UNIVERSITAT POLITÈCNICA DE VALÈNCIA

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Institute for the Preservation and Improvement of Valencian Agro-diversity

Resistance to virus yellows in sugar beet

Master's Thesis

European Master Degree in Plant Breeding

AUTHOR: Joshi , Pratikshya

Tutor: Fita Fernández, Ana María

External cotutor: KVARNHEDEN, ANDERS

Experimental director: PUTHANVEED, VINITHA

ACADEMIC YEAR: 2021/2022

RESISTANCE TO VIRUS YELLOWS IN SUGAR BEET

Pratikshya Joshi

Master's Thesis

Erasmus Mundus Master's in Plant Breeding (emPLANT)

Supervisor: Prof. Anders Kvarnheden

Co-supervisor: Vinitha Puthanveed

Tutor at UPV: Prof. Ana Fita Fernández

Swedish University of Agricultural Sciences (SLU)

Uppsala, Sweden

September 12, 2022

Abstract

Sugar beet (*Beta vulgaris*), a biennial crop from the Amaranthaceae family, is cultivated for sucrose. Virus yellows is an economic disease affecting sugar beet. It is caused by a complex of viruses, namely, beet yellows virus (BYV), beet mild yellowing virus (BMYV), beet chlorosis virus (BChV), beet mosaic virus (BtMV), beet western yellows virus (BWYV) and beet leaf yellowing virus (BLYV). Neonicotinoids had been an effective control measure preventing aphid-mediated transmission of virus yellows until their near-total ban by the EU in 2018. Sugar beet production stands vulnerable against virus yellows as no resistant cultivars have been released yet, and the resistance mechanism of plants against virus yellows is not fully known. This project aims to understand the resistance mechanism of sugar beet and eventually develop resistant cultivars using the identified wild resistant source. For this purpose, an experiment involving aphid-mediated inoculation of BMYV and BYV, containing three treatments (virulent aphids, healthy aphids, and insecticide spray), was carried out for susceptible elite and resistant wild genotypes. The total RNA (containing RNA from both the plant and the virus) extracted from these samples went through RNA sequencing and analysis to determine the differentially expressed genes across different genotypes, treatments and time points. cDNA was synthesised, and RT-qPCR for absolute quantification was performed to determine the virus titre in BMYV inoculated resistant and susceptible plants. Furthermore, RT-qPCR for relative quantification was performed to determine the differential expression of the pathogenesis-related genes *AGO1*, *PR1a,* and *PR5* in the virus inoculated and non-inoculated healthy plants at different time points. The virus titre was higher in susceptible plants as compared to resistant plants. The genes *PR1a* and *PR5* were not expressed in the cDNA mix of samples from resistant and susceptible genotypes. Meanwhile, in susceptible plants, *AGO1* showed increased expression in healthy plants compared to inoculated plants. In resistant plants, it was probably expressed but probably could not be detected using primers designed for cultivated sugar beet. For studying gene expression in the resistant genotype, it will be necessary to design new primers, which could be done using obtained sequence data for the resistant genotype.

Keywords- *sugar beet, virus yellows, BMYV, BYV, absolute quantification, RT-qPCR, RNA sequencing, relative quantification*

Acknowledgements

I started this thesis work with the motivation to bridge the gap in my knowledge and get proficient in molecular biology. I express my gratitude to the people who helped make it possible. First, I would like to express my sincere appreciation to Dr Anders Kvarnheden for his excellent supervision and feedback. I will remember most how he reminded me that the point is always to learn, regardless of the results. I am equally thankful to Vinitha Puthanveed for being a great mentor and supervising and guiding me on this thesis project. I have learned a lot from her during the thesis period. My sincere gratitude to Dr Ana Fita for her valuable guidance and continuous motivation throughout the academic period. I am very grateful to Dr Alicia Ayerdi-Gotor for her constant support and helpful advice in every stage of this master's programme. My heartfelt appreciation also goes to SLU Bioinformatic Infrastructure (SLUBI) and Dr Abu Bakar Siddique for the training in RNA sequencing analysis. I am thankful to the Department of Plant Biology for the opportunity to carry out my thesis here. This project would not have been possible without the funding from SLU Grogrund and DLF Beet Seed and the plant samples received from DLF Beet Seed. My special thanks go to the lab members in the department for your helpful advice on the experiments and the interesting conversations at lunchtime and meetings. I am grateful for my warm welcome in Uppsala and the wonderful experiences during my stay. Lastly, I want to acknowledge my colleagues in the emPLANT programme for their helpful insights and beautiful memories and my family for their unwavering support.

Table of Contents

List of Figures v
List of Tables vi
List of Abbreviations vii
1 Introduction 8
1.1 Sugar beet 8
1.2 Diseases in sugar beet 8
1.3 Polerovirus 9
1.4 Closterovirus 10
1.5 Control Measures 11
1.6 Resistance Mechanism 11
1.7 RNA sequencing 12
1.8 Aim of the project 12
2 Materials and methods 13
2.1 Plant materials 13
2.2 Inoculation using aphids 13
2.3 RNA extraction 14
2.4 RNA sequencing 14
2.5 cDNA synthesis 14
2.6 RT-qPCR for absolute quantification of viral RNA 14
2.7 RT-qPCR for relative quantification 17
2.8 Bioinformatic analysis 18
2.9 Statistical analysis 19
3. Results 20
3.1 Absolute quantification 20
3.2 Relative quantification 22
3.3 RNA sequencing analysis 24
4 Discussion 26
5 Conclusion 28
6 References 29
7 Appendix 34
Script for differential gene expression analysis on RStudio 35
Supplementary Tables 1-4 43

List of Figures

Figure 1: Genomic organisation schema of poleroviruses, with ORFs 9
Figure 2: Schematic representation of the genome organisation of closteroviruses..... 10
Figure 3: Schema of the qPCR run cycles, including amplification cycles and melt curve analysis 16
Figure 4: Pipeline for RNA-seq analysis 19
Figure 5: Melt curves generated for qPCR runs with primer pair I and primer pair II..... 21
Figure 6: Bar graph showing the copy number of BMYV in 1 µg RNA in older infected leaves from resistant and susceptible plants at different time points..... 21
Figure 7: Bar graph with BMYV copy numbers of BMYV in 1 µg RNA in infected young leaves from resistant and susceptible plants at different time points..... 22
Figure 8: Bar graph with *AGO1* gene expression ratio of BMYV inoculated and healthy leaves in susceptible genotype at 14 and 21 dpi 23
Figure 9: Principal component analysis for gene expression in susceptible and resistant plants inoculated with BMYV or BYV 24

List of Tables

Table 1: List of samples for absolute quantification of BMYV…….15

Table 2: Sequences of primers for target genes and reference genes with their forward and reverse sequences……...17

Table 3: Coefficient of determination (R^2), efficiency, and slope for different qPCR runs for absolute quantification of BMYV…….20

Table 4: Primer efficiencies of housekeeping and target genes tested with qPCR…….23

List of Abbreviations

cDNA- Complementary deoxyribonucleic acid

RNA- Ribonucleic acid

RT- qPCR- Reverse transcription quantitative polymerase chain reaction

Dpi- Days post-inoculation

1 Introduction

1.1 Sugar beet

Sugar beet (*Beta vulgaris* spp. vulgaris, L.) is the second largest source of sugar after sugar cane (OECD, 2020). It is a biennial crop from the Amaranthaceae family and among the four cultivated beets in the subspecies vulgaris: red beet (beetroot), fodder beet, leafy beet (chard), and sugar beet. With an estimated genome size of 714-758 megabases, it is a diploid species with nine pairs of chromosomes, while polyploid varieties also exist (Zicari et al., 2019). It has a wide cultivation range, and its leading producers are Russia, France, Germany, the United States and Turkey (FAOSTAT, 2020). Sugar extracted from its hypocotyl as sucrose, forming 15-20% of its root mass, accounts for about 28% of the world's sugar production (Zicari et al., 2019). Besides sugar production, other uses of sugar beet include sugar beet pulp, molasses, bioplastics, and biofuel (Mukharjee & Gantait, 2022). An important crop in Swedish agriculture, sugar beet production in Sweden accounts for 1.07% of arable land and is grown in southern parts of Sweden (Ruus, 2020, Jordbruksverket, 2019).

1.2 Diseases in sugar beet

Fungal and viral diseases are common in sugar beets, which express symptoms in the leaves or roots, resulting in reduced sugar yields. Major fungal diseases of sugar beet include Cercospora leaf spot caused by *Cercospora beticola*, downy mildew caused by *Peronospora farinosa* and powdery mildew caused by *Erysiphe betae*. Major viral diseases in sugar beet include rhizomania caused by beet necrotic yellow vein virus (BNYVV) and virus yellows (Bayer Crop Science UK, 2018).

Virus yellows is a disease complex caused by viruses from three virus families. Many of these viruses can infect the plant at the same time, increasing the disease severity (Wintermantel, 2005). Beet yellows virus (BYV) is a member of the family *Closteroviridae*, and beet mosaic virus (BtMV) belongs to the *Potyviridae* family. In a survey in Europe by Hossain et al. (2021), BYV, BMYV and BChV were commonly detected, but not BtMV. The four viruses causing virus yellows in sugar beet belonging to the family *Solemoviridae* are beet mild yellowing virus (BMYV), beet western yellows virus (BWYV), beet chlorosis virus (BChV) and beet leaf yellowing virus (BLYV) (Yoshida & Tamada, 2019). It is a vector-transmitted disease, the major vector being the green peach potato aphid (*Myzus persicae*). The visible symptoms on leaves are

necrosis, yellowing and eventual thickening, brittleness, and reduced growth (Nilsson & Larsson 1990). The necrotic leaf spots cause reduced photosynthesis, leading to reduced sucrose storage in the hypocotyl. This reduced photosynthesis, as well as the increase in juice impurities, results in reduced sugar yields (Bennet, 1960; Stevens et al., 2004).

1.3 Polerovirus

The genus *Polerovirus*, consisting of viruses with a genome of positive sense, single-stranded RNA (+ssRNA), was recently moved from the family *Luteoviridae* to *Solemoviridae* (Walker et al., 2021). BMYV, BChV and BWYV are classified as poleroviruses, while BLYV is considered an unclassified polerovirus (LaTourrette et al., 2021). Poleroviruses have a non-enveloped, spherical virion with icosahedral symmetry and a genome size of 5.3 to 5.7 kb, comprising at least seven open reading frames (ORFs), as illustrated in figure 1. The genome of poleroviruses lacks a poly-A tail at the 3' end, and their 5' end contains a viral genome-linked protein (VPg) (Wetzel et al., 2018). The three ORFs towards the 5' end (ORF0, -1 and -2) are translated directly from the genomic RNA, while the remaining downstream ORFs (ORF3a, -3, -4 and -5) are translated from the subgenomic RNAs (sgRNAs) (Stevens et al., 2005a).

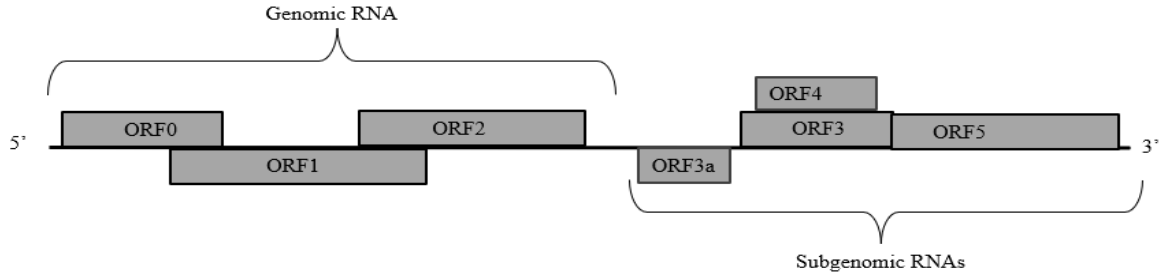


Figure 1: Genomic organisation schema of poleroviruses, with ORFs. Adapted from Delfosse et al. (2021).

ORF0 codes for the P0 protein, which functions as a viral RNA silencing suppressor to counter post-transcriptional gene silencing (PTGS), an RNA-silencing mechanism mediated by ARGONAUTE 1 (AGO1) as a component of the RNA-induced silencing complex (RISC) in plants (Bortolamiol et al., 2007). Moreover, ORF1 codes for the P1 protein related to viral replication and VPg (Delfosee et al., 2021). P2, encoded by ORF2 and containing the RNA-dependent polymerase motif (RdRp), forms a fusion with P1. Furthermore, ORF3a codes for protein P3a, which is involved in the long-distance movement of the virus, while the P3 protein encoded by

ORF3 is the major coat protein (capsid). The P3P5 fusion protein is formed through the translation of ORF3 and ORF5 together, and its functions are related to aphid transmission and the systemic spread of the virus. Finally, ORF4 encodes protein P4, which plays a role in host-specific viral movement (Wetzel et al., 2018).

The transmission of poleroviruses is persistent, circulative, and non-propagative. Once ingested from the phloem sap by an aphid, poleroviruses move through its guts, tissues, hemocoel and finally reach the salivary glands. When the aphid feeds on a new host plant, the viruses can be injected from the salivary glands into the phloem sap, thus following a circulative transmission pattern (Gray & Gildow, 2003; Brault et al., 2007). The transmission is also persistent as the virus is retained by the aphid and remains infective for long periods. However, the virus is non-propagative, meaning that the virus cannot replicate inside the aphid. Moreover, in the host plant, the poleroviruses are restricted to the phloem (Hipper et al., 2013). The most efficient vector for polerovirus transmission in sugar beets is *Myzus persicae*, while *Macrosiphum euphorbiae* is also known to transmit them (Kozlowska-Makulska et al., 2009).

1.4 Closterovirus

Closteroviruses, belonging to the family *Closteroviridae*, have a genome of +ssRNA with a size of 15.9 to 19.3 kilobases (Igori et al., 2009). Besides *M. persicae*, the black bean aphid (*Aphis fabae*) is also known to transmit virus yellows causing closteroviruses in sugar beets. The transmission of closteroviruses by aphids is semi-persistent, meaning that the virus particles are stored in the mouth parts of aphids for a shorter time (Carr et al., 2020). The 3' end of their genome lacks a polyA tail, and their virus particles contain two coat proteins (one minor coat protein) (Igori et al., 2009). The organisation of the genome is illustrated in figure 2.

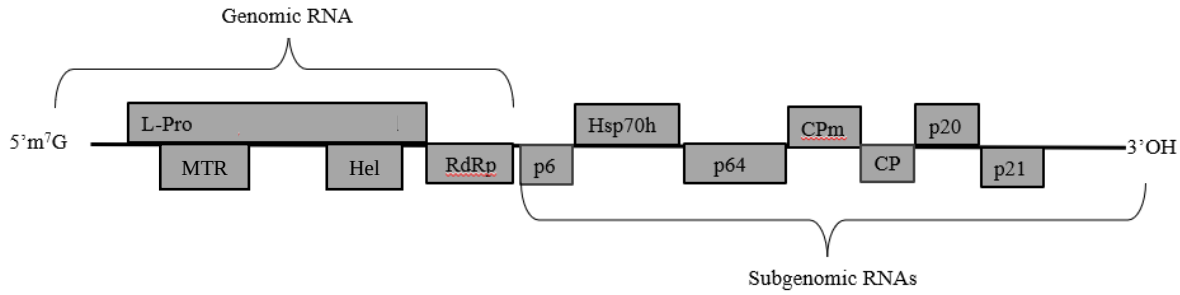


Figure 2: Schematic representation of the genome organisation of closteroviruses. L-PRO, papain-like

protease; MTR, methyltransferase; HEL, helicase; RdRp, RNA-dependent RNA-polymerase; HSP70h, heat shock protein 70 homologue; CPm, minor coat protein; CP, coat protein; p20, 20-kDa protein; p21, 21-kDa protein.

1.5 Control Measures

Several methods have been used for the control of virus yellows. Cultural practices include minimising the time sugar beets are placed in a clamp and early sowing (Ruus, 2020). The most efficient control method previously in practice is the use of insecticides of the class neonicotinoids. Neonicotinoids are nicotine-like substances that target the nervous system of the aphids feeding on sugar beet leaves (Bass & Field, 2018). They are used as seed coatings, taken up by the plant, and transported to all plant tissues. The problem lies in the transport of neonicotinoids to pollen and nectar, which poses risks to pollinators like bees, causing the European Union (EU) to establish a near-total ban on neonicotinoids (European Commission, 2018). Other insecticides still in use are some organophosphates and carbamates (Hauer et al., 2017).

1.6 Resistance Mechanism

In response to viral attacks, plants have developed effective defence mechanisms of resistance by limiting or preventing the damage caused by virus infection. One of the major defence mechanisms is through resistance (R) genes which are specific to the pathogen. At the local infection site, this begins with the hypersensitive response (HR) followed by programmed cell death (PCD) (Soosar et al., 2005). In tissues that are far from the region of infection, systemic acquired resistance (SAR) comes into action and activates defence signalling pathways: salicylic acid (SA) pathway and jasmonic acid (JA) pathway. SA pathway is activated upon biotrophic infection, while the JA pathway is associated with defence against necrotrophic pathogens and insects. These pathways induce increased expression of pathogenesis-related proteins. For example, upon viral infection, the accumulation of pathogenesis-related protein 1a (PR1a) and pathogenesis-related protein 5 (PR5) is induced by the SA pathway in non-infected regions to block further virus propagation (Ali et al., 2018).

Furthermore, another major defence mechanism is RNA-mediated resistance or PTGS. It is an RNA-silencing mechanism involved in the cleavage of double-stranded RNA (dsRNA) into small interfering RNAs (siRNAs) by the RNAse Dicer. These siRNAs bind to RISC, which degrades the viral RNA (Lin et al., 2007). Viruses have developed an RNA-silencing suppression

method to counter this. Bortolamiol et al. (2007) showed that the P0 protein encoded by the viral ORF0 degrades the AGO1 protein, a key component of the RISC complex.

Moreover, tolerance may be present in plants, where the symptoms are either absent or mild in response to an infection but have a similar virus titre to susceptible plants (Bruening, 2006; Palukaitis & Carr, 2008)

1.7 RNA sequencing

RNA sequencing is a high throughput sequencing technology that uses next-generation sequencing (NGS) and elucidates the transcriptional structure of the genes. Comparison of RNA-seq data of the susceptible and resistant sugar beet genotypes followed by differential gene expression analysis can show the differentially expressed genes involved in defence pathways (Liu et al., 2013).

1.8 Aim of the project

Without protection from this insecticide, the sugar beet crop is vulnerable to virus yellows as no resistant cultivars have yet been released. Moreover, the resistance mechanism is not fully known as virus yellows did not receive much attention while neonicotinoids remained a viable and efficient option for control. This ban has severely reduced yields in Belgium, France, Germany, the Netherlands and the UK. In Sweden, the colder winters do not allow the overwintering of adult aphids, and they migrate only late in the season to Sweden when the plants have matured, and the impact of the viruses is less severe. However, that is predicted to change with milder winters due to climate change favouring aphids overwintering and spreading early in the season in Swedish fields. There is a dire need for a more sustainable control method against virus yellows.

The aim of this project, which is carried out in collaboration between the Swedish University of Agriculture (SLU) and the breeding company DLF Beet Seed (previously MariboHilleshög), is to understand the resistance mechanism of sugar beet against virus yellows and eventually develop resistant cultivars using a resistant source identified by DLF Beet Seed. The thesis project involves RT-qPCR for absolute quantification to measure the BMYV concentration, RNA sequencing analysis and differential expression analysis, and RT-qPCR for relative quantification to assess the differential expression of some defence genes.

2 Materials and methods

As this thesis was a part of a bigger project, some experiments: experimental setup of sugar beet plants, RNA extraction and RNA sequencing, were performed before the start of the thesis.

2.1 Plant materials

The sugar beet plant materials, one resistant genotype and one susceptible genotype, used for the experimental setup were prepared at the DLF Beet Seed site in Landskrona, Sweden. The seeds were sown in the greenhouse at different time intervals, and the plants with the first true leaf pair were selected for the experiment.

2.2 Inoculation using aphids

The sugar beet plants, resistant and susceptible, with emerged first true leaf pair were separated into three treatment groups: aphids with virus, aphids without virus and insecticidal spray. For the virus inoculation treatment, aphids were fed on virus-infected leaves, and 10 aphids were placed on each plant. The same process was followed for BYV (a closterovirus) and BMYV (a polerovirus). As a control, plants were exposed to aphids, which had fed on healthy leaves instead of virus-infected sugar beet leaves. As an additional control, plants were sprayed with insecticide four days after the start of the experiment. Before starting the experiment, six samples per genotype were taken. These are considered time point 0 samples and were used as common control samples for BYV and BMYV. For each treatment and genotype, six samples were taken at eight time points: 0, 1, 2, 4, 7, 14, 21 and 28 days post-inoculation (dpi).

Until 4 dpi, samples were only taken from plants exposed to viruliferous or non-viruliferous aphids, but after this, samples were also collected from plants sprayed with insecticide. For the BYV experiment, the virus could not be detected at earlier time points, so only samples from 14, 21, and 28 dpi were included for the RNA sequencing analysis. In addition, samples of the upper leaf pair directly above the first leaf pair were also collected at 14, 21, and 28 dpi for both viruses. At 28 dpi, the inoculated leaves had wilted and died, so the inoculated leaves or older leaves were collected up to 21 dpi. The six samples were taken from six plants and considered biological replicates. The plants were covered with fleece to avoid cross-contamination through the movement of aphids across different treatment areas.

2.3 RNA extraction

The collected leaf samples were received at the SLU, Alnarp Campus, where total RNA (containing both plant and viral RNA) was extracted. The extraction was carried out using RNAqueous™-4PCR Total RNA Isolation Kit (ThermoFisher Scientific) following the kit protocol. The RNA was then quantified using BioAnalyser and sent to SLU, Uppsala Campus.

2.4 RNA sequencing

Three samples were selected for the RNA sequencing out of the six samples collected per treatment per genotype. For BMYV, five time points (1, 4, 14, 21 and 28 dpi) were chosen, while for BYV, three time points (14, 21, and 28 dpi) were selected. The control samples collected at time point 0 were used as common for both BYV and BMYV. cDNA was synthesised from the extracted total RNA, whose quality and RIN number were checked. The samples were sent to SciLifeLab (Science for Life Laboratory) at Uppsala University, where Illumina paired-end sequencing was used. The sequencing results were analysed with the help of SLU Bioinformatics Infrastructure (SLUBI).

2.5 cDNA synthesis

The thesis project started with cDNA synthesis from RNA extracted in section 2.3. cDNA synthesis was performed using Maxima First Strand cDNA Synthesis Kit for RT-qPCR according to the kit protocol in a BioRad T100 Thermo Cycler. For all samples, 1 µg of RNA was used for cDNA synthesis, except for two samples with low RNA concentration, where 500 ng of RNA was taken. Few randomly selected samples were used for reverse transcriptase minus (RT-) negative controls, and a single no template control (NTC) was taken for each round of cDNA synthesis. The thermocycler was set for 10 minutes at 25°C, 15 minutes at 50°C and five minutes at 85°C and was held at 4°C after completion. The cDNA was diluted 10-fold and stored at -20 °C.

2.6 RT-qPCR for absolute quantification of viral RNA

Absolute quantification of viral RNA with RT-qPCR was performed using Maxima SYBR Green/ROX qPCR Master Mix (2X) according to the manufacturer's protocol using samples shown in Table 1. The experiment was set up in a 96-well plate (Bio-Rad Hard-Shell) with three technical replicates of each sample and water as a negative control. A standard curve was produced using five serial dilutions of a plasmid containing the coat protein gene of BMYV, already created in the research group, in the order of: 100 pg/µL, 10 pg/µL, 1 pg/µL, 0.1 pg/µL and 0.01 pg/µL.

Table 1: List of samples for absolute quantification of BMYV.

Time points	Genotypes	Treatments	Name of samples
0 dpi (old leaves)	Resistant	Virus with aphids	1, 2, 3
	Susceptible	Virus with aphids	4, 5, 6
1 dpi (old leaves)	Resistant	Aphids with virus	7, 8, 9
		Aphids without virus	10, 11, 12
	Susceptible	Aphids with virus	13, 14, 15
		Aphids without virus	16, 17, 18
4 dpi (old leaves)	Resistant	Aphids with virus	31, 32, 33
		Aphids without virus	34, 35, 36
	Susceptible	Aphids with virus	37, 38, 39
		Aphids without virus	40, 41, 42
14 dpi (young + old leaves)	Resistant	Aphids with virus	61, 63, 65 (old leaves) 62, 64, 66 (young leaves)
		Aphids without virus	67, 69, 71 (old leaves) 68, 70, 72 (young leaves)
		Insecticide spray	73, 75, 77 (old leaves) 74, 76, 78 (young leaves)
	Susceptible	Aphids with virus	79, 81, 83 (old leaves) 80, 82, 84 (young leaves)
		Aphids without virus	85, 87, 89 (old leaves) 86, 88, 90 (young leaves)
		Insecticide spray	91, 93, 95 (old leaves) 92, 94, 96 (young leaves)
21 dpi (young + old leaves)	Resistant	Aphids with virus	97, 99, 101 (old leaves) 98, 100, 102 (young leaves)
		Aphids without virus	103, 105, 107 (old leaves) 104, 106, 108 (young leaves)
		Insecticide spray	109, 111, 113 (old leaves) 110, 112, 114 (young leaves)
	Susceptible	Aphids with virus	115, 117, 119 (old leaves) 116, 118, 120 (young leaves)
		Aphids without virus	121, 123, 125 (old leaves) 122, 124, 126 (young leaves)
		Insecticide spray	127, 129, 131 (old leaves) 128, 130, 132 (young leaves)
28 dpi (young leaves)	Resistant	Aphids with virus	134, 136, 138
		Aphids without virus	140, 142, 144
		Insecticide spray	146, 148, 150
	Susceptible	Aphids with virus	152, 154, 156
		Aphids without virus	158, 160, 162
		Insecticide spray	164, 166, 168

The qPCR was executed using BioRad CFX Connect Real-Time System using the qPCR program (Figure 3). The results were obtained as viral copy numbers in the BioRad CFX Manager Version 3.1.

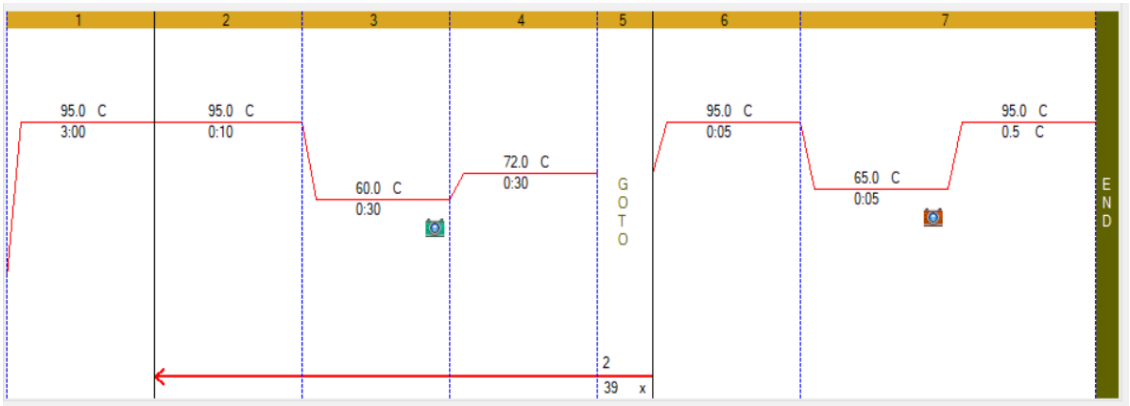


Figure 3: Schema of the qPCR run cycles, including amplification cycles and melt curve analysis

The results obtained were checked for their coefficient of determination (R^2>0.99), slope (between -3.1 and -3.6), efficiency (90-110%), and melt curve for the primers (Illumina, 2010). Primer pair I (forward sequence 5'-CAAAAGTTATGAGCGGTAGGCA-3' and reverse sequence 5'-CTTGTTCGTTGAACCGACTGC-3') and primer pair II (forward sequence 5'-AAAGTTATGAGCGGTAGGCAA-3' and reverse sequence 5'-CACTTGTTCGTTGAACCGACT-3') were used for quantification of BMYV. The following equation was used for calculating the viral copy number of standards, based on which the software calculated the viral copy numbers of the samples.

$$ng\ of\ dsDNA\ x\ Avogadro\ constant\ x\ \frac{1}{template\ length}\ x\ \frac{1}{660g}\ x\ \frac{1}{10^9}\ \frac{ng}{g}$$

Here, the template length signifies the total length of the PCR template for the standards and ng of dsDNA means the amount of template (plasmid) used. The results were then analysed through Microsoft Excel, where an unpaired student's t-test of significance with unequal variance was performed with a p-value of 0.05.

2.7 RT-qPCR for relative quantification

Based on the literature review, three genes related to pathogenesis were selected to test their differential expression between resistant and susceptible genotypes and across different treatments and time points. This was done using RT-qPCR for relative quantification.

Two genes induced by salicylic acid (SA) mediated defence pathway, pathogenesis-related protein 1a (*PR1a,* accession no: AM932128) and pathogenesis-related protein 5 (*PR5*, accession no: XM_010680025.2), as well as the gene Argonaute 1 (*AGO1*, accession no: XM_010688933), which is involved in RNA-silencing of the virus, were tested. Elongation Factor 2 (*EF2*) and glyceraldehyde-3-phosphate-dehydrogenase (*GAPD*/*GAPD*H) were used as reference genes. A study by Fernando Gil et al. (2020) showed these genes to be stable for data normalisation; thus, they were considered suitable reference genes, and the primer sequences were retrieved from their publication. Using the gene sequence information in sugar beet from the NCBI website, two primer pairs were designed for each gene of interest using the Primer-BLAST tool (Ye et al., 2012), considering basic primer requirements (Table 2).

Table 2: Sequences of forward and reverse primers for target genes and reference genes

Primer sets	Forward primer (5'-3')	Reverse primer (5'-3')
AGO1 I	ATGACCTCGCAACAGTCTGG	CACCTCTTGGATGCTGCTGA
II	GGTCCTTTCGTCCACCAGTT	CAGCCTGCATAGGTGACACA
EF2	AGCTGCGAAAATGGTGAAGT	AGCGTTGATTTCCCGTGA
GAPD	CACCACCGATTACATGACATACA	GGATCTCCTCTGGGTTCCTG
PR1a I	GCTTCTGGCAAAGTGTGTGG	CGACAAAGTTGCCTGGTGGA
II	TTGCAACACTAGCCCTAGCC	AGGCTGCCACTTGATCATCC
PR5 I	CAGGCTGCAATTTCGACACC	AAGTCTTGTCCACCTGCACC
II	GAGCGTGGTGGCTTGTAAGA	TGTAAGCTGTTGGACACGCA

The RT-qPCR was performed using Maxima SYBR Green/ROX qPCR Master Mix (2X), following the manufacturer's protocol. Initially, the primers were tested for their efficiency using a standard curve containing a cDNA mixture of 12 samples (six resistant and six susceptible). The cDNA stock solution for each of the 12 samples was diluted 10-fold. These diluted solutions were pooled to make a cDNA mix, from which six serial dilutions were prepared with five-fold dilution

at each stage. On the first qPCR run, the primer efficiency of the two housekeeping genes was tested, where water was used as a negative control. Next, the efficiency was tested for both primer pairs of the three target genes using the cDNA mix.

Furthermore, two separate cDNA mixes were used to determine if the target genes of both the resistant and susceptible genotypes bind to the primers: one for resistant and another for susceptible samples. For each genotype, the stock cDNA solutions for three samples were diluted 10-fold. The diluted solutions were pooled, and five serial dilutions were made with a five-fold dilution at each step.

The BioRad CFX Connect Real-Time System was used for the qPCR, and the results were obtained in BioRad CFX Manager. The efficiency for each primer was calculated using Microsoft Excel.

Among the genes whose primer pairs had high efficiency and gave lower Ct values, *AGO1*, *EF2* and *GAPD*, the second primer pair of *AGO1* and *EF2* as housekeeping gene were selected for further testing based on the efficiency and Ct values. For the quantification, cDNA of samples collected at time points 14 and 21 dpi from older leaves of susceptible plant exposed to viruliferous or non-viruliferous aphids were used. For each time point, three biological replicates and two technical replicates from each biological replicate were tested. On the same plate, the expression was analysed for *AGO1* and *EF2*. The relative expression was calculated using the Pfaffl method.

2.8 Bioinformatic analysis

A schematic representation of the RNA-seq analysis is given in Figure 4. The RNA sequencing results obtained at SciLifeLab were received by SLU Bioinformatics Infrastructure (SLUBI) and stored in the project repository: SUPR-SNIC. The Linux operating system was required to analyse results and thus was performed using ThinLinc Client 4.14.0. The supercomputer cluster UPPMAX was used for the bioinformatic analysis. Additionally, Ubuntu 18.04.05 for Windows was used for other operations related to the data in the Linux platform. The raw reads for a total of 174 samples comprising plant samples inoculated with BYV (54) or BMYV (120), including the corresponding controls, were then analysed using the nfcore/rnaseq analysis pipeline, which uses a Nextflow workflow (Ewels et al., 2020), and was personalised for this project. The workflow starts with visualising the quality (QC) of the raw reads in FASTQ files using FASTQC, followed by adapter trimming using the tool TrimGalore. The reference genome of sugar beet was added from the NCBI website (Dohm et al., 2014). Salmon was used for pseudo-

alignment, and the detailed results for all samples were obtained in a multiqc report after quality control (Ewels et al., 2016; Patro et al., 2017).

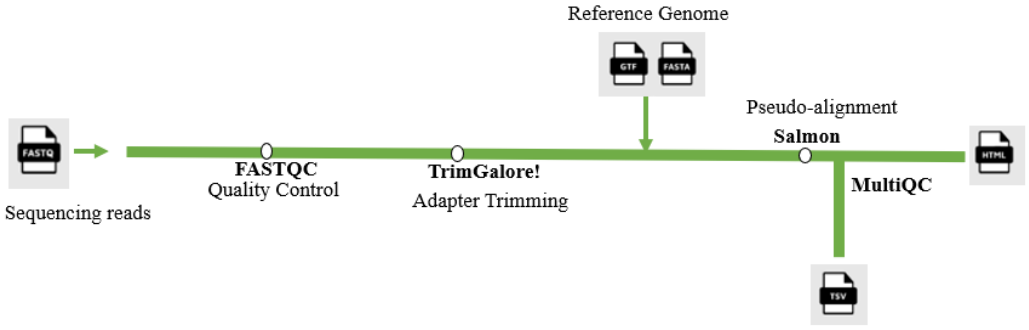


Figure 4: Pipeline for RNA-seq analysis

2.9 Statistical analysis

Differential expression analysis was performed on the RNA-seq data using RStudio with R version 4.2.0. The packages DESeq2, in addition to edgeR, were used for this purpose (Love et al., 2014; Robinson et al., 2010).

First, using the package tximport, the pseudo counts generated by Salmon as normalised transcripts per million (TPM) were transformed to non-normalised count estimates for DESeq2 analysis (Piper et al., 2017a). Then, the gene duplications were removed, followed by removing genes that were not expressed in any samples. The genes were further filtered by removing genes with less than a total of ten reads across all samples. The samples then underwent DESeq analysis (Love et al., 2022). For samples from resistant or susceptible genotypes exposed to BYV or BMYV, comaparison was between healthy plants and virus-inoculated plants (separate comparisons for older and younger leaves). Wald Test was performed for hypothesis testing, where the p-value was set as 0.05.

Moreover, the Benjamini-Hochberg method was used to adjust the p-value and correct the multiple testing to reduce the number of false positives. The differentially expressed genes (DEGs) were then filtered based on significance (p-adjusted value < 0.05) (Piper et al., 2017b). And the results were visualised using PCA plots. Bar plots were generated showing sequencing depth (total number of gene reads per sample), scatterplots showing mean vs variance for each comparison, and the number of upregulated and downregulated genes for comparisons between healthy and

inoculated samples at each timepoint from each leaf age and each virus. (Supplementary Tables 1-4). The script provided by SLUBI for differential gene expression analysis, with some changes, is included in the appendix.

3. Results

3.1 Absolute quantification

After virus inoculation, foliar symptoms were evaluated in the plants found on the susceptible plants, while resistant plants did not show any symptoms (unpublished data from the research group).

For all runs with primer pair II except for samples 31-42, the R^2 value was optimum (>0.99), while it was slightly lower when using primer pair I (0.976). Furthermore, the efficiency was within the desired range for some samples, while samples from 97-168 and repeats of samples 1-6 with primer pair II showed lower efficiency. The slope value was also lower than within the desired range for the samples with low efficiency (Table 3)

Table 3: Coefficient of determination (R^2), efficiency, and slope for different PCR runs for absolute quantification of BMYV

BMYV samples	R^2 value	Efficiency	Slope value
1-6 Primer I	0.976	102.2 %	-3.271
7-18 + 61-71 Primer II	0.991	92.6%	-3.514
31-42 Primer II	0.978	100.7%	-3.306
72-96 Primer II	0.997	98.4%	-3.453
97-119 Primer II	0.998	87.8%	-3.653
120-152 Primer II	0.994	83.8%	-3.784
154-168 + 1-6 Primer II	0.991	79.4%	-3.939

The melt curve analysis showed multiple peaks for runs using primer pair I (samples 1-6), while primer pair II showed a single peak for all qPCR runs (Figure 5).

A)

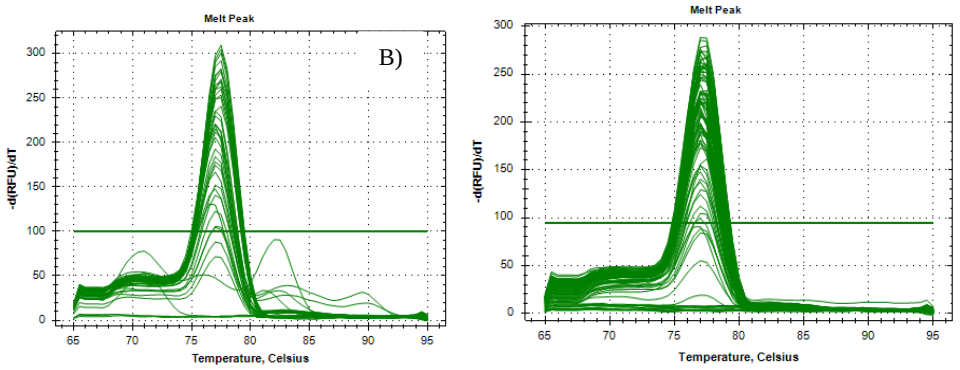


Figure 5: Melt curves generated for qPCR runs with a) primer pair I and b) primer pair II.

The virus titres from the plant samples of susceptible and resistant genotypes exposed to BMYV were measured using qPCR in older and younger leaves at five time points (Figure 6, 7). The older leaves showed an increase in virus titre at later time points. For resistant plants, the highest viral load was seen at 21 dpi; for susceptible plants, the viral load was highest at 14 dpi, which decreased at 21 dpi. While the virus titre in susceptible plants is higher than that of resistant plants at all four time points (1, 4, 14, and 21 dpi), the difference was not statistically significant.

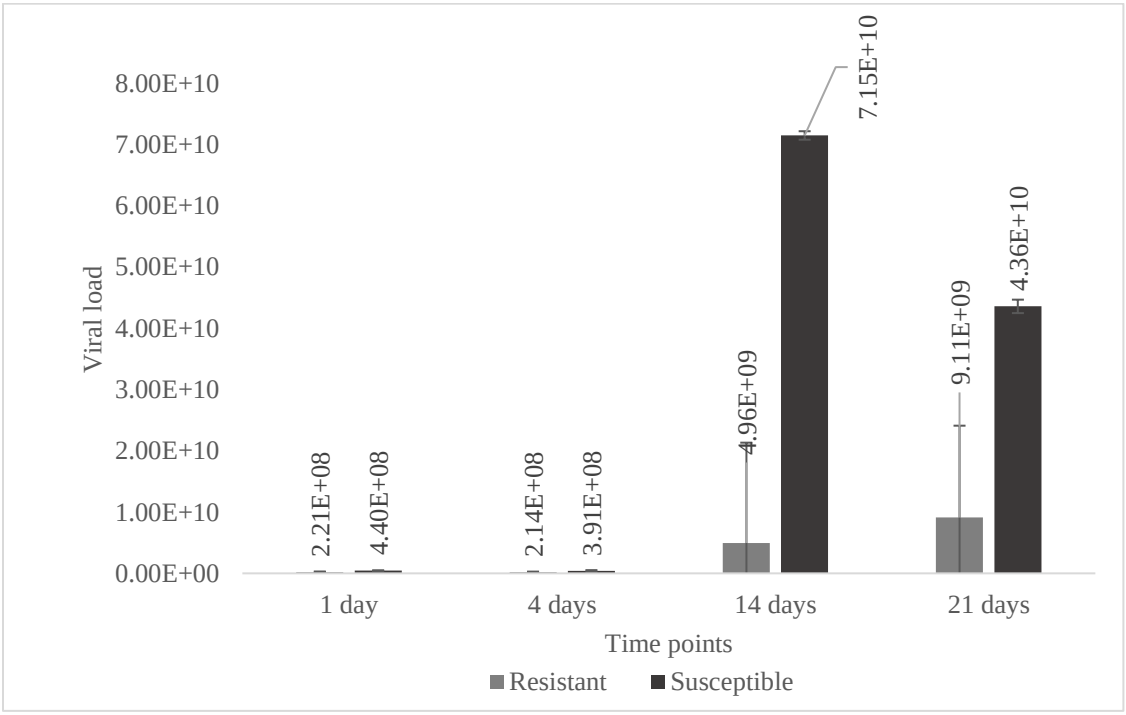


Figure 6: Bar graph showing the copy number of BMYV in 1 µg RNA in older infected leaves from resistant and susceptible plants at different time points. Error bar shows standard deviation.

In younger leaves (leaves directly above the inoculated leaves) exposed to BMYV, the virus titre in susceptible plants was highest at 21 dpi. The viral load increased at 21 dpi from 14 dpi and decreased at 28 dpi. However, in resistant plants, the virus titre decreased from 14 dpi to 21 dpi and then increased at 28 dpi. The viral load at 14 and 21 dpi was lower in younger leaves compared to older leaves, which was true for both resistant and susceptible genotypes. In yougner leaves, the virus titre was higher in susceptible plants at all time points; however, this difference between resistant and susceptible genotypes was only significant at 21 dpi.

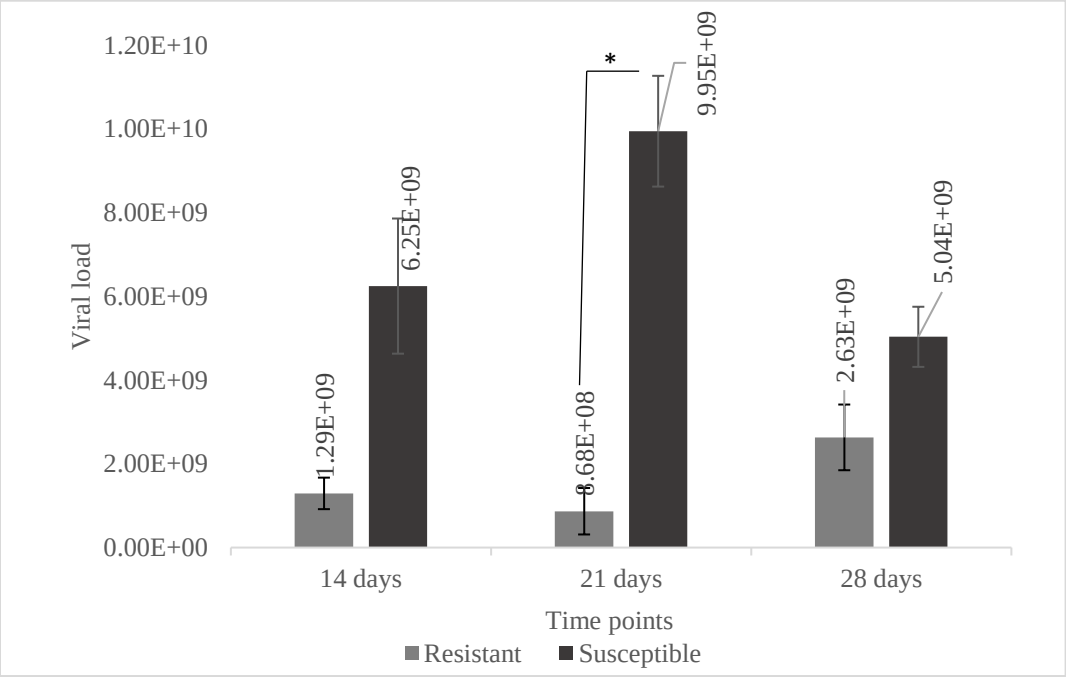


Figure 7: Bar graph showing the copy number of beet mild yellowing virus in 1 µg RNA in younger infected leaves from resistant and susceptible plants at different time points. Error bar shows standard deviation. Analysis performed by unpaired student's t-test with p-value ≤ 0.05*.

3.2 Relative quantification

The qPCR test for primer efficiency of target genes and housekeeping genes using a cDNA mix of susceptible and resistant genotypes generated very high Ct values (Ct >30) for *PR1a* and *PR5*; thus, they were not tested further. However, the two housekeeping genes (*EF2* and *GAPD*) and target gene *AGO1* showed Ct values lower than 30 and were further tested using separate cDNA mixes for resistant and susceptible plants. They showed higher efficiency (>100%) and lower Ct values for samples from susceptible plants but lower efficiency (<60%) and higher Ct

values for samples from resistant plants (Table 4). Thus, qPCR analyses were only carried out to monitor the expression of *AGO1* and *EF2*.

Table 4: Primer efficiencies of primers of housekeeping and target genes tested with qPCR

Genotypes	**Genes tested**	**Efficiency**
Susceptible + Resistant	*EF2*	90.20%
	GAPD	91.87%
	PR1a	n/a
	PR5	n/a
	AGO1	87.03%, 72.50%
Resistant	*EF2*	53.05%
	GAPD	51.64%
	AGO1	54.96%
Susceptible	*EF2*	105.35%
	GAPD	105.63%
	AGO1	101.48%

The results were analysed using the Pfaffl method and the true efficiencies for each run. Susceptible plants inoculated with BMYV showed lower expression of *AGO1* as compared to the untreated group at both time points (14 and 21 dpi). However, this difference in expression was not significant (Figure 8).

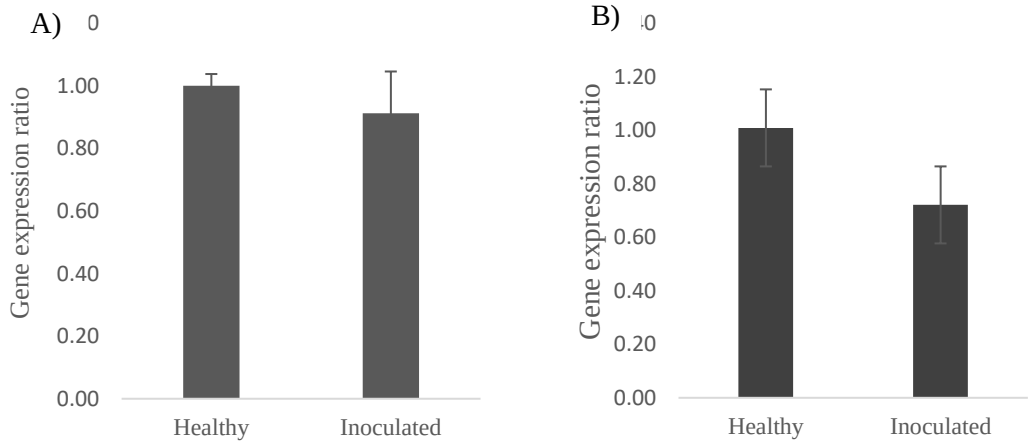
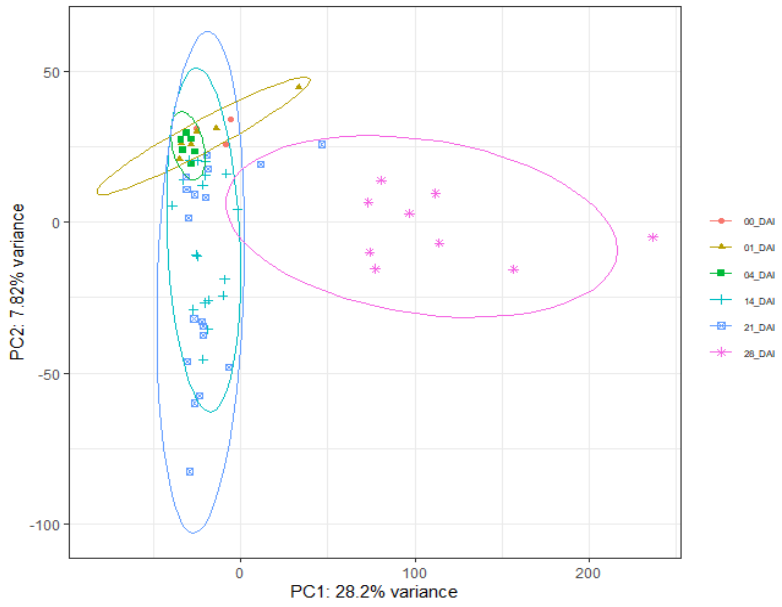


Figure 8: Relative gene expression of *AGO1* comparing BMYV-inoculated and healthy plants of the susceptible genotype across a) 14 dpi and b) 21 dpi. Error bars show standard deviation.

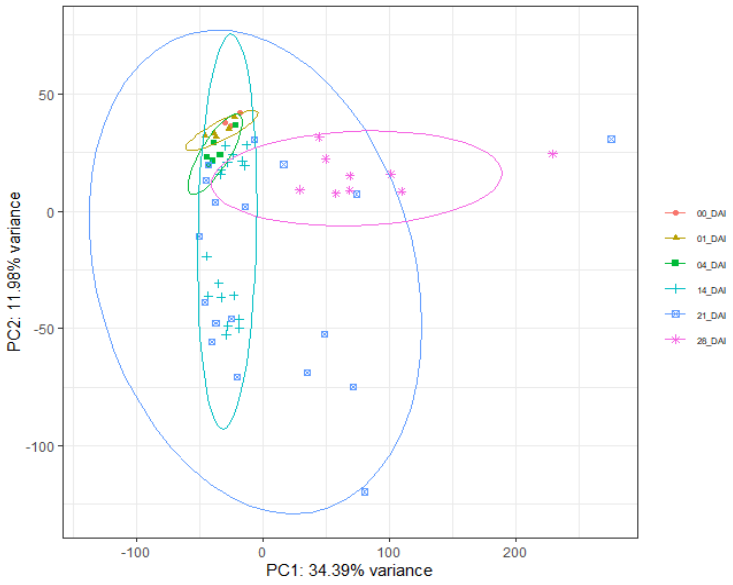
3.3 RNA sequencing analysis

RNA sequencing analysis using nf-core/rnaseq (with Nextflow workflow) and DESeq2 was performed to examine transcript level differences across the two genotypes and their treatments with BYV and BMYV. All four plots of principal component analysis (PCA) in figure 9 show a distinct clustering of gene expression for plants inoculated with BMYV or BYV at different time points.

A)



B)



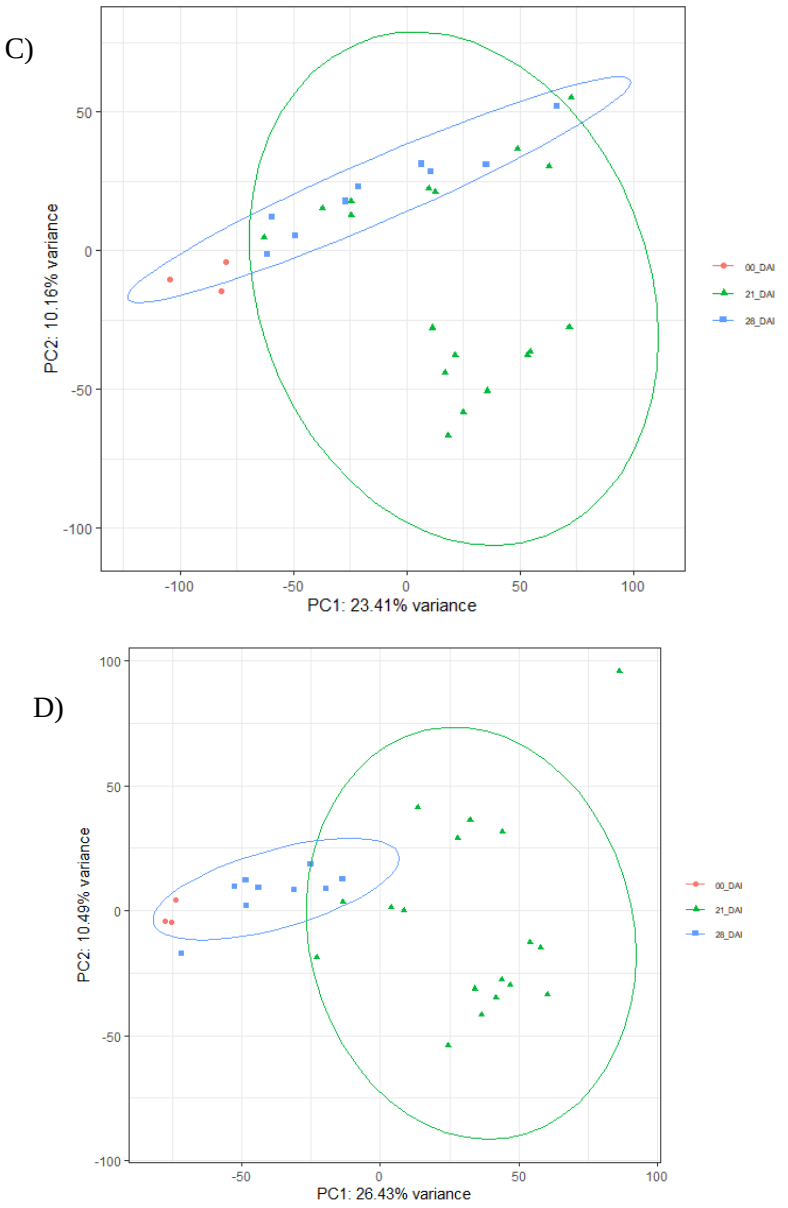


Figure 9: Principal-component analysis (PCA) visualisation of variation in gene expression between resistant and susceptible plants exposed to BYV or BMYV at different time points. The colours and symbols in the plots represent time points for A) resistant plants inoculated with BMYV, B) susceptible plants inoculated with BMYV, C) resistant plants inoculated with BYV, and D) susceptible plants inoculated with BYV.

Additional visualisation of results through bar plot of sequencing depth, scatterplot for mean vs variance and MA plot portraying the relationship between the base means and log2fold change of the counts were generated after normalisation. (Supplementary Tables 1-4). For each of

the mean vs variance comparisons, the variance was higher than the mean, and thus variance stabilising transformation was performed during differential gene expression analysis in RStudio.

Moreover, an increase in differentially expressed genes (DEGs) was observed at later time points. In samples from susceptible plants exposed to BYV, older leaves showed some DEGs, while no DEGs were present for the younger leaves. In contrast, a comparison of transcript levels for older leaves from resistant genotypes exposed to BYV showed no DEGs compared to healthy plants, while 23 genes were downregulated in younger leaves at 28 dpi. Older leaves from susceptible plants exposed to BMYV showed an increasing number of DEGs with later time points (14 and 21 dpi), and the same pattern was seen in younger leaves at 14 and 21 dpi, while the DEGs decreased at 28 dpi. This decrease at 28 dpi was also seen across the comparisons in younger leaf samples from resistant plants exposed to BMYV, while in older leaves, the DEGs decreased from 1 to 0 at 21 dpi (Supplementary Tables 1-4).

4 Discussion

Virus yellows is a widespread disease in Europe as well as in the United States (Stevens et al., 2005b). Molecular approaches such as virus quantification facilitate reliable assessment of viral load. Due to the high sensitivity of this method, the results cannot be swayed by the effect of environmental factors (Shirima et al., 2017). Such an approach has been applied to measure the viral load of southern rice black-streaked dwarf virus (SRBSDV) transmitted by the white-backed planthoppers at different developmental stages in rice (An et al., 2015).

The measurement of BMYV titre at different time points (in copy numbers) for the two genotypes showed a general trend of higher viral load in plants of the susceptible genotype compared to the resistant genotype. Moreover, at later time points (14, 21, and 28 dpi), the virus titre was found to be higher than at earlier time points (1 and 4 dpi). The patterns of increased or decreased viral load at later time points differ between the resistant and susceptible genotypes. Moreover, the absence of symptoms but a significant virus titre (for example, 9.11×10^9 viral RNA copies at 21 dpi in older leaves) suggests that the resistant genotype may not be expressing complete resistance but tolerance. Further analysis of DEGs obtained from RNA-seq analysis in resistant genotypes could help better understand the type of resistance and the number of genes involved.

We assessed the expression levels of *PR-1a*, *PR-5* and *AGO1* genes in sugar beet leaf tissue exposed to viruliferous or non-viruliferous aphids. ARGONAUTES (AGOs) are effector proteins that play significant roles in RNA silencing/degradation. Wu et al. (2015) revealed that *AGO18* positively regulated *AGO1* for its antiviral activity in rice against rice stripe tenuivirus (RSV) and rice dwarf phytoreovirus (RDV). Moreover, the role of *AGO1* in temperature-dependent symptom recovery in *N. benthaminana* plants infected with tomato ringspot virus was recorded in the study by Ghoshal and Sanfaçon (2014). As AGO1, a component of RISC, is involved in antiviral defence, it would be logical that its expression is induced/increased by virus infection. The results matched our expectations in susceptible genotypes when compared for time points 14 and 21 dpi (albeit statistically not significant). Further qPCR runs at more time points would more clearly show the role of *AGO1* in resistance to virus yellows.

Generally, PR proteins are implicated in plant defence against biotrophic pathogens, and the accumulation of PR proteins is induced by SA to confer SAR against fungi, bacteria, and viruses (Ali et al., 2018, Padmanabhan et al., 2004). Okushima et al. (2000) found that PR-17 stimulates defence against tobacco mosaic virus (TMV) in *N. tabacum*, whereas Park et al. (2004) recorded the ribonucleolytic role of PR-10 in an antiviral pathway in *Capsicum annum* against TMV. A study by Elvira et al. (2008) noted the resistance induction by *PR-5* and *PR-1* when induced by pepper mild mottle virus (PMMoV) in *Capsicum chilense*. An increase in expression of PR1a and PR5 would mean that the SA-induced defence has been activated. However, this study could not verify this as the primer testing or *PR1a* and *PR5* showed very low expression when tested with a cDNA mix of resistant and susceptible samples. The primer efficiency for the genes *PR1a* and *PR5* should be tested separately in resistant and susceptible samples to understand their amplification.

The expression of the target and housekeeping genes was not detected in plant samples of the resistant genotype. The primer sequences used for both the resistant and susceptible genotypes were derived from the same reference genome of sugar beet. Without an available genome sequence for the resistant genotype, the difference in gene expression between the genotypes could not be determined. Transcriptome assembly from the RNA-seq data for the resistant genotype could be performed to optimise primer design.

The RNA-sequencing analysis showed a clear distinction in gene expression between the samples at different time points. Further comparisons should be made (comparing genotypes or virus treatments) to further understand the variance between samples. The DEGs were increasing at later time points in older leaves of susceptible plants exposed to BMYV. A similar trend was not observed for resistant plants. Although at 21 dpi, older leaves of susceptible samples plants exposed to BYV showed some DEGs (compared to very few DEGs in younger leaf samples of resistant plants), this number was much lower for samples of plants inoculated with BMYV. It could be implied that the resistance mechanisms against these two viruses are different.

5 Conclusion

The absolute quantification of BMYV was performed to analyse virus titre in sugar beet leaf samples. Moreover, relative expression of pathogenesis-related genes was assessed across different time points comparing virus inoculated and healthy samples from susceptible sugar beet. This was carried out with the main objective of understanding the resistance mechanism of sugar beet against virus yellows and developing a breeding program for resistant cultivars. Further analysis of upregulated and downregulated genes obtained from the RNA sequencing analysis is required to better understand the resistance mechanism. Details comparisons, as done in this study, can be followed with bioinformatic analyses and qPCRs to validate the expression of these DEGs for specific time points or treatments to draw a clearer picture of the defence mechanisms against BYV and BMYV in resistant and susceptible genotypes.

6 References

Ali, S., Ganai, B. A., Kamili, A. N., Bhat, A. A., Mir, Z. A., Bhat, J. A., Tyagi, A., Islam, S. T., Mushtaq, M., Yadav, P., Rawat, S., & Grover, A. (2018). Pathogenesis-related proteins and peptides as promising tools for engineering plants with multiple stress tolerance. Microbiological Research, 212–213, 29–37. https://doi.org/10.1016/j.micres.2018.04.008

An, X.-K., Hou, M.-L., & Liu, Y.-D. (2015). Relation Between the Viral Load Accumulation of Southern Rice Black-Streaked Dwarf Virus and the Different Developmental Stages of *Sogatella furcifera* (Hemiptera: Delphacidae). Journal of Economic Entomology, 108(3), 917–924. https://doi.org/10.1093/jee/tov065

Bass, C., & Field, L. M. (2018). Neonicotinoids. Current Biology, 28(14), R772–R773. https://doi.org/10.1016/j.cub.2018.05.061

Bayer Crop Science UK (2018). Sugar Beet Diseases. https://cropscience.bayer.co.uk/threats/diseases/sugar-beet-diseases/

Bennet, C. W (1960). Sugar beet virus yellows in the United States. Agricultural Research Service. Technical Bulletin, 1218. https://naldc.nal.usda.gov/download/CAT87201161/PDF

Bortolamiol, D., Pazhouhandeh, M., Marrocco, K., Genschik, P., & Ziegler-Graff, V. (2007). The polerovirus F box protein P0 targets ARGONAUTE1 to suppress RNA silencing. Current Biology, 17(18), 1615–1621. https://doi.org/10.1016/j.cub.2007.07.061

Brault, V., Herrbach, É., & Reinbold, C. (2007). Electron microscopy studies on luteovirid transmission by aphids. Micron, 38(3), 302–312. https://doi.org/10.1016/j.micron.2006.04.005

Bruening, G. (2006). Resistance to Infection. In G. Loebenstein & J. P. Carr (Eds.), Natural Resistance mechanisms of plants to viruses (pp. 211–240). Springer Netherlands. https://doi.org/10.1007/1-4020-3780-5_10

Carr, J. P., Tungadi, T., Donnelly, R., Bravo-Cazar, A., Rhee, S.-J., Watt, L. G., Mutuku, J. M., Wamonje, F. O., Murphy, A. M., Arinaitwe, W., Pate, A. E., Cunniffe, N. J., & Gilligan, C. A. (2020). Modelling and manipulation of aphid-mediated spread of non-persistently transmitted viruses. Virus Research, 277, 197845. https://doi.org/10.1016/j.virusres.2019.197845

Delfosse, V. C., Barrios Barón, M. P., & Distéfano, A. J. (2021). What we know about poleroviruses: Advances in understanding the functions of polerovirus proteins. Plant Pathology, 70(5), 1047–1061. https://doi.org/10.1111/ppa.13368

Dohm, J. C., Minoche, A. E., Holtgräwe, D., Capella-Gutiérrez, S., Zakrzewski, F., Tafer, H., Rupp, O., Sörensen, T. R., Stracke, R., Reinhardt, R., Goesmann, A., Kraft, T., Schulz, B., Stadler, P. F., Schmidt, T., Gabaldón, T., Lehrach, H., Weisshaar, B., & Himmelbauer, H. (2014). The genome of the recently domesticated crop plant sugar beet (*Beta vulgaris*). Nature, 505(7484), 546–549. https://doi.org/10.1038/nature12817

Elvira, M. I., Galdeano, M. M., Gilardi, P., García-Luque, I., & Serra, M. T. (2008). Proteomic analysis of pathogenesis-related proteins (PRs) induced by compatible and incompatible interactions of pepper mild mottle virus (PMMoV) in *Capsicum chinense* L3 plants. Journal of Experimental Botany, 59(6), 1253–1265. https://doi.org/10.1093/jxb/ern032

European Commission (2018). Neonicotinoids. https://ec.europa.eu/food/plants/pesticides/approval-active-substances/renewal-approval/neonicotinoids_en

Ewels, P. A., Peltzer, A., Fillinger, S., Patel, H., Alneberg, J., Wilm, A., Garcia, M. U., Di Tommaso, P., & Nahnsen, S. (2020). The nf-core framework for community-curated bioinformatics pipelines. Nature Biotechnology, 38(3), 276–278. https://doi.org/10.1038/s41587-020-0439-x

Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: Summarize analysis results for multiple tools and samples in a single report. Bioinformatics, 32(19), 3047–3048. https://doi.org/10.1093/bioinformatics/btw354

FAOSTAT. (2020). Crops and livestock products. https://www.fao.org/faostat/en/#data/QCL/visualize

Fernando Gil, J., Wibberg, D., Eini, O., Savenkov, E. I., Varrelmann, M., & Liebe, S. (2020). Comparative transcriptome analysis provides molecular insights into the interaction of beet necrotic yellow vein virus and beet soil-borne mosaic virus with their host sugar beet. Viruses, 12(1), 76. https://doi.org/10.3390/v12010076

Ghoshal, B., & Sanfaçon, H. (2014). Temperature-dependent symptom recovery in *Nicotiana* benthamiana plants infected with tomato ringspot virus is associated with reduced translation of

viral RNA2 and requires ARGONAUTE 1. Virology, 456–457, 188–197. https://doi.org/10.1016/j.virol.2014.03.026

Gray, S., & Gildow, F. E. (2003). Luterovirus-aphid interactions. Annual Review of Phytopathology, 41, 539-566. https://doi.org/10.1146/annurev.phyto.41.012203.105815

Hauer, M., Hansen, A. L., Manderyck, B., Olsson, Å., Raaijmakers, E., Hanse, B., Stockfisch, N., & Märländer, B. (2017). Neonicotinoids in sugar beet cultivation in Central and Northern Europe: Efficacy and environmental impact of neonicotinoid seed treatments and alternative measures. Crop Protection, 93, 132–142. https://doi.org/10.1016/j.cropro.2016.11.034

Hossain, R., Menzel, W., Lachmann, C., & Varrelmann, M. (2021). New insights into virus yellows distribution in Europe and effects of beet yellows virus, beet mild yellowing virus, and beet chlorosis virus on sugar beet yield following field inoculation. Plant Pathology, 70(3), 584–593. https://doi.org/10.1111/ppa.13306

Igori, D., Lim, S., Cho, H. S., Kim, H.-S., Park, J. M., Lee, H.-J., Kwon, S.-Y., & Moon, J. S. (2021). Complete genome sequence of platycodon closterovirus 1, a novel putative member of the genus *Closterovirus*. Archives of Virology, 166(7), 2051–2054. https://doi.org/10.1007/s00705-021-05081-3

Illumina (2010). Absolute Quantification of Gene Expression using SYBR Green. https://www.illumina.com/Documents/products/technotes/technote_eco_absolute_quantification_using_sybrgreen.pdf

Jordbruksverket (2019) Jordbruksmarkens användning 2019—JO10SM1902—Kommentarer. https://djur.jordbruksverket.se/webdav/files/SJV/Amnesomraden/Statistik,%20fakta/Arealer/JO10/JO10SM1902/JO10SM1902_kommentarer.htm

Kozłowska-Makulska, A., Beuve, M., Syller, J., Szyndel, M. S., Lemaire, O., Bouzoubaa, S., & Herrbach, E. (2009). Aphid transmissibility of different European beet polerovirus isolates. European Journal of Plant Pathology, 125(2), 337–341. https://doi.org/10.1007/s10658-009-9474-7

LaTourrette, K., Holste, N. M., & Garcia-Ruiz, H. (2021). Polerovirus genomic variation. Virus Evolution, 7(2), veab102. https://doi.org/10.1093/ve/veab102

Lin, S.-S., Henriques, R., Wu, H.-W., Niu, Q.-W., Yeh, S.-D., & Chua, N.-H. (2007). Strategies and mechanisms of plant virus resistance. Plant Biotechnology Reports, 1(3), 125–134. https://doi.org/10.1007/s11816-007-0021-8

Liu, Y., Zhou, J., & White, K. P. (2014). RNA-seq differential expression studies: More sequence or more replication? Bioinformatics, 30(3), 301–304. https://doi.org/10.1093/bioinformatics/btt688

Love, M., Ahlmann-Eltze, C., Forbes, K., Anders, S., Huber, W., FP7, R. E., NHGRI, N., & CZI. (2022). DESeq2: Differential gene expression analysis based on the negative binomial distribution (1.36.0). Bioconductor version: Release (3.15). https://doi.org/10.18129/B9.bioc.DESeq2

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology, 15(12), 550. https://doi.org/10.1186/s13059-014-0550-8

Mukherjee, E., & Gantait, S. (2022). Genetic transformation in sugar beet (*Beta vulgaris* L.): Technologies and applications. Sugar Tech. https://doi.org/10.1007/s12355-022-01176-6

Nilsson, B. & Larsson, H. (1990). Virusgulsot i sockerbetor - historik, biologi, spridning, effekter på skörden samt möjligheter till prognos och bekämpning. Svenska växtskyddskonferensen, vol. 31, pp. 299–318

OECD. (2020). World sugar projections. Organisation for Economic Co-operation and Development. https://www.oecd-ilibrary.org/agriculture-and-food/world-sugar-projections_b62e023c-en

Palukaitis, P., & Carr, J. (2008). Plant resistance responses to viruses. J Plant Pathol, 90.

Park, C.-J., Kim, K.-J., Shin, R., Park, J. M., Shin, Y.-C., & Paek, K.-H. (2004). Pathogenesis-related protein 10 isolated from hot pepper functions as a ribonuclease in an antiviral pathway. The Plant Journal, 37(2), 186–198. https://doi.org/10.1046/j.1365-313X.2003.01951.x

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., & Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. Nature Methods, 14(4), 417–419. https://doi.org/10.1038/nmeth.4197

Piper, M., Mistry, M., & Khetani, R. (2017a). Set up and overview for gene-level differential

expression. Introduction to DGE. analysis. https://hbctraining.github.io/DGE_workshop_salmon/lessons/01_DGE_setup_and_overview.htm

Piper, M., Mistry, M., & Khetani, R. (2017b). Gene-level differential expression analysis

With DESeq2. Introduction to DGE. https://hbctraining.github.io/DGE_workshop_salmon/lessons/05_DGE_DESeq2_analysis2.html

Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics, 26(1), 139–140. https://doi.org/10.1093/bioinformatics/btp616

Ruus, A. G. (2020). Virus yellows in sugar beets in Sweden and Europe. SLU. BSc Thesis

Shirima, R. R., Maeda, D. G., Kanju, E., Ceasar, G., Tibazarwa, F. I., & Legg, J. P. (2017). Absolute quantification of cassava brown streak virus mRNA by real-time qPCR. Journal of Virological Methods, 245, 5–13. https://doi.org/10.1016/j.jviromet.2017.03.003

Stevens, M., Freeman, B., Liu, H.-Y., Herrbach, E., & Lemaire, O. (2005a). Beet poleroviruses: Close friends or distant relatives? Molecular Plant Pathology, 6(1), 1–9. https://doi.org/10.1111/j.1364-3703.2004.00258.x

Stevens, M., Hallsworth, P. B., & Smith, H. G. (2004). The effects of beet mild yellowing virus and beet chlorosis virus on the yield of UK field-grown sugar beet in 1997,1999 and 2000. Annals of Applied Biology, 144(1), 113–119. https://doi.org/10.1111/j.1744-7348.2004.tb00323.x

Stevens, M., Patron, N. J., Dolby, C. A., Weekes, R., Hallsworth, P. B., Lemaire, O., & Smith, H. G. (2005b). Distribution and properties of geographically distinct isolates of sugar beet yellowing viruses. Plant Pathology, 54(2), 100–107. https://doi.org/10.1111/j.1365-3059.2005.01155.x

Walker, P. J., Siddell, S. G., Lefkowitz, E. J., Mushegian, A. R., Adriaenssens, E. M., Alfenas-Zerbini, P., Davison, A. J., Dempsey, D. M., Dutilh, B. E., García, M. L., Harrach, B., Harrison, R. L., Hendrickson, R. C., Junglen, S., Knowles, N. J., Krupovic, M., Kuhn, J. H., Lambert, A. J., Łobocka, M., … Zerbini, F. M. (2021). Changes to virus taxonomy and to the International Code of Virus Classification and Nomenclature ratified by the International Committee on Taxonomy of Viruses (2021). Archives of Virology, 166(9), 2633–2648. https://doi.org/10.1007/s00705-021-05156-1

Wetzel, V., Brault, V., & Varrelmann, M. 2018. (n.d.). Production of a beet chlorosis virus full-length cDNA clone by means of Gibson assembly and analysis of biological properties. Journal of General Virology, 99(11), 1522–1527. https://doi.org/10.1099/jgv.0.001146

Wintermantel, W. M. (2005). Co-infection of beet mosaic virus with beet yellowing viruses leads to increased symptom expression on sugar beet. Plant Disease, 89(3), 325–331. https://doi.org/10.1094/PD-89-0325

Wu, J., Yang, Z., Wang, Y., Zheng, L., Ye, R., Ji, Y., Zhao, S., Ji, S., Liu, R., Xu, L., Zheng, H., Zhou, Y., Zhang, X., Cao, X., Xie, L., Wu, Z., Qi, Y., & Li, Y. (2015). Viral-inducible Argonaute18 confers broad-spectrum virus resistance in rice by sequestering a host microRNA. ELife, 4. https://doi.org/10.7554/eLife.05733

Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., & Madden, T. L. (2012). Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. BMC Bioinformatics, 13(1), 134. https://doi.org/10.1186/1471-2105-13-134

Yoshida, N., & Tamada, T. (2019). Host range and molecular analysis of beet leaf yellowing virus, Beet western yellows virus-JP and brassica yellows virus in Japan. Plant Pathology, 68(6), 1045–1058. https://doi.org/10.1111/ppa.13023

Zicari, S., Zhang, R., & Kaffka, S. (2019). Chapter 13—Sugar Beet. In Z. Pan, R. Zhang, & S. Zicari (Eds.), Integrated processing technologies for food and agricultural by-products. (pp. 331–351). Academic Press. https://doi.org/10.1016/B978-0-12-814138-0.00013-7

7 Appendix

Script for differential gene expression analysis on RStudio

```
# DGE analysis ----
# Script to perform differential gene expression analysis using DESeq2 package
# ---- Abu B Siddique, SLUBI, SLU, Uppsala, Sweden

## Burning questions that we will answer today!
# 1. How many genes do we have in total from our reference data?
# 2. How many genes remain after filtering?
# 3. How many genes are expressed and how many are not expressed?
# 4. Sequencing depth reads or genes (how many genes per sample)?
# 5. How many up-regulated & down-regulated genes?
# 6. Plots: PCA, MA, Volcano, Heat maps?
-------------------------------------------------------------------------------
# Step 1. Preparing packages, libraries & uploading data ----

# clean your r environment
rm(list = ls())

# install packages and load libraries
if (!require("BiocManager", quietly = TRUE))
   install.packages("BiocManager")
 BiocManager::install("DESeq2")
 BiocManager::install("tximport")
 BiocManager::install("edgeR")

library(DESeq2)
library(tidyverse)
library(readr)
library(tximport)
library(edgeR)
library(ggplot2)

## **** Prepare sample metadata before coming here in RStudio!!
## Keep all the salmon samples in one folder, and take other files to another sub-folder
## then do as follows ->

## 1.1. set working directory ---where the salmon folder is.
setwd("your directory where you have the SALMON folder")
getwd() # you can see where it is

## 1.2. sample meta data (coldata) ----
# Note: Load all metadata (or sample sheet) about your experimental set up and sequencing IDs,
```

```
#You may modify them first according to your wish and then upload it with the following
command
samples <- read_delim("final_salmon_analysis/metadata_samples.csv",
                      delim = ",", escape_double = FALSE, trim_ws = TRUE)
# you can call it 'colData' too

### 1.2.1. Subsampling ----you can regroup your data and do new analysis from here
# you can do similar subsamplings for BYV

## BMYV
#samples <- subset(samples, virus == 'BMYV') # BMYV (all BMYV including resistant and
susceptible)
samples <- subset(samples, virus == 'BMYV' & genotype == 'Resistant' & time_point ==
'04_DAI') # BMYV + only Susceptible +04_DAI
samples <- subset(samples, virus == 'BMYV' & genotype == 'Resistant') # Only Resistant
BMYV samples
view(samples)

## 1.3. count data ----
# Load file named as 'tx2gene' means transcripts to genes (translation table from the salmon
folders)

library(readr)
tx2gene <- read_delim("final_salmon_analysis/salmon_tx2gene.tsv",
                      delim = "\t", escape_double = FALSE,
                      col_names = c("transcript_id","gene_id1","gene_id2"), trim_ws = TRUE)
head(tx2gene) # Column '1' is transcript ID and '2' is gene ID , column 3 = column 2

## Note: Now, load the sample name from salmon folders (files as a list from the salmon folder)
#and load all the quantification data files from the salmon folder named 'quant.sf'.

files <- file.path("final_salmon", samples$SampleID, "quant.sf")
#file path-folder inside the working directory where your salmon folder is.
# The folder name for each sample should be the same as the first column name of the
#samplesheet (in this case SampleID column)

names(files) <- paste0(samples$SampleID)
head(files) # see the names
# or by names(files) <- samples$SampleID

## 1.4. import quantification data (tximport) ----
# We import the necessary quantification data for DESeq2 using the tximport function.
# For further details on use of tximport, including the construction of the tx2gene
```

```
# table for linking transcripts to genes in your dataset, please refer to the
# tximport package vignette = http://bioconductor.org/packages/tximport .

txi <- tximport(files, type = "salmon", tx2gene = tx2gene)

# Components or attributes or sheet of txi file:
attributes(txi)
# you will see four attributes like this:
# [1] "abundance"           "counts"              "length"
# [4] "countsFromAbundance"

# Explanation: "abundance is provided by quantification tools like TPM (transcripts-per-million),
# while the counts are estimated counts (possibly fractional), and
# the "length" matrix contains the effective gene lengths.
# The "length" matrix can be used to generate an offset matrix for downstream gene-level
differential analysis of count matrices"

## 1.5. *** Quality Control -----
### 1.5.1. meta data = count data?----
# Now make sure that the sample name in the meta data (samples) matches the name of the
#Counts (txi$counts). First, see if the 'row names in samples' matches to 'column names in
#counts_data'

all(colnames(counts) %in% rownames(samples))
# false (that means they are not same) !!! But why?
colnames(counts)
rownames(samples) ## they are not same because the samples has no row names

# solution: so make row names and save the new sample sheet as 'samples_updated'
samples <- as.data.frame(samples)
samples_updated <- as.data.frame(samples[,-1])
rownames(samples_updated) <- samples[,1]
# are they in the same order now?
all(colnames(counts) == rownames(samples_updated))
# TRUE! yes they are.

### 1.5.2. How many genes are not expressed at all in our samples? ----
# the data that may cause problems or are not needed for further analysis
## check all the 'unexpressed genes' and their count data
sel <- rowSums(counts) == 0
sprintf("%s%% percent (%s) of %s genes are not expressed",
      round(sum(sel) * 100/ nrow(counts),digits=2),
      sum(sel),
```

```
      nrow(counts)) # *** important results, it will print for example :"27.3% percent (10484) of
38338 genes are not expressed"
### ## 1.5.3. ** sequencing depth ----
## Second, check all samples and do they have robust sequencing depth or are there outliers?
#* Let us take a look at the sequencing depth, colouring by different "groups"

# create a data frame from 'counts' having all the column names and sum of the each sample
column
dat <- tibble(x=colnames(counts),y=colSums(counts)) %>%
  bind_cols(samples)

# plot sum of all gene counts # genotype wise
ggplot(dat,aes(sample_condition ,y,fill=genotype)) + geom_col() +
  scale_y_continuous(name="Gene_reads") +
  theme(axis.text.x=element_text(angle=90,size=4),axis.title.x=element_blank())

## *** do the same for other attributes (e.g., time_points, treatment,leaf_age etc)

##  1.6 Detailed data exploration (new but optional) ----
### 1.6.1. mean vs variance ---- why do we compare mean and variance here?
# we would like to see mean variance relationship between first three samples,as they are from
same genotype
data_counts <- txi$counts
mean_counts <- apply(data_counts[,1:3], 1, mean)
#The second argument '1' of 'apply' function indicates the function being applied to rows. Use '2'
if applied to columns
variance_counts <- apply(data_counts[,1:3], 1, var)
df_counts <- data.frame(mean_counts, variance_counts)
# or you can see the mean vs variance for all the samples within a single subsetting. change 3 to
the last number of sample in that case

# mean vs variance plot
ggplot(df_counts) +
  geom_point(aes(x=mean_counts, y=variance_counts)) +
  scale_y_log10(limits = c(1,1e9)) +
  scale_x_log10(limits = c(1,1e9)) +
  geom_abline(intercept = 0, slope = 1, color="red")
# what does it mean and why this relationship is important for further data analysis?

# STOP!? ----
# you must have a choice which group or meta column or variable you are interested on
# *** you can manipulate this script according to your interest of research question :)
```

```
# Step 2. construct a DESeqDataSet object ----------
dds <- DESeqDataSetFromTximport(
  txi=txi,
  colData = samples,
  design = ~ treatment) ## ? you can modify design (or condition)
#condition means what you are comparing (time_point, treatment, genotype, etc)
# in this case, I am choosing 'treatment' as the condition because I am comparing between the
healthy and virus treated samples.

#plot Dispersion Estimates
dds = estimateSizeFactors(dds)
dds = estimateDispersions(dds)
plotDispEsts(dds)

## 2.1. pre-filtering (keep '> 10 reads') ----
# removing rows with low gene counts---keeping rows that have at least 10 reads total
keep <- rowSums(counts(dds)) >= 10
dds <- dds[keep,]
dds
-------------------------------------------------
  ### Or !
  # it can be done in another way--- remove genes with low counts
  # keep genes that have minimum 1 CPM across 3 samples (since group has three replicates)
  # keepgenes <- rowSums(edgeR::cpm(dds)>1) >= 173 # fix according tor your row numbers
  # keepgenes
  # dds <- dds[keepgenes,]
  # dds

## 2.2. *** run deseq2 -----
dds <- DESeq(dds)
dds
# it will print like this:
## class: DESeqDataSet
## dim: 14599 7
## metadata(1): version
## assays(1): counts
## rownames(14599): FBgn0000003 FBgn0000008 ... FBgn0261574 FBgn0261575
## rowData names(0):
## colnames(7): treated1 treated2 ... untreated3 untreated4
## colData names(2): condition type

## 2.3. export results or vst counts ----
```

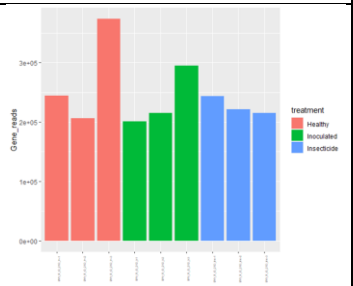
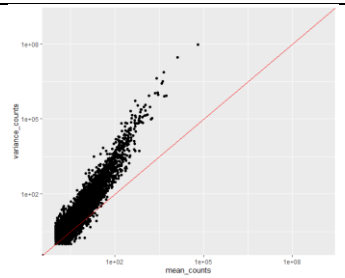
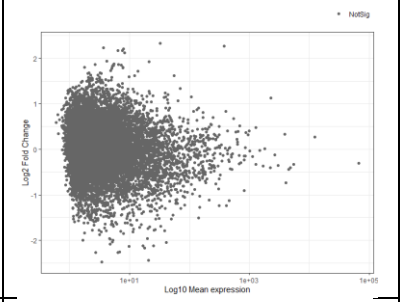
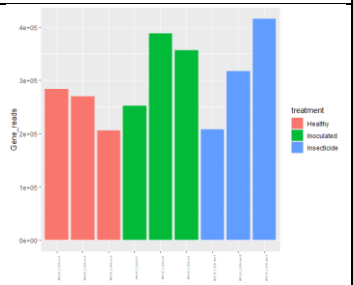
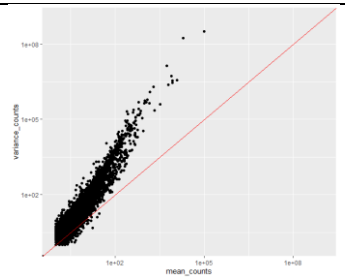
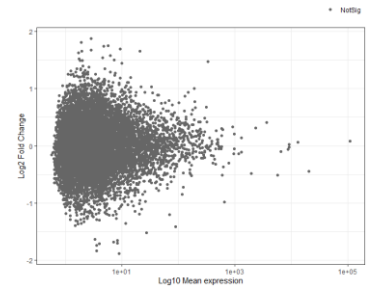
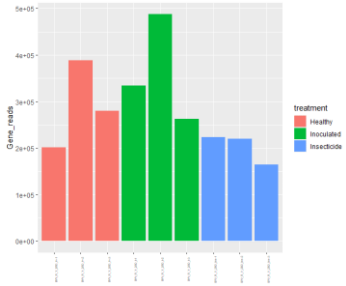
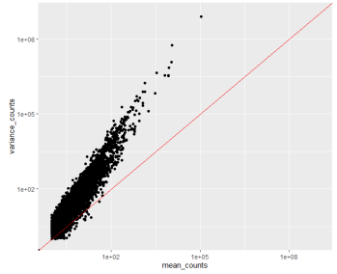
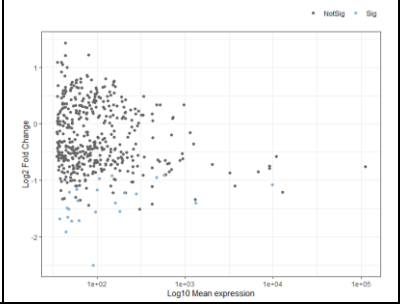
```
## vst means 'variance stabilizing transformations'. why vst is important? correction for high variance
cv <- as.data.frame(assay(varianceStabilizingTransformation(dds,blind=T)),check.names=F)
write.table(cv,"counts_vst_full.txt",sep="\t",dec=".",quote=FALSE)
saveRDS(cv,"counts_vst_full.Rds")

## 2.4. *** Up-regulated and down-regulated genes? ----
dg <- nbinomWaldTest(dds) # Wald test
# Why is Wald test important?
dg
print(resultsNames(dg))    #this is important

### 2.4.1 *** comparing across "groups" of samples
# It is also called as res by using 'contrast' option
res <- results((dg),alpha=0.05)
summary(res)
*** important results, it will say how many genes are up & down-regulated in inoculated samples for the above contrast

## 2.5. lfc shrink-log2fold change shrink ----
res1 <- lfcShrink(dg, contrast=c("treatment", "Healthy","Inoculated"), res=res, type="normal")
res1

## 2.6. convert table to data frame and save  ----
table_res <- as.data.frame(res1) # we try for 1st combination 'res1'
table_res$ensembl_gene_id <- rownames(table_res)
write.table(table_res,"res1",sep="\t",dec=".",quote=FALSE)

 # Step 3. PCA analysis -----
## 3.1. load vst count data ----
#table_vst <- readRDS("../5_dge/counts_vst_full.Rds")
table_vst <- cv

## 3.2. metadata as dataframe ----
table_meta <- as.data.frame(samples_updated)

## 3.3. match order of counts and metadata ----
mth <- match(colnames(table_vst),rownames(table_meta))
mth <- match(colnames(table_vst),rownames(samples_updated))
mth

# Note??: this next two line will take away your gene names from the table vst for rest of the analysis so you may avoid running next 2 lines
```

```
table_vst <- table_vst[,mth]
all.equal(rownames(table_meta),colnames(table_vst))

## 3.4. pca commands ----
pcaobj <- prcomp(x=t(table_vst))
pcs <- round(pcaobj$sdev^2/sum(pcaobj$sdev^2)*100,2)
pcs
pcamat1 <- as.data.frame(pcaobj$x)
pcamat2 <- merge(pcamat1,samples_updated,by=0)

## 3.5. PCA plots ----
p1 <- ggplot(pcamat2,aes(PC1,PC2,col=time_point,shape=time_point))+
  geom_point()+
  theme_bw()+
  stat_ellipse()+
  xlab(paste0("PC1: ",pcs[1],"% variance")) +
  ylab(paste0("PC2: ",pcs[2],"% variance")) + # https://www.biostars.org/p/472836/
  geom_text(aes(label=leaf_age),size=2,nudge_x=1,hjust="inward")+
  theme(legend.title=element_blank(),
        legend.text=element_text(size=6),
        legend.position="right",
        legend.justification="right")
p1

# to save the plot
ggsave("pca.png",p,height=12,width=12,units="cm",dpi=250)

# Step 4. MA plot -----

# load DESeq data and categorize gene expression as significant and non significant
table_res <- table_res[!is.na(table_res$padj),]
table_res$sig <- ifelse(table_res$padj<0.05,"Sig","NotSig")

p2 <- ggplot(table_res,aes(x=baseMean,y=log2FoldChange,colour=sig))+
  geom_point()+
  scale_x_log10()+
  scale_colour_manual(values=c("grey40","#80b1d3"))+
  labs(x="Log10 Mean expression",y="Log2 Fold Change")+
  theme_bw()+
  theme(legend.title=element_blank(),
        legend.position="top",
        legend.justification="right")
p2
```

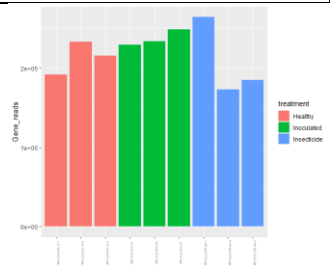
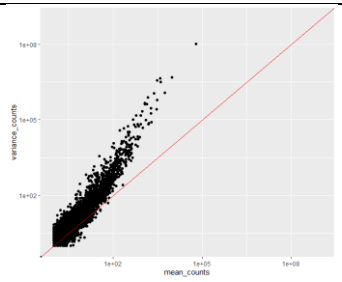
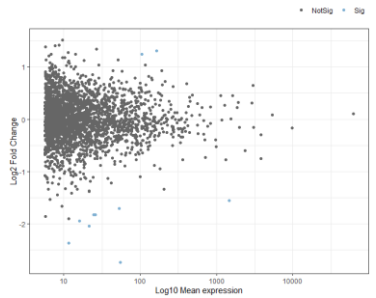
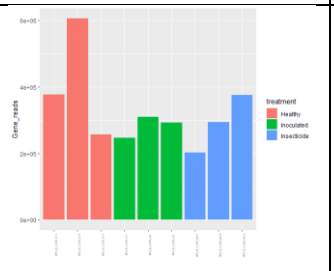
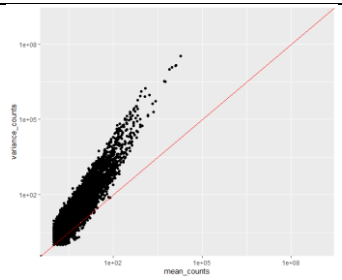
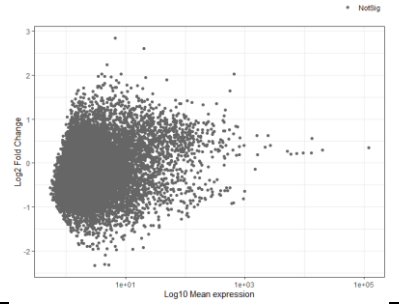
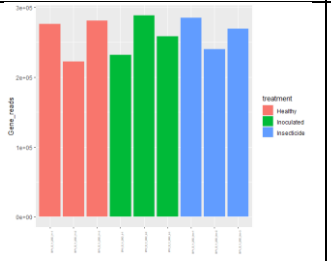
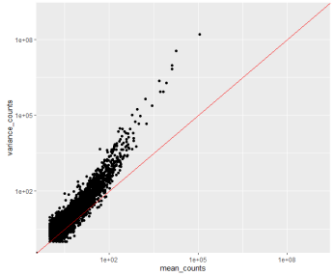
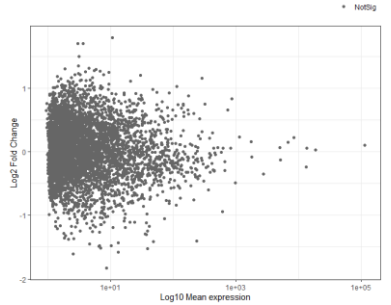
```
# save the plot as png file
ggsave("ma.png",p,height=12,width=12,units="cm",dpi=250)

# want to check where are you saving all these plots and results data??
getwd()

# Step 5. Volcano plot -----

p4 <- ggplot(table_res,aes(x=log2FoldChange,y=-log10(padj),colour=sig))+
  geom_point()+
  scale_colour_manual(values=c("grey40","#80b1d3"))+
  labs(x="Log2 Fold Change",y="-Log10 BH adjusted p-value")+
  theme_bw()+
  theme(legend.title=element_blank(),
        legend.position="top",
        legend.justification="right")
# BH = Benjamini Hochberg = FDR
p4
# to save the plot
ggsave("volcano.png",p,height=12,width=12,units="cm",dpi=250)

----------------------------------------------------

#  References. ----
#http://bioconductor.org/packages/release/bioc/vignettes/DESeq2/inst/doc/DESeq2.html#differential-expression-analysis
#https://hbctraining.github.io/DGE_workshop_salmon/lessons/01_DGE_setup_and_overview.html
# https://www.hadriengourle.com/tutorials/rna/#differential-expression-using-deseq2
```

Supplementary Tables 1-4 (from page 43)

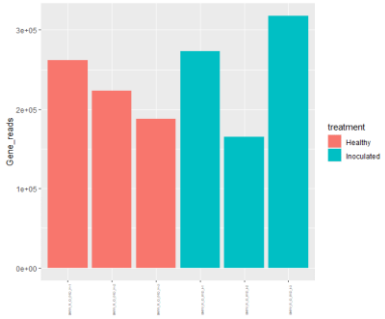
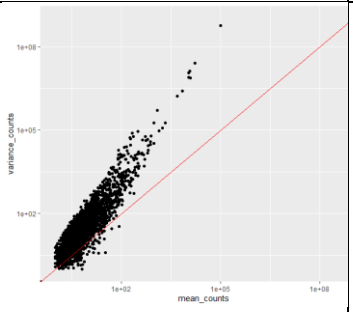
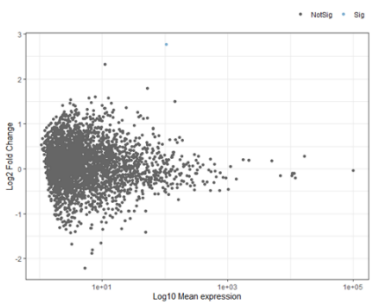
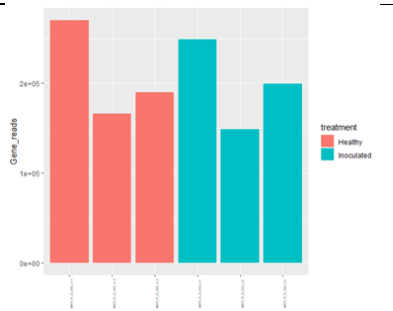
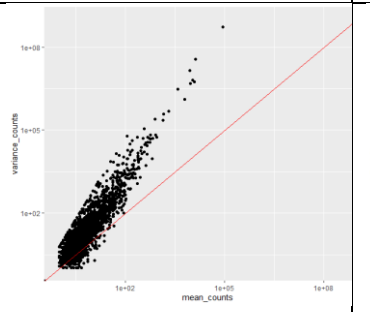
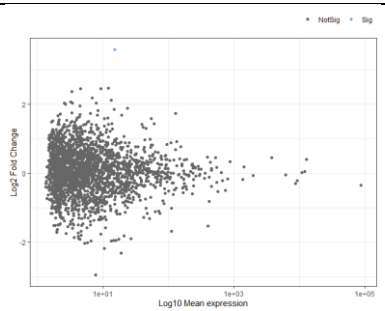
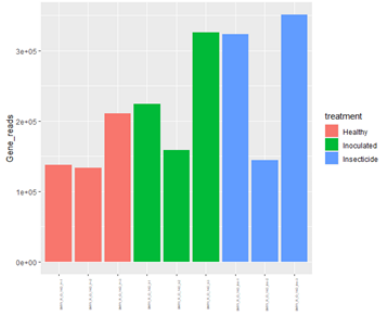
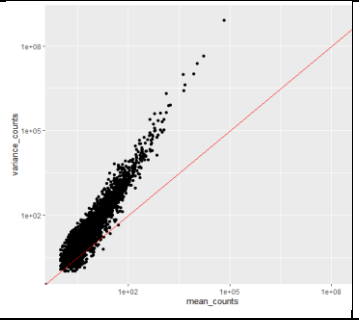
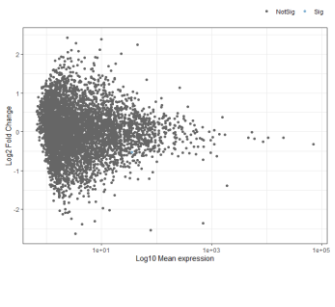
Supplementary Table 1: Differential expression analysis between resistant plant samples exposed to viruliferous or non-viruliferous aphids with BYV

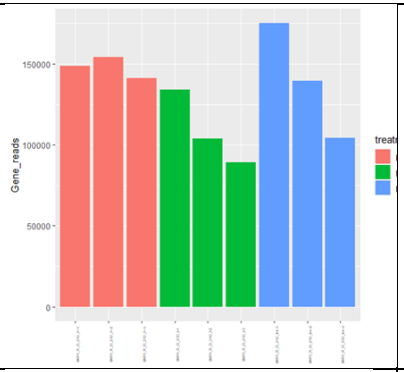
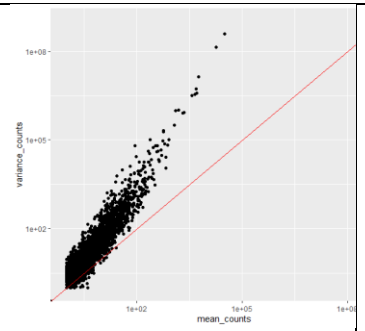
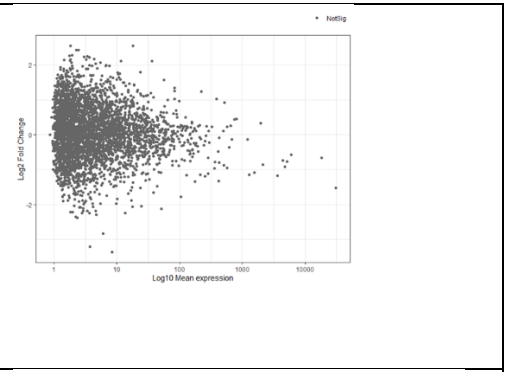
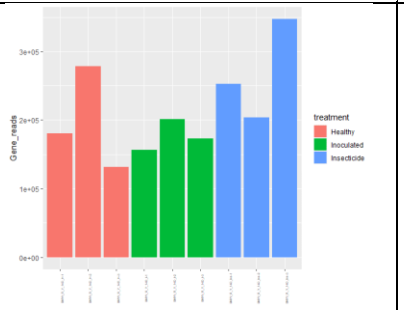
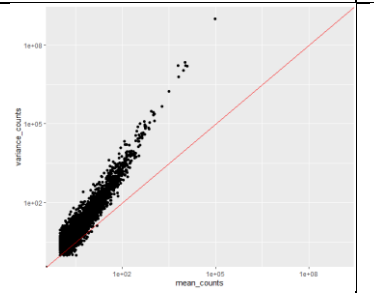
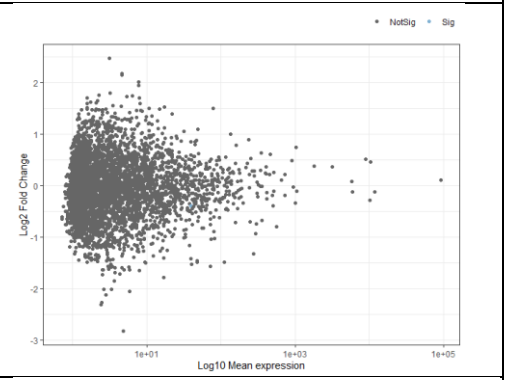
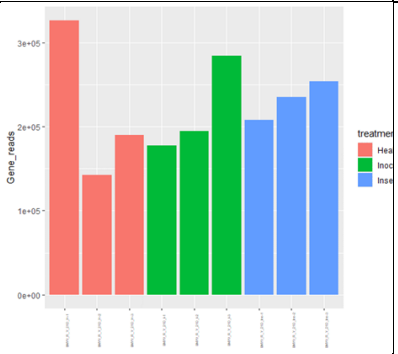
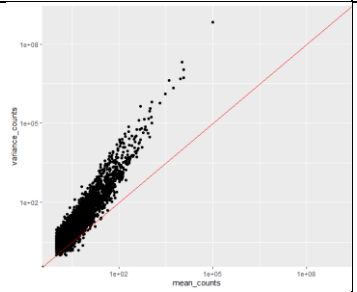
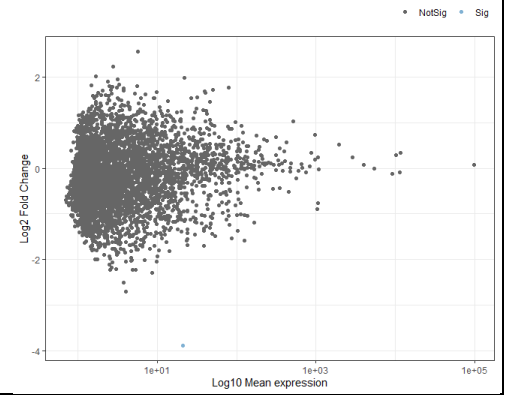
Leaf age	Time point	Number of genes not expressed	Gene reads per sample	Differentially expressed genes (DEGs)	Variance vs mean	MA plot
Old	21 days	38.53 % (10465) of 27161 genes are not expressed.		out of 9646 with nonzero total read count adjusted p-value < 0.05 LFC > 0 (up) : 0, LFC < 0 (down) : 0, 0 outliers [1] : 4, 0.041% low counts [2] : 0, 0% (mean count < 6)		
Young	21 days	38.86% percent (10554) of 27161 genes are not expressed		out of 9607 with nonzero total read count adjusted p-value < 0.05 LFC > 0 (up) : 0, 0% LFC < 0 (down) : 0, 0% outliers [1] : 5, 0.052% low counts [2] : 0, 0% (mean count < 1)		
Young	28 days	39.23% percent (10654) of 27161 genes are not expressed		out of 9145 with nonzero total read count adjusted p-value < 0.05 LFC > 0 (up) : 0, 0% LFC < 0 (down) : 23, 0.25% outliers [1] : 4, 0.044% low counts [2] : 8686, 95% (mean count < 35)		

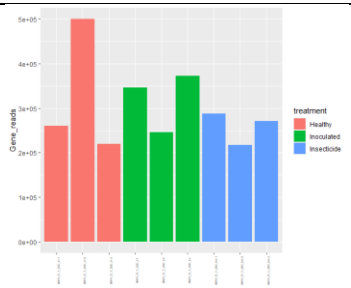
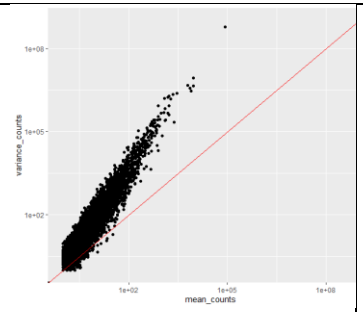
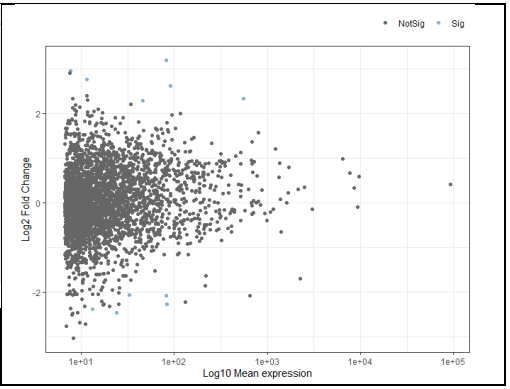
Supplementary Table 2: Differential expression analysis between susceptible plant samples exposed to viruliferous or non-viruliferous aphids with BYV

Leaf age	Time point	Number of genes not expressed	Gene reads per sample	Differentially expressed genes (DEGs)	Variance vs mean	MA plot
Old	21 days	40.63% percent (11036) of 27161 genes are not expressed		out of 8451 with nonzero total read count adjusted p-value < 0.05 LFC > 0 (up) : 8, 0.095% LFC < 0 (down) : 2, 0.024% outliers [1] : 4, 0.047% low counts [2] : 5898, 70% (mean count < 6)		
Young	21 days	40.35% percent (10959) of 27161 genes are not expressed		out of 9238 with nonzero total read count adjusted p-value < 0.05 LFC > 0 (up) : 0, 0% LFC < 0 (down) : 0, 0% outliers [1] : 1, 0.011% low counts [2] : 0, 0% (mean count < 1)		
Young	28 days	48.35% percent (13132) of 27161 genes are not expressed		out of 4966 with nonzero total read count adjusted p-value < 0.05 LFC > 0 (up) : 0, 0% LFC < 0 (down) : 0, 0% outliers [1] : 0, 0% low counts [2] : 0, 0% (mean count < 1)		

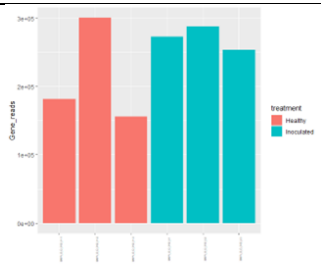
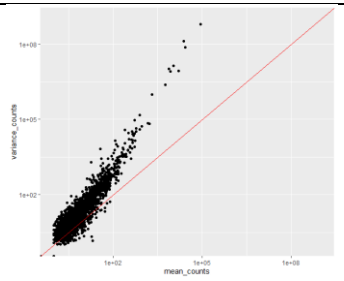
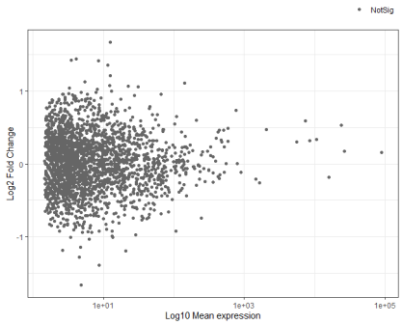
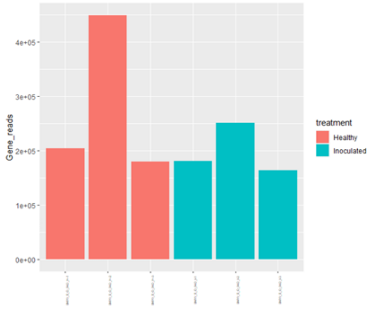
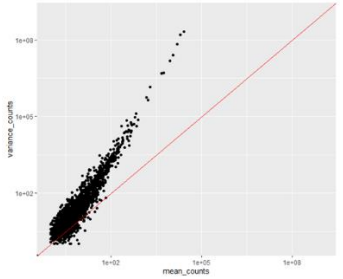
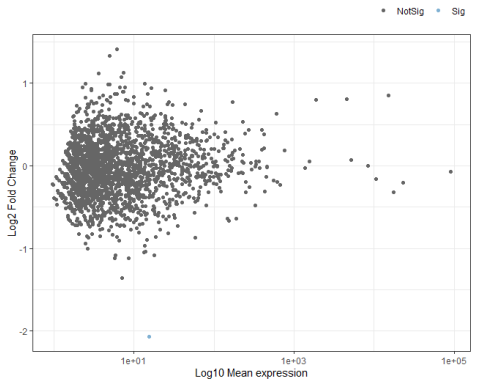
Supplementary Table 3: Differential expression analysis between resistant plant samples exposed to viruliferous or non-viruliferous aphids with BMYV

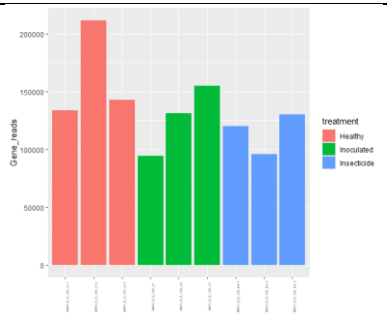
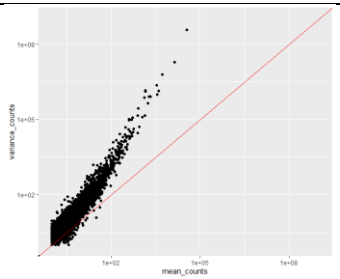
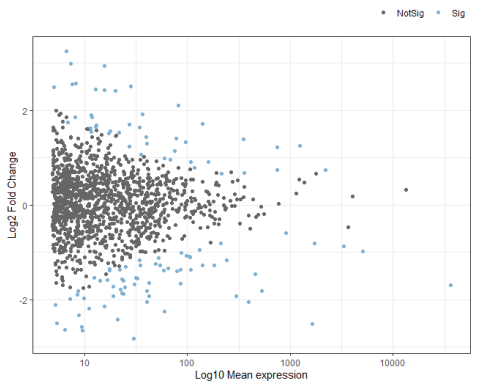
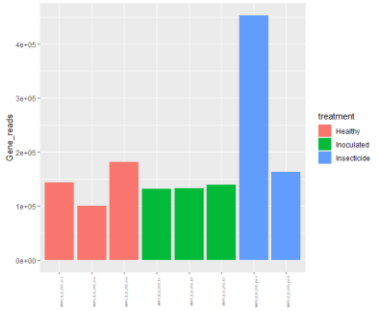
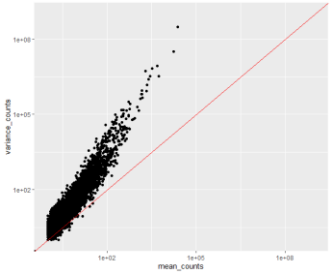
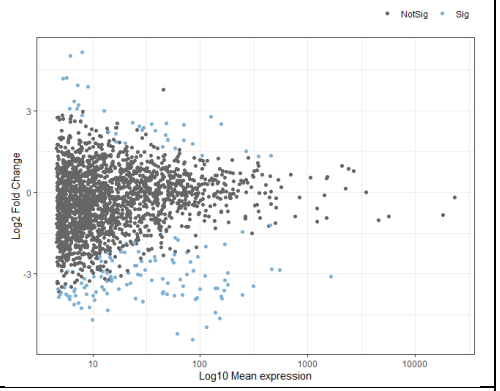
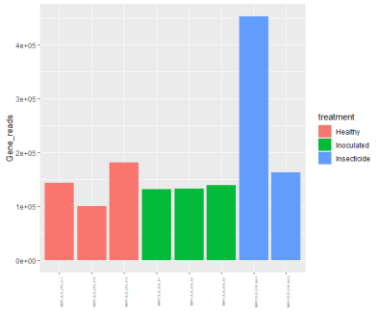
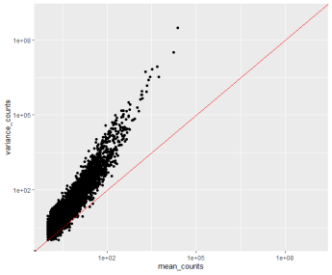
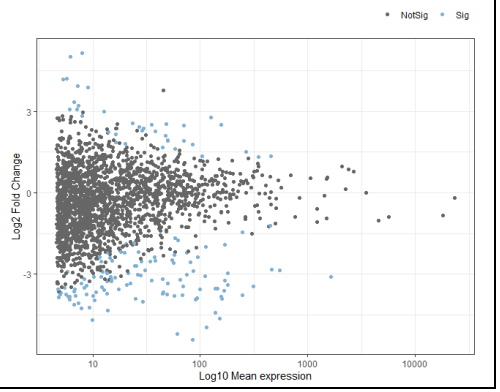
Leaf age	Time point	Number of genes not expressed	Gene reads per sample	Differentially expressed genes (DEGs)	Variance vs mean	MA plot
Old	1 day	59.38% percent (16128) of 27161 genes are not expressed		"out of 3106 with nonzero total read count adjusted p-value < 0.05 LFC > 0 (up) : 1, 0.032% LFC < 0 (down) : 0, 0% outliers [1] : 3, 0.097% low counts [2] : 0, 0% (mean count < 1)"		
Old	4 days	66.41% percent (18037) of 27161 genes are not expressed		out of 2444 with nonzero total read count adjusted p-value < 0.05 LFC > 0 (up) : 1, 0.041% LFC < 0 (down) : 0, 0% outliers [1] : 12, 0.49% low counts [2] : 0, 0% (mean count < 1)		
Old	14 days	52.41% percent (14234) of 27161 genes are not expressed		out of 4831 with nonzero total read count adjusted p-value < 0.05 LFC > 0 (up) : 0, 0% LFC < 0 (down) : 1, 0.021% outliers [1] : 6, 0.12% low counts [2] : 0, 0% (mean count < 1)		

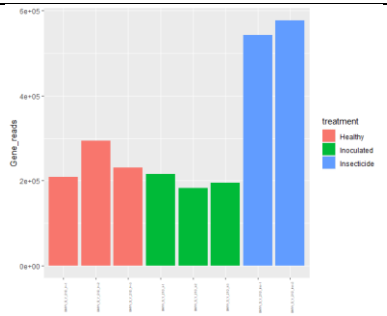
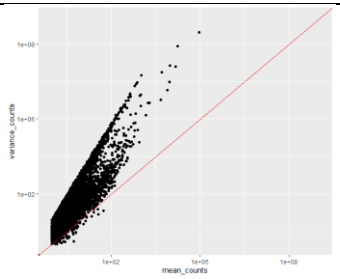
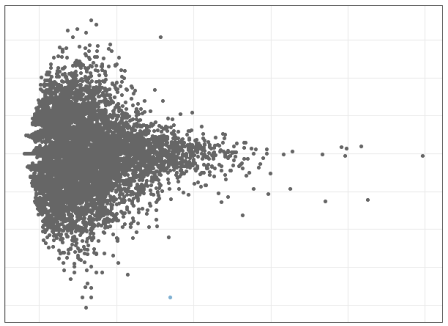
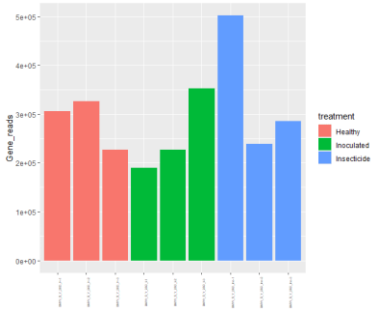
Old	21 days	56.78% percent (15421) of 27161 genes are not expressed		out of 3655 with nonzero total read count adjusted p-value < 0.05 LFC > 0 (up) : 0, 0% LFC < 0 (down) : 0, 0% outliers [1] : 10, 0.27% low counts [2] : 0, 0% (mean count < 1)		
Young	14 days	56.64% percent (15384) of 27161 genes are not expressed		out of 3852 with nonzero total read count adjusted p-value < 0.05 LFC > 0 (up) : 0, 0% LFC < 0 (down) : 1, 0.026% outliers [1] : 1, 0.026% low counts [2] : 0, 0% (mean count < 1)		
Young	21 days	53.41% percent (14507) of 27161 genes are not expressed		out of 4118 with nonzero total read count adjusted p-value < 0.05 LFC > 0 (up) : 0, 0% LFC < 0 (down) : 1, 0.024% outliers [1] : 5, 0.12% low counts [2] : 0, 0% (mean count < 1)		

Young	28 days	35.52% percent (9647) of 27161 genes are not expressed		out of 11293 with nonzero total read count adjusted p-value < 0.05 LFC > 0 (up) : 6, 0.053% LFC < 0 (down) : 5, 0.044% outliers [1] : 0, 0% low counts [2] : 8101, 72% (mean count < 7)		

Supplementary Table 4: Differential expression analysis between susceptible plant samples exposed to viruliferous or non-viruliferous aphids with BMYV

Leaf age	Time point	Number of genes not expressed	Gene reads per sample	Differentially expressed genes (DEGs)	Variance vs mean	MA plot
Old	1 day	66.21% percent (17984) of 27161 genes are not expressed		out of 2502 with nonzero total read count adjusted p-value < 0.05 LFC > 0 (up) : 0, 0% LFC < 0 (down) : 0, 0% outliers [1] : 2, 0.08% low counts [2] : 0, 0% (mean count < 1)		
Old	4 days	68.5% percent (18606) of 27161 genes are not expressed		out of 2268 with nonzero total read count adjusted p-value < 0.05 LFC > 0 (up) : 0, 0% LFC < 0 (down) : 1, 0.044% outliers [1] : 0, 0% low counts [2] : 0, 0% (mean count < 1)		

Old	14 days	56.3% percent (15291) of 27161 genes are not expressed		out of 3885 with nonzero total read count adjusted p-value < 0.05 LFC > 0 (up) : 46, 1.2% LFC < 0 (down) : 68, 1.8% outliers [1] : 1, 0.026% low counts [2] : 2485, 64% (mean count < 5)		
Old	21 days	45.14% percent (12260) of 27161 genes are not expressed		out of 6605 with nonzero total read count adjusted p-value < 0.05 LFC > 0 (up) : 38, 0.58% LFC < 0 (down) : 114, 1.7% outliers [1] : 8, 0.12% low counts [2] : 4482, 68% (mean count < 5)		
Young	14 days	45.14% percent (12260) of 27161 genes are not expressed		out of 6605 with nonzero total read count adjusted p-value < 0.05 LFC > 0 (up) : 38, 0.58% LFC < 0 (down) : 114, 1.7% outliers [1] : 8, 0.12% low counts [2] : 4482, 68% (mean count < 5)		

Young	21 days	40.71% percent (11056) of 27161 genes are not expressed		out of 9088 with nonzero total read count adjusted p-value < 0.05 LFC > 0 (up) : 0, 0% LFC < 0 (down) : 1, 0.011% outliers [1] : 82, 0.9% low counts [2] : 0, 0% (mean count < 1)		
Young	28 days	38.32% percent (10409) of 27161 genes are not expressed		out of 10707 with nonzero total read count adjusted p-value < 0.05 LFC > 0 (up) : 1, 0.0093% LFC < 0 (down) : 0, 0% outliers [1] : 3, 0.028% low counts [2] : 0, 0% (mean count < 0)