10D_DHPSPI); 22 D resistant Parent Inoculated (22D_RPI); 22 D doubled haploid pool of resistant population inoculated (22D_DHPRPI); 22 D susceptible parent Inoculated (22D_SPI); 22 D doubled haploid pool of susceptible population inoculated (22D_DHPSPI). Differentially expressed small RNA higher in resistant reaction

are indicated by red rectangle and higher in susceptible reaction are indicated by green rectangles.

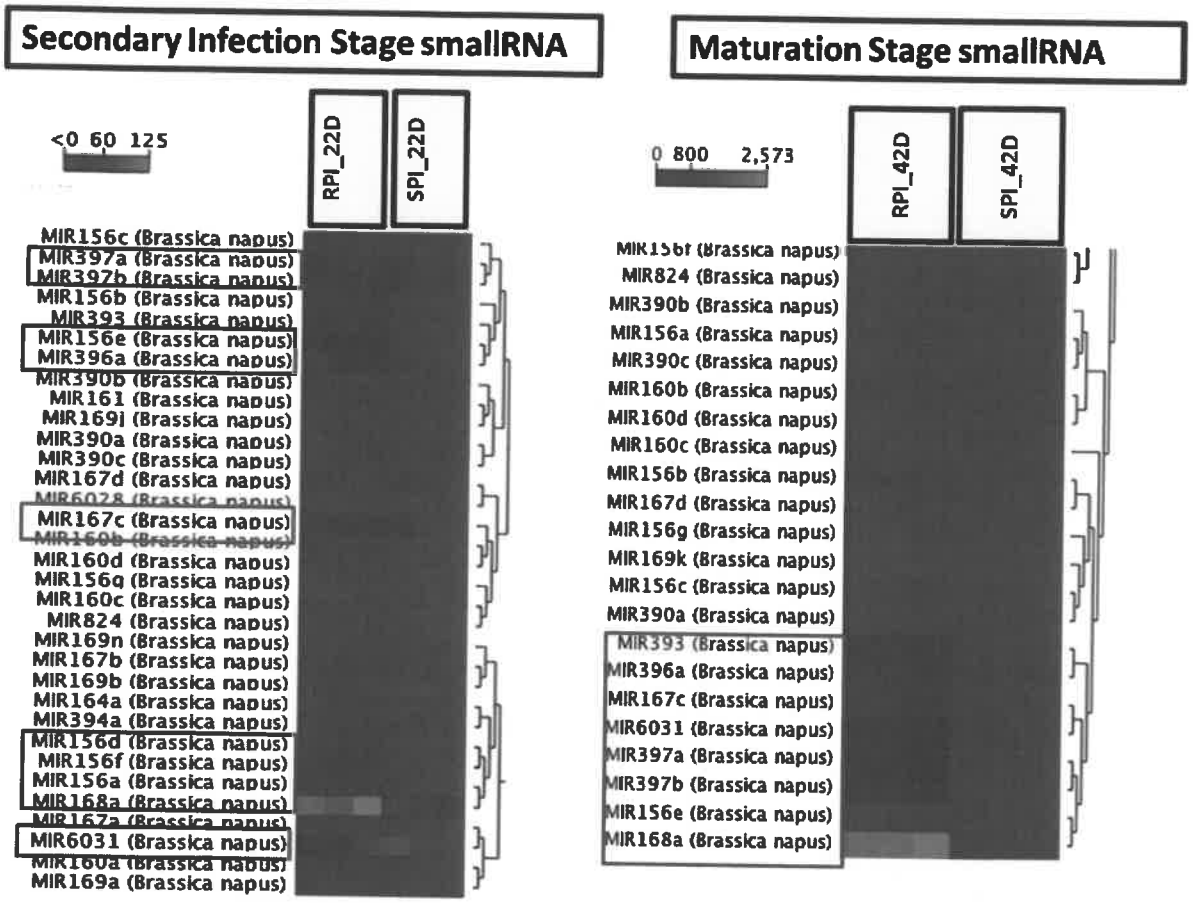


Fig.18.Cluster analysis indicating small RNA consistently expressed at higher level in resistant reaction at secondary and maturation phase of disease development (22 and 42 DPI stage). Pearson's correlation with complete linkage was used for clustering

Feature ID	Small RNA - Length	Small RNA - Name	Small RNA - Resource	Small RNA - Match type
AAAAGGGATAAGTGTCGATTAGATCATGTTCGTCGTTTCAC	41	MIR167c	Brassica napus	Precursor variant
AAGAAGTTCGGAGCGGTTTGAAGC	24	MIR6031	Brassica napus	Mature 5' variant
ATCAACGTTGGCTCAATTATGT	22	MIR397a//MIR397b	Brassica napus	Precursor
CGACAGAAGAGAGTGAGCACAA	22	MIR156a//MIR156d//MIR156f	Brassica napus	Mature 5' super variant
CGACAGAAGAGAGTGAGCACAT	22	MIR156a//MIR156d//MIR156e//MIR156f	Brassica napus	Mature 5' super variant
CGACAGAAGAGAGTGAGCACCC	22	MIR156a//MIR156d//MIR156f	Brassica napus	Mature 5' super variant
CGACAGAAGAGAGTGAGCACCT	22	MIR156e	Brassica napus	Mature 5' super variant
CGACAGAAGAGAGTGAGCACTC	22	MIR156a//MIR156d//MIR156e//MIR156f	Brassica napus	Mature 5' super variant
CGACAGAAGAGAGTGAGCACTT	22	MIR156e	Brassica napus	Mature 5' super variant
TCCACAGCTTTCTTGAACTTTCT	23	MIR396a	Brassica napus	Precursor
TCGCATGGTGCAGGTCGGGAAC	22	MIR168a	Brassica napus	Mature 5' super variant
TCGCATGGTGCAGGTCGGGAAT	22	MIR168a	Brassica napus	Mature 5' super variant
TCGCGTGGTGCAGGTCGGGAAC	22	MIR168a	Brassica napus	Mature 5' super variant
TCGCTTGATGCAGGTCGGGAAC	22	MIR168a	Brassica napus	Mature 5' super variant
TCGCTTGGAGCAGGTCGGGAAC	22	MIR168a	Brassica napus	Mature 5' super variant
TCGCTTGGCGCAGGTCGGGAAC	22	MIR168a	Brassica napus	Mature 5' super variant
TCGCTTGGGGCAGGTCGGGAAC	22	MIR168a	Brassica napus	Mature 5' super variant
TCGCTTGGGGCAGGTCGGGAAT	22	MIR168a	Brassica napus	Mature 5' super variant
TCGCTTGGTGCAGGACGGGAAC	22	MIR168a	Brassica napus	Mature 5' super variant
TCGCTTGGTGCAGGCCGGGAAC	22	MIR168a	Brassica napus	Mature 5' super variant
TCGCTTGGTGCAGGTAGGGAAC	22	MIR168a	Brassica napus	Mature 5' super variant
TCGCTTGGTGCAGGTCGCGAAC	22	MIR168a//MIR168b	Brassica napus	Mature 5' super variant
TCGCTTGGTGCAGGTCGGGAAA	22	MIR168a	Brassica napus	Mature 5' super variant
TCGCTTGGTGCAGGTCGGGAAC	22	MIR168a	Brassica napus	Mature 5' super
TCGCTTGGTGCAGGTCGGGAACC	23	MIR168a	Brassica napus	Mature 5' super variant
TCGCTTGGTGCAGGTCGGGAACT	23	MIR168a	Brassica napus	Mature 5' super
TCGCTTGGTGCAGGTCGGGAACTG	24	MIR168a	Brassica napus	Precursor
TCGCTTGGTGCAGGTCGGGAACTGATTGACTGACACCGCCACG	43	MIR168a	Brassica napus	Precursor variant
TCGCTTGGTGCAGGTCGGGAAG	22	MIR168a	Brassica napus	Mature 5' super variant
TCGCTTGGTGCAGGTCGGGAAT	22	MIR168a	Brassica napus	Mature 5' super variant
TCGCTTGGTGCAGGTCGGGAATG	23	MIR168a	Brassica napus	Mature 5' super variant
TCGCTTGGTGCAGGTCGGGAATT	23	MIR168a	Brassica napus	Mature 5' super variant
TCGCTTGGTGCAGGTCGGGACA	22	MIR168a	Brassica napus	Mature 5' super variant
TCGCTTGGTGCAGGTCGGGACC	22	MIR168a	Brassica napus	Mature 5' super variant
TCGCTTGGTGCAGGTCGGGACT	22	MIR168a	Brassica napus	Mature 5' super variant
TCGCTTGGTGCAGGTCGGGACTT	23	MIR168a	Brassica napus	Mature 5' super variant
TCGCTTGGTGCAGGTCGGGATC	22	MIR168a	Brassica napus	Mature 5' super variant
TCGCTTGGTGCAGGTCGGGATT	22	MIR168a	Brassica napus	Mature 5' super variant
TCGCTTGGTGCAGGTCGGGGAC	22	MIR168a	Brassica napus	Mature 5' super variant
TCGCTTGGTGCAGGTCTGGAAC	22	MIR168a	Brassica napus	Mature 5' super variant
TCGCTTGGTTCAGGTCGGGAAC	22	MIR168a	Brassica napus	Mature 5' super variant
TCGCTTGTTGCAGGTCGGGAAC	22	MIR168a	Brassica napus	Mature 5' super variant
TCCAAAGGGATCGCATTGATTT	22	MIR393	Brassica napus	Mature 5' super variant
TCCAAAGGGATCGCATTGATTTT	23	MIR393	Brassica napus	Mature 5' super variant

Table.1 important small RNA identified in this project differentially expressed in resistant and susceptible reaction.

Identification of P3 isolates specific transcript and their differential expression in resistant and susceptible roots.

Inoculated and un-inoculated transcripts were mapped to *B. napus* v.3 genome (inoculated include canola root + Plasmodiophora transcripts). Unmapped reads from inoculated samples were collapsed against uninoculated sample reads to remove possible non-Plasmodiophora transcripts. We obtained around 41 million reads with this exercise after analyzing and mapping more than 2.5 billion transcripts we generated including all three stages of disease development (10D; 22D and 42 D after inoculation covering primary, secondary and advanced disease development stage.

These 41 million reads were de-novo assembled using CLC Bio to get contigs. We got 21,811 contigs (about 90% - 1Kb; 10% - 2 Kb; criteria was above 500 bp). We also run a small subset (10,000 reads) BLAST against available Plasmodiophora sequences at NCBI and around 8% shows hit. Another way we checked is ran a blast against de-novo assembled contigs which shows around 10% hits against similar NCBI sequences for Plasmodiophora. It should also be considered that these are really small number of sequences available at NCBI (January, 2015) possibly covering really small amount of genome. Schematic of this approach is given in Fig. 19.

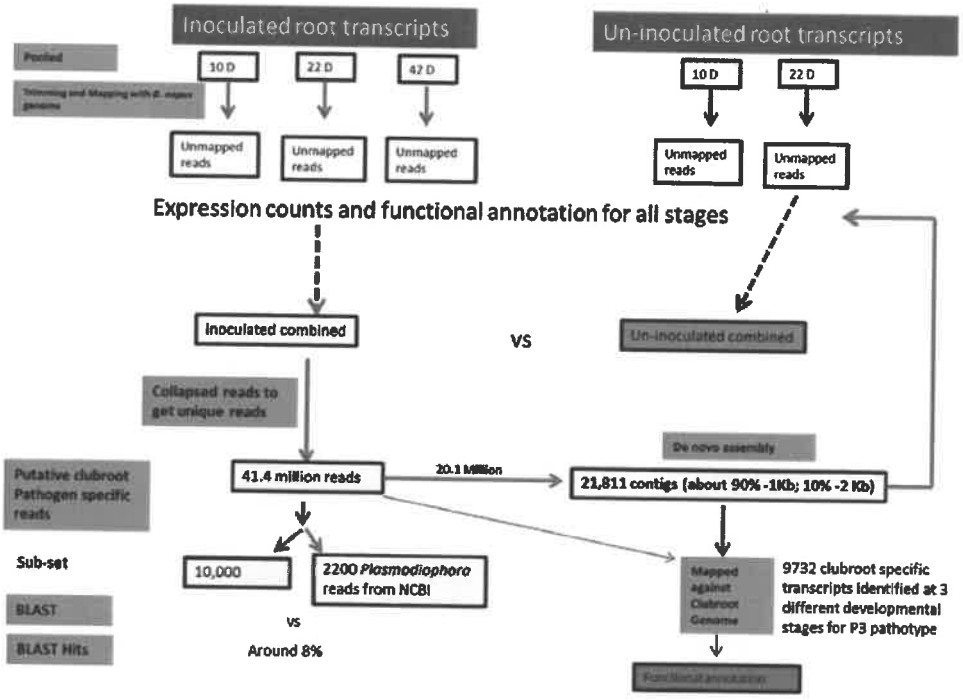


Fig.19.Schematics of bioinformatics approach to identify P3 pathotype transcripts from RNAseq data generated to study the pant genes differentially regulated under 3 clubroot disease development stages (transcriptome data earlier reported in 2014)

These stage specific reads were Tophat aligned. Out of total assembled transcripts, 38% (8353 out of 21811) had a blastn (e<1e-50) to clubroot genome assembly. 25% (5831) were really similar to the assembly (identity >99%). In

total, 10831156 (25%) unique reads mapped and 32425787 did not map. Expanding the collapsed reads, 31292652 (25%) reads mapped. Abundance of bacterial transcripts was the reason for lower transcript proportions or possibly difference. Differentially expressed genes were identified using DESeq2. Recently published Plasmodiophora genome (Schwelm et al., 2015) was used for this alignment in collaboration with Dr. Arne Schwelm (University of Uppsala, Sweden). In all, 9732 transcripts of Plasmodiophora brassicae were used to obtain FPKM values to indicate transcript abundance and putative functional annotations with BLAST hits. These datasets were further used to hierarchical clustering to obtain enriched dataset representing Pathotype 3 (Canadian) transcripts expressed at higher level in susceptible reaction and expressed higher level in resistant reaction (Fig.20). blue color indicate transcripts expressed at higher level in susceptible reaction). These includes genes like Actin1, HSP, ATPase AAA, Alcohol dehydrogenase, SWIM Zinc finger protein expressed at higher level in susceptible reaction at secondary and maturation phase of disease development. On the other hand, Calcium dependent Kinases, Chromosome condensation regulator protein, CathepsinL3, Lipase, Ankyrin2 were expressed at higher level in resistant reaction. This indicates differential regulation of transcripts in not only plant but also in pathotype depending on disease response.

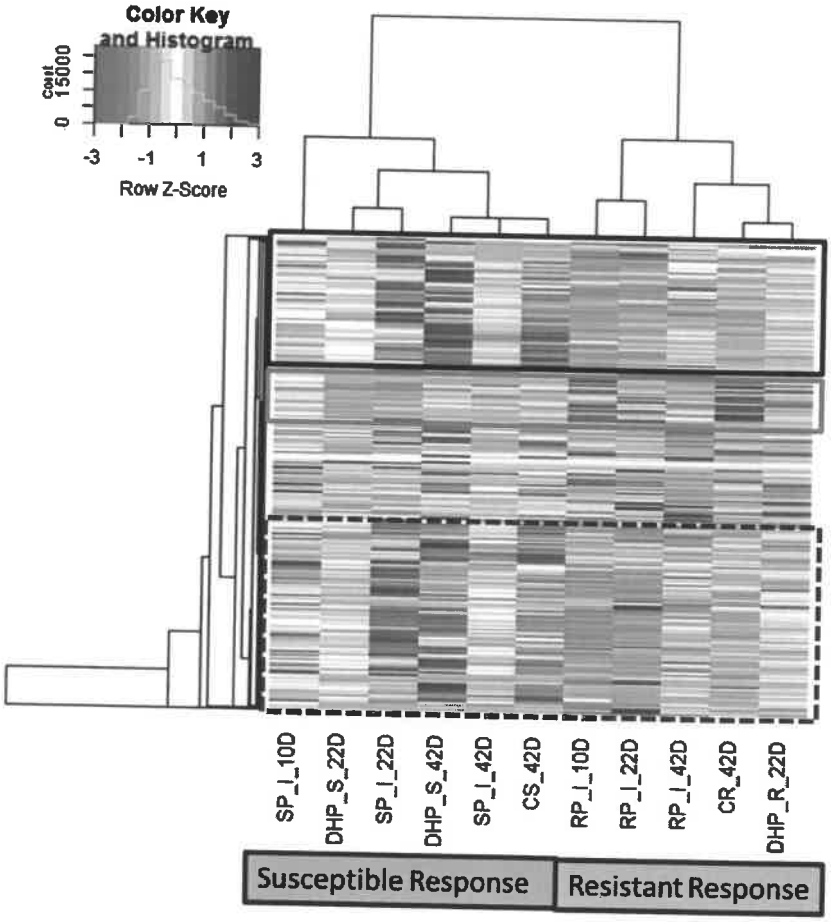


Fig.20. Hierarchical clustering of Plasmodiophora brassicae W. Pathotype 3 transcripts differentially expressed under susceptible and resistant response. Datasets represents total 9732 transcripts used for cluster analysis at 10,

22 and 42 DAI stage obtained from same RNAseq dataset which were used to understand plant transcript differential response.

8. Discussion: *Provide discussion necessary to the full understanding of the results. Where applicable, results should be discussed in the context of existing knowledge and relevant literature. Detail any major concerns or project setbacks.*

The clubroot pathogen *P. brassicae*'s pathotype P3 is very virulent and a major contributor to clubroot disease in Alberta and Saskatchewan[Rahman et al., 2013]. Single gene resistance does not appear to be promising to control clubroot since major QTL loci can only afford up to 40% resistance (Jubault et al., 2008, Chen et al., 2013), necessitating in addition complementary strategies. We took a multidisciplinary approach to understand the clubroot resistance response in spring canola (*B. napus*) to Pathotype 3 . We identified the genomic regions associated with resistance; we conducted genetic linkage mapping, transcriptome analysis and small RNA analysis. Genetic linkage mapping provided a tightly linked marker for resistance (sv Mendel-derived) which can be further used for Marker Assisted Selection (MAS) in future. This locus can be stacked with other known clubroot resistance loci on other chromosomes (like A1, A8) to breed stronger resistance. It should be noted that this loci covers up to around 50% genetic variation for clubroot resistance. This indicates the need to find other resistance factors that cannot be identified by genetic linkage mapping alone. Transcriptome analysis provided a window to identify these potential resistance pathways/factors. We developed a comprehensive atlas of susceptibility and resistance reactions at important disease developmental stages (primary, secondary and advanced disease). This analysis identified genes that are specifically expressed at higher levels in resistance reaction at the secondary and maturation phases. These genes include lignin and phenylpropanoid, SA mediated resistance response genes. The candidates can be further investigated to understand their role in imparting a stronger resistant phenotype. Earlier research has indicated role of lignin biosynthesis and phenylpropanoid genes in disease resistance (Miedes et al., 2014). Overall, the disease resistance cascades are in accordance with transcriptome analysis for clubroot disease earlier done in *Brassica rapa* (Chu et al., 2014). We also performed small RNA sequencing and analysis to identify small RNA categories expressed at a higher level in resistance reaction. We mapped these to the *B. napus* genome and identified differential response in the case of resistance and susceptible reactions. Future research needs to be done to identify how clubroot resistance genomic regions, transcriptome and small RNA interplay clubroot resistance and susceptibility response.

9. Conclusions and Recommendations: *Highlight significant conclusions based on the previous sections, with emphasis on the project objectives specified above. Provide recommendations for the application and adoption of the project.*

A major locus for clubroot resistance in spring canola (*B. napus*) has been mapped and a marker has been identified. This locus, derived from a European winter rapeseed, alone can contribute to about 50% of the resistant phenotype, suggesting that additional unmapped regions are also responsible for the resistance phenotype of the European parent and also the spring canola derivative.

The locus is molecularly complex. The multiplicity of resistance gene-like sequences need to be investigated

further to determine if a single protein-coding gene is responsible or if interactions with other gene(s) in the proximity are required.

Transcriptome analysis at the primary, secondary and maturation phase of clubroot disease reaction identified a number of pathways/genes that associated with susceptibility and resistance. These provide a valuable starting point to dissect the ancillary resistance pathways. These need to investigated further to present a prioritized list of genes to select for, for transferring resistance to target varieties.

The small RNA data generated here provides another window to clubroot disease susceptibility and resistance. The analyses can be used in assembling additional arsenal for combatting the disease by genetics.

10. Success stories/ practical implications for producers or industry: *Identify new innovations and /or technologies developed through this project; and elaborate on how they might impact the producers /industry.*

- A tightly linked marker for clubroot resistance on N3 chromosome mapped here can be used for Marker Assisted Selection of clubroot resistance
- A comprehensive transcriptional atlas of resistance and susceptibility reactions generated
- Disease resistance genes identified here can be validated and stacked for durable clubroot resistance

11. Patents/ IP generated/ commercialized products: *List any products developed from this research.*

- Internal assessment in progress

12. List technology transfer activities: *Include presentations to conferences, producer groups or articles published in science journals or other magazines.*

Publications from/pertaining to this project is as below:

1. Sharma et al., (2013) Reaction of Lines of the Rapid Cycling Brassica Collection and Arabidopsis thaliana to Four Pathotypes of Plasmodiophora brassicae. Plant Disease 97:6, 720-727
2. Rahman H., Peng G., Yu F., Falk K.C., Kulkarni M., & Selvaraj G. (2013) Genetics and breeding for clubroot resistance in Canadian spring canola (*B. napus* L.). Canadian Journal of Plant Pathology special issue for clubroot resistance.36 (s1):122-134 DOI:10.1080/07060661.2013.862571
3. **Oral Presentation by Gopalan Selvaraj**: Kulkarni, M., Ashe, P., Fredua-Ageyman, R., Akhov, L., Rahman, H., G., Selvaraj (2014).Transcriptome analyses of Brassica napus roots after infection with Plasmodiophora brassicae Woronin indicate differential dynamics of gene expression in resistant and susceptible lines. 11th Conference of the European Foundation for Plant Pathology, 8-13 Sept, 2014 (Krakow, Poland)
4. **Oral Presentation by Dr. Manoj Kulkarni:** Kulkarni, M., Ashe, P., Fredua-Ageyman, R., Akhov, L., Rahman, H., G., Selvaraj (2014). Insights into susceptibility and resistance to clubroot pathogen in Brassica napus. 11th Applied Genomics Workshop, Canola Industry Meeting, 3-4 December, 2014, Saskatoon.
5. **Poster Presentation by Dr. Manoj Kulkarni:**
Kulkarni, M., Ashe, P., Fredua-Ageyman, R., Akhov, L., Rahman, H., G., Selvaraj (2015) Identification of genome-wide pathways for enhancing clubroot (Plasmodiophora brassicae W.) resistance in Brassica napus L. (Abstract ID. 11019). 14th International Rapeseed Congress, 2015 Saskatoon (July 5-9, 2015)

13. List any industry contributions or support received.

This project was co-funded by Western Grains Research Foundation, SaskCanola and ADF program of the Saskatchewan Ministry of Agriculture

14. Is there a need to conduct follow up research? *Detail any further research, development and/or communication needs arising from this project.*

Further research would be required to clone the CR gene from the mapped region; validation of gene leads from the transcriptomics and small RNA work reported will also be useful in arriving at a suite of genes for stacking for durable clubroot resistance.

15. Acknowledgements. *Include actions taken to acknowledge support by the Ministry of Agriculture and the Canada-Saskatchewan Growing Forward 2 bilateral agreement.*

We thank Western Grains Research Foundation (WGRF), SaskCanola, and Agriculture Development Fund of the Saskatchewan Ministry of Agriculure for funding support. We are grateful to Dr. H. Rahman and his post-doctoral associate Dr R. Fredua-Ageyman at the University of Alberta for access to the genetic material

and for disease scoring. We are grateful to Dr. Paula Ashe for bioinformatics in the transcriptome analysis work.

16. Appendices: *Include any additional materials supporting the previous sections, e.g. detailed data tables, maps, graphs, specifications, literature cited*

Chen et al., (2013) Identification of Novel QTLs for Isolate-Specific Partial Resistance to Plasmodiophora brassicae in Brassica rapa. PLoS ONE 8(12): e85307. doi:10.1371/journal.pone.0085307

Chu et al., (2014) Fine mapping of Rcr1 and analyses of its effect on transcriptome patterns during infection by Plasmodiophora brassicae. BMC Genomics 2014, 15:1166 doi:10.1186/1471-2164-15-1166

Chu et al., (2013) Identification of the clubroot resistance gene Rpb1 and Introgression of the resistance gene into canola breeding lines using a marker-assisted selection approach. Acta Hortic. 2013;1005:599–605.

Hammond-Kosack and Jones (1996). Resistance gene-dependent plant defense responses. Plant Cell 8: 1773–1791.

Hatakeyama., et al. (2013) Identification and Characterization of Crr1a, a Gene for Resistance to Clubroot Disease (Plasmodiophora brassicae Woronin) in Brassica rapa L. PLoS ONE 8(1): e54745. doi:10.1371/journal.pone.0054745

Jubault et al (2008). Identification of quantitative trait loci controlling partial clubroot resistance in new mapping populations of Arabidopsis thaliana. Theor Appl Genet. 2008; 117(2):191–202. doi: 10.1007/s00122-008-0765-8.

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

Miedes et. al. 2014. The role of the secondary cell wall in plant resistance to pathogens. Front. Plant Sci., doi: 10.3389/fpls.2014.00358

Rahman H., Peng G., Yu F., Falk K.C., Kulkarni M., & Selvaraj G. (2013) Genetics and breeding for clubroot resistance in Canadian spring canola (Brassica napus L.). Canadian Journal of Plant Pathology special issue for clubroot resistance.36 (s1):122-134 DOI:10.1080/07060661.2013.862571

Schwelm et al. (2015) The Plasmodiophora brassicae genome reveals insights in its life cycle and ancestry of chitin synthases. Scientific Reports 5, Article number: 11153. Doi: 10.1038/srep11153

Sharma et al., (2013) Reaction of Lines of the Rapid Cycling Brassica Collection and Arabidopsis thaliana to Four Pathotypes of Plasmodiophora brassicae. Plant Disease 2013 97:6, 720-727

Thimm et. al. 2004. MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. Plant J. 37: 914–939.

Ueno et al., (2012) Molecular characterization of the CRa gene conferring clubroot resistance in Brassica rapa. Plant Mol Biol (2012) 80:621–629 DOI 10.1007/s11103-012-9971-5.

Usadel et. al. 2005. Extension of the visualization tool MapMan to allow statistical analysis of arrays, display of corresponding genes, and comparison with known responses. Plant Physiol. 138: 1195–1204.