NKFIH-K 119701

FINAL REPORT

Identification of plant genes responsible for viral symptom development

György Szittya PhD

National Agricultural Research and Innovation Centre
Agricultural Biotechnology Institute
Institute of Plant Biotechnology

16.12.2020

Abbreviations:

AMV: alfalfa mosaic virus
AOP2: Alkenyl Hydroxalkyl Producing 2
A. thaliana: Arabidopsis thaliana
B. ricinifolia: Begonia ricinifolia
BRSV: Beet ringspot virus
CMV: Cucumber mosaic virus
CymRSV: Cymbidium ringspot virus
DCL: Dicer-like
DE: differential expression
DFL2: Dwarf in Light 2
DMR6: Downy Mildew Resistant 6
dpi: days post inoculation
dsRNA: double-stranded RNA
FBL: F-box/LRR protein
kDa: kiloDalton
miRNA: microRNA
mRNA: messenger RNA
M. truncatula: Medicago truncatula
N. benthamiana: Nicotiana benthamiana
NB-LRR: immune receptor genes of the nucleotide binding and leucine-rich repeat class
nm: nanometer
nt: nucleotide
ORF: open reading frame
pha-siRNA: phased siRNA
RIL: Recombinant Inbred Line
RISC: RNA-induced silencing complex
RMV: Ribgrass Mosaic Virus
SAM: shoot apical meristem
siRNA: short interfering RNA
SIZ1: SUMO E3 ligase
S. lycopersicum: Solanum lycopersicum
sRNA: small RNA
ta-siRNA: trans-acting siRNA
TuMV: Turnip Mosaic Virus
TVCV: Turnip vein clearing virus
UTR: untranslated regions
vasiRNA: virus activated, host derived siRNAs
vsiRNA: virus-derived short interfering RNA
WT: wild type
WUS: Wuschel

Summary

Plant viruses cause a variety of diseases in susceptible hosts and therefore can possess serious threats on agricultural crop production. Virus infections are often systemic and cause damaging symptoms. The symptoms of such infections have been the subject of many descriptive and (fewer) mechanistic studies, mainly focusing on identifying viral symptom determinants. However, host determinants influencing symptom development are mostly unknown, and the mechanisms underlying these processes are still poorly understood. During this project we identified and characterized several plant viruses infecting ornamental (Beet ringspot virus) and crop plants (Alfalfa mosaic virus, Turnip vein clearing virus), described their effect on the gene expression of their host and also identified host gene candidates that could influence of TVCV and TuMV induced symptoms.

During this research project we used two *Arabidopsis thaliana* ectypes (Col-0 and Bur-0) to reveal those molecular factors which contribute to symptom formation during TVCV infections. TVCV causing a very mild symptom on the Col-0 ecotype and severe leaf curling and deformation on Bur-0. These *Arabidopsis* ecotypes could be considered as a molecular biological tool kit for viral infection and symptom development studies as they bearing natural genetic variations in homozygous form, caused by the geographical isolation. We believe that the observed differences in the symptoms between the two *Arabidopsis* ecotypes are based on genetic differences that manifest on gene expression level. We used high throughput sequencing technologies to analyse changes in the expression of the whole transcriptome of mRNAs, sRNAs and their cleaved mRNA targets (degradome) between mock and virus infected ecotypes. Comparison of the mRNA, sRNA and degradome profiles of two genetically divergent, virus infected *Arabidopsis* ecotypes exposed significant differential expression patterns and we identified several candidate host genes as the major host determinants responsible for the distinct symptom development phenotypes caused by TVCV on the Bur-0 ecotype. To determine if the candidate genes obtained from the genome wide expressional studies play a general role in symptom development, we used genetic mapping, using Recombinant Inbred Lines (RIL) of the two ecotypes. The correlation between genotypes and TVCV caused symptoms observed on the infected RILs was assessed using the Windows QTL Cartographer software. A single marker association test revealed that four consecutive markers (c4_02133, c4_03833, c4_04877, and c4_05629) on the top of chromosome 4 were linked to the viral leaf deformation symptom with a P-value of 0.007. These markers represent a 3.5 Mb region on chromosome 4. Combining the results of the candidate genes that are differentially expressed only in the virus-infected Bur-0 ecotype samples and fall within the mapping region of the virus infected RIL lines resulted the list of most likely symptom determinant host genes (AOP2, Small GTPase-like protein, DFL, FBL and FBL23). To validate these results, we deleted the mentioned candidate genes using the CRISPR/Cas9 system in the Bur-0 ecotypes.

The results of this research were published (9 peer reviewed article and 1 book chapter) and continuously presented on international and national conferences (4 presentations). Preparation of a manuscript and a doctoral dissertation based on the results on the identification of plant genes responsible for viral symptom development is still expected. The results and figures shown in this report outline the high quality of this study, which is anticipated to be published in an academic peer reviewed journal. The results of this work are expected to be a highly cited reference for future studies.

Content:

1. Brief Introduction

Virus-infected plants may develop various symptoms on all or some of their parts. The symptoms may be either nonspecific or quite characteristic of the particular virus. Common symptoms are dwarfing, stunting, various leaf deformations, chlorotic mosaics, vein-clearing, etc. These effects may be either severe or hardly detectable. Many viruses may infect certain hosts without causing visible symptoms. Such viruses are called latent viruses and the hosts are called symptomless carriers. The development of symptoms usually depends on the environmental conditions, like temperature and light. Finally, plants may show acute or severe symptoms soon after inoculation that may lead to death of the host. If the host survives the initial shock, the symptoms tend to become milder in the subsequently developing part of the plant, leading to partial or total recovery. On the other hand, symptoms may progressively increase in severity and may result in gradual decline of the plant (Hull 2014). Virus infection and symptom development is a multistep process. First, the virus must invade the plant cell either by mechanical means or by a vector. After uncoating, inside the host cytoplasm the viral RNA is translated by the host translational machinery and the viral encoded proteins become active. Movement proteins help the viral nucleic acid move into another cell through plasmodesmata (Carrington et al. 1996). At the site of entry, plants develop a hypersensitive response which involves the overproduction of reactive oxygen species and programmed cell death which prevents further spreading of the virus into the neighbouring cells. This manifests in local tissue lesions (local symptoms). If the virus manages to overcome this barrier and reaches the veins, it can spread into other tissues or organs causing systemic symptoms, i.e. leaf deformations, mosaics, yellowing, ringspots, mottling, stunting, necrosis. In compatible interactions between plants and viruses that result in systemic infection, symptom development is a major phenotypic trait. However, host determinants governing this trait are little known, and the mechanisms underlying it are still poorly understood.

To restrict viral accumulation, plants and other eukaryotes utilize a highly adaptive, sequence-specific mechanism called RNA silencing (Ding and Voinnet 2007). Upon virus infection, the viral nucleic acid (i.e. positive single stranded RNA) gets into the host cytoplasm, where it functions as a genuine mRNA. Double-stranded hairpin-like structures of this RNA molecule or the double-stranded RNA (dsRNA) intermedier of viral replication trigger antiviral silencing (Molnár et al. 2005). The dsRNA is randomly cut into 21–22 bp-long fragments called virus-derived short interfering RNAs (vsiRNAs) by specific Dicer-like (DCL) endonucleases (Deleris et al. 2006). One of the strands of these fragments called the guide strand is then loaded into an RNA-induced silencing complex (RISC). By the help of the guide RNA, the RISC binds to the complementary single stranded viral RNA molecules and either catalytically cleaves them or blocks their translation. In the first case, the cleavage products are completely degraded by exonucleases purging the viral RNA from the cytoplasm. For an efficient virus infection, viruses should be able to evade the silencing machinery. In the evolutionary arms race, viruses developed a plethora of different silencing suppressor molecules that can interfere with various steps of the silencing pathway, including small RNA production, processing, stability and activity of the RISC (Voinnet 2005; Burgyán and Havelda 2011; Csorba, Kontra, and Burgyán 2015).

RNA silencing functions not only as an antiviral defence system but also control gene expression. Endogenous small RNAs (sRNA) are an important gene expression regulator

involved in many developmental processes and also stress responses. In plants sRNAs are very diverse and based on their biogenesis we distinguish two main classes: microRNAs (miRNAs) and siRNAs. MiRNAs function in post-transcriptional manner by down regulating target mRNAs in a variety of cellular processes (Bartel 2004; Jones-Rhoades et al., 2006; Mallory and Vaucheret, 2006) while siRNAs mainly responsible for maintaining genome integrity (Chen 2009). Since VSRs interfere with RNA silencing pathways it is not surprising that in several cases they were shown to be viral symptoms determinants (Szittya and Burgyán, 2013). In addition, virus derived siRNAs (vsiRNAs) incorporate into RISC and they provide sequence specificity to the antiviral defence system of plants. These vsiRNAs, however, can find complementary target sequences also among the host mRNAs triggering their silencing and causing viral symptoms (Shimura et al., 2011; Smiths et al., 2011). Since the interference with the host gene expression highly depends on the degree of sequence complementarity, natural selection for less complementary host sequences can attenuate the severity of symptoms. Furthermore, the production of virus activated, host derived siRNAs (vasiRNAs) (Cao et al., 2014) may also present a conserved host response to virus infection by silencing of host genes.

Since viral infections often lead to severe diseases in plants that result in significant economic loss, we have started to study symptom development in two *Arabidopsis thaliana* ecotypes (Clo-0 and Bur-0) as a model system. *Arabidopsis* ecotypes are genetically distinct geographic variants in which genetic heterogeneity is greatly reduced as a result of inbreeding. This makes them particularly useful for comparative genetic analysis and genetic mapping. We used two viruses (TVCV and TuMV) that showed different symptoms on the selected *Arabidopsis* ecotypes for further studies. Some part of the results written here are not published yet, therefore we give a detailed description of our results here.

2. **Molecular characterization of a beet ringspot nepovirus causing light dependent symptoms on *Begonia ricinifolia***

As a preliminary research of this NKFIH project, we characterized a nepovirus which was isolated from *Begonia ricinifolia* showing chlorotic ringspot and line pattern symptoms (Fig 1). The purified virus had spherical particles of ca. 30 nm and contained a single coat protein subunit of ca. 56 kDa. The complete nucleotide sequence of the bipartite viral genome was determined. RNA 1 is 7394 nucleotides long, flanked by 5’ and 3’ untranslated regions (UTR), and followed by a 3’ poly-A tail. It contains a single 6810 nt long open reading frame (ORF), which is translated into a 255 kDa polyprotein composed of 2269 amino acids. The 4684 nt long RNA 2 has a 4053 nt long ORF which encodes a single polyprotein of 1350 amino acids with a molecular weight of 149 kDa (Fig 2). Sequence comparisons revealed that the virus isolated from *B. ricinifolia* has the highest sequence similarity to beet ringspot virus and should be considered as a strain of BRSV. This was the first report on the occurrence of BRSV in *B. ricinifolia* and the presence of this virus outside Scotland (Kis et al., 2017).

Fig. 1. Ringspot and line pattern symptoms caused by BRSV.

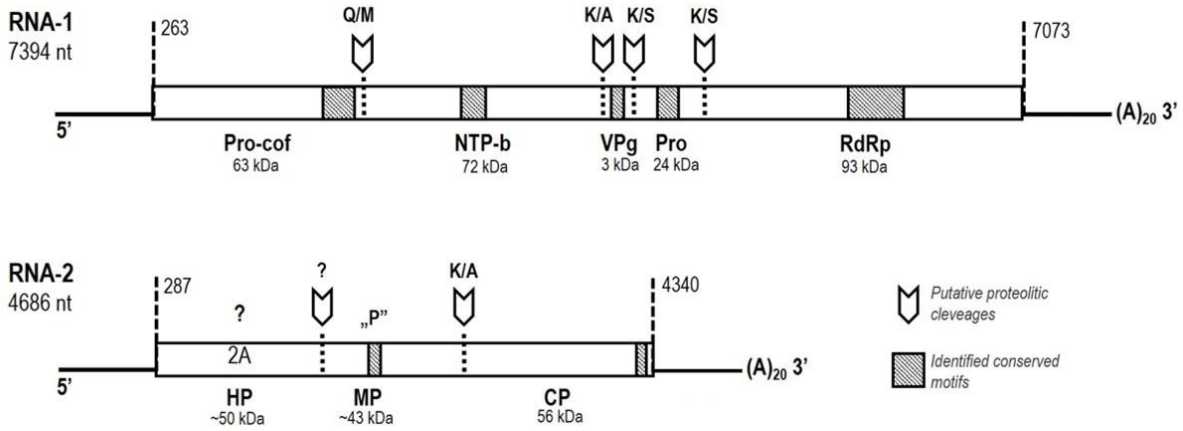

Fig. 2. Schematic representation of the genome organization of BRSV.

3. Virus infection modulates the expression level of NB-LRR regulating miRNAs

During this NKFIH project we tested the effect of virus infection on the host plant "immune system". Plant genomes contain large numbers of NB-LRR type resistance genes that recognize specific pathogen effectors and trigger resistance responses (Meyers et al., 2003) and as a consequence protecting plants from diseases caused by devastating pathogens. However, the unregulated expression of NB-LRR genes can inhibit growth and may result in autoimmunity in the absence of pathogen infection (Weaver et al., 2006). It was shown previously that a subset of plant miRNAs (miR482 family) can target NB-LRR genes and regulate plant immunity (Li et al., 2012; Shivaprasad et al., 2012; Zhai et al., 2011). It was proposed that in the absence of pathogens the NB-LRR resistance genes are transcribed, but cleaved by miRNAs, keeping their translation at a very low level. During pathogen attack the level of miRNA are reduced and this should result in increased level of resistance gene transcript (Shivaprasad et al., 2012). To test the effect of virus infection on the expression level of NB-LRR-regulating miRNAs (miR1507, miR2109, and miR2118 – member of the miR482 family) in legumes, we infected *M. truncatula* plants with AMV (Salamon et al., 2018). The level of the NB-LRR-targeting miRNAs was monitored in the inoculated cotyledons of mock and virus-infected plants with northern blots 4 dpi. We found that similarly to previous reports in Solanaceae species (Li et al., 2012; Shivaprasad et al., 2012), the expression level of these three NB-LRR regulating miRNAs decreased during AMV infection in *M. truncatula* cotyledons and the largest reduction was detected in the expression of miR2118. This result showed that during virus infection the modulation of the NB-LRR regulating endogenous miRNA level could result in higher NB-LRR level and increase defence reaction of the virus infected plants against pathogen attack (Sós-Hegedűs et al., 2020).

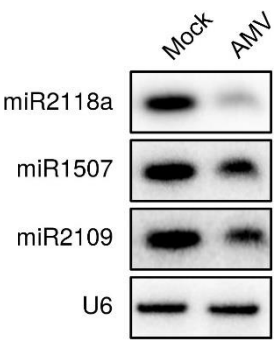

Fig. 3. Expression of the NB-LRR regulating miRNAs in *Medicago truncatula* roots during pathogen attack.

4. Transcriptome reprogramming in the shoot apical meristem of CymRSV-infected *Nicotiana benthamiana* plants

During this project we collaborated with József Burgyán's lab to characterise symptom recovery in virus infected plants. It has been observed in some plant-virus interaction under certain circumstances that the plants show a sign of healing from virus-infection. However, the nature of molecular processes behind symptom recovery has been an enigma in plant virology for a long time. Discovery of RNA silencing provided a possible mechanism to explain some aspects of recovery but not all. We provided evidence that RNA silencing is not the reason for recovery in CymRSV-infected *Nicotiana benthamiana* plants. Transcriptome analysis followed by *in situ* hybridization experiments shed light on the expressional changes in the shoot apical meristem (SAM) region upon virus infection. We observed a severe loss of meristem function in the SAM which was accompanied by the downregulation of meristem-specific genes including WUSCHEL (WUS). However, WUS was not downregulated in the SAM of plants infected with TVCV and CMV suggesting that the loss of meristem function is not the reason for shoot necrosis (Fig 4).

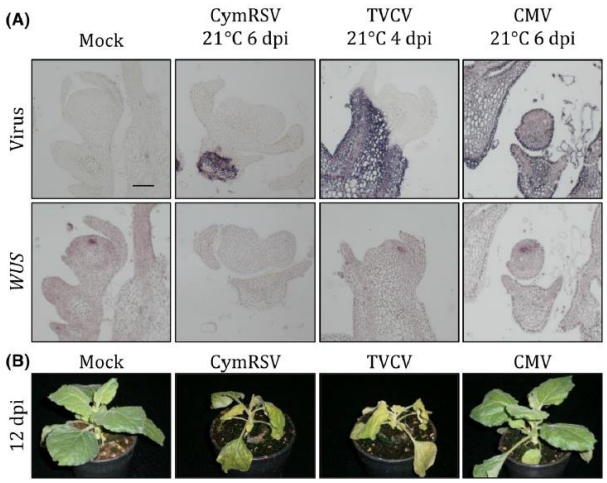

Fig. 4. WUS expression and viral symptoms of plants infected with three different viruses. (A) *In situ* hybridization of consecutive longitudinal sections of mock-, CymRSV-, TVCV- and CMV-infected shoot apical meristems hybridized with probes to detect the virus or WUSCHEL (WUS). Bar = 100 μm. (B) Symptoms of mock-, CymRSV-, TVCV- or CMV-infected *Nicotiana benthamiana* plants 12 days after infection.

Among the upregulated genes we found an enrichment of salicylate-metabolism and signalling (i.e. DMR6, SIZ1, and others), and autophagy-related functions (among many others) suggesting that these genes might play a role in shoot necrosis and symptom recovery (Medzihradszky et al., 2019).

5. Novel functions in the auxin signalling network revealed by comparative analysis of Turnip vein-clearing virus infected *Arabidopsis thaliana* ecotypes

Symptom development in virus-infected plants depends on both viral-encoded and host-encoded factors. The viral-encoded symptom determinants have been extensively studied in the last decades, however host genes determining symptom development are largely unknown, and the mechanisms underlying their actions are still poorly understood. To design new strategies to alleviate virus symptoms, it is necessary to understand the host-pathogen interactions at molecular levels. Therefore, we used two *Arabidopsis thaliana* ectypes (Col-0 and Bur-0) to reveal those molecular factors which contribute to symptom formation during RMV infections. These *Arabidopsis* ecotypes could be considered as a molecular biological tool kit for viral infection and symptom development studies as they bearing natural genetic variations in homozygous form, caused by the geographical isolation. We think that the observed differences

in the symptoms between the two *Arabidopsis* ecotypes are based on genetic differences that manifest on gene expression level. We used high throughput sequencing technologies to analyse changes in the expression of the whole transcriptome of mRNAs, sRNAs and their cleaved mRNA targets (degradome). To determine if the candidate genes obtained from the experiments play a general role in antiviral defence, we used genetic mapping by using Recombinant Inbred Lines (RIL) of the two ecotypes.

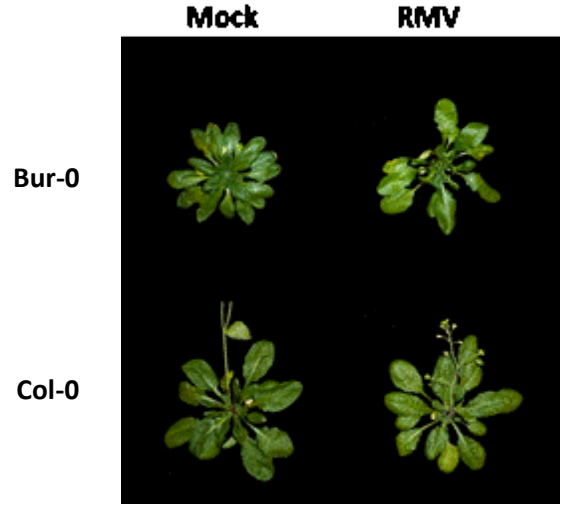

Fig. 5. Viral symptom development in Arabidopsis thaliana ecotypes. Plants were infected with RMV or mock inoculated and picture was taken 3 weeks after inoculation.

5.1. Molecular characterization and in vitro synthesis of infectious RNA of Turnip vein-clearing virus

We started the project with a virus which was isolated from *Alliaria petiolata* in Hungary. Host range experiments revealed a high similarity of this isolate to ribgrass mosaic viruses therefore we referred to this virus as RMV. To further characterize our RMV virus isolate that causes different phenotypes on the *Arabidopsis* ecotypes we cloned and determined the complete nucleotide sequence of the viral genome. The genomic RNA is composed of 6312 nucleotides and contains four open reading frames (ORF). ORF1 is 3324 nt-long and encodes a polypeptide of about 125.3 kDa. The ORF1 encoded putative replication protein contains an Alphavirus-like methyltransferase

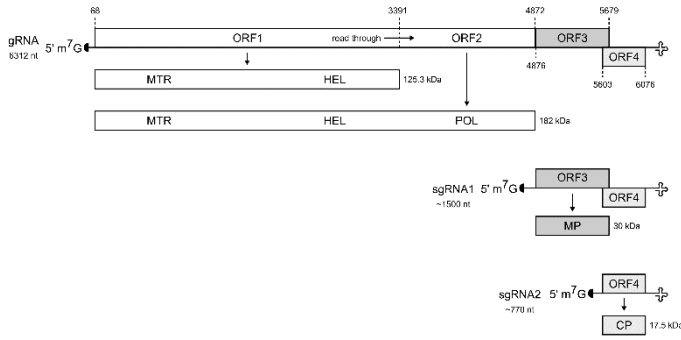

Fig. 6. Schematic representation of the genome organization of TVCV.

domain. ORF2 is 4806 nt-long and encodes a polypeptide of about 182 kDa. The ORF2 encoded putative replication protein contains an RNA-dependent RNA polymerase, catalytic domain. ORF3 encodes the putative cell-to-cell movement protein with a molecular weight of 30.1 kDa. ORF4 overlaps with ORF3 and encodes the coat protein with a size of 17.5 kDa (Fig 6). Sequence comparisons revealed that the RMV isolate has the highest similarity to turnip vein-clearing viruses and should be considered an isolate of Turnip vein-clearing virus (TVCV). This was the first report on the occurrence of TVCV in Hungary. *In vitro* transcripts prepared from the full-length cDNA clone of TVCV were highly infectious and induced typical symptoms characteristic to the original isolate of the virus. Since infectious clones of TVCV and crTMV (another isolate of TVCV) markedly differed in respect to recovery phenotype in *Arabidopsis thaliana*, it is feasible to carry out gene exchange or mutational studies to determine viral factors responsible for the symptom recovery phenotype (Tóth et al., 2019).

5.2. Comparing transcriptomic, sRNA-omic and degradome profiles of mock- and virus infected two *Arabidopsis* ecotypes.

Upon viral infection, the transcriptome of the infected host changes significantly (Pitzalis et al., 2020). The virus modulates the host environment in favour of its replication and spread, while the host tries to restrict the intruder and initiates an immune response. We expected that the two ecotypes would respond to the TVCV infection similarly in many ways. However, since there were significant differences in the observable symptoms, we anticipated that there would be gene expressional changes that reflect these differences. In order to reveal these differential gene expression profiles, we performed an RNA-seq experiment. We collected leaves from mock- and TVCV infected Col-0 and Bur-0 plants in four biological replicates and purified total RNA from them. For the RNA-seq analysis, the polyA-tail mRNA fraction was purified and stranded sequencing libraries were prepared. The libraries were sequenced with a 2×150 bp paired-end chemistry on an Illumina HiSeq4000 platform. Raw sequences were quality filtered and the adaptors were trimmed by the sequencing facility. The clean reads were mapped to the *Arabidopsis thaliana* reference genome TAIR10 with HiSat2 using the AtRTDv2 reference transcriptome to predict novel transcripts. The amended reference transcriptome was used for differential expression analysis with kallisto and sleuth (Pimentel et al, 2017). This suit can assess and remove technical variance from total variance and this way it can better estimate the actual biological variance. During the differential expression analysis, we considered genes differentially expressed between the mock and virus-infected samples if the Q-value from the statistical test was lower than 0.05 (5% false discovery rate), the absolute value of the effect (b parameter) was higher than log(2), and the mean expression (mean_obs) was at least 10.

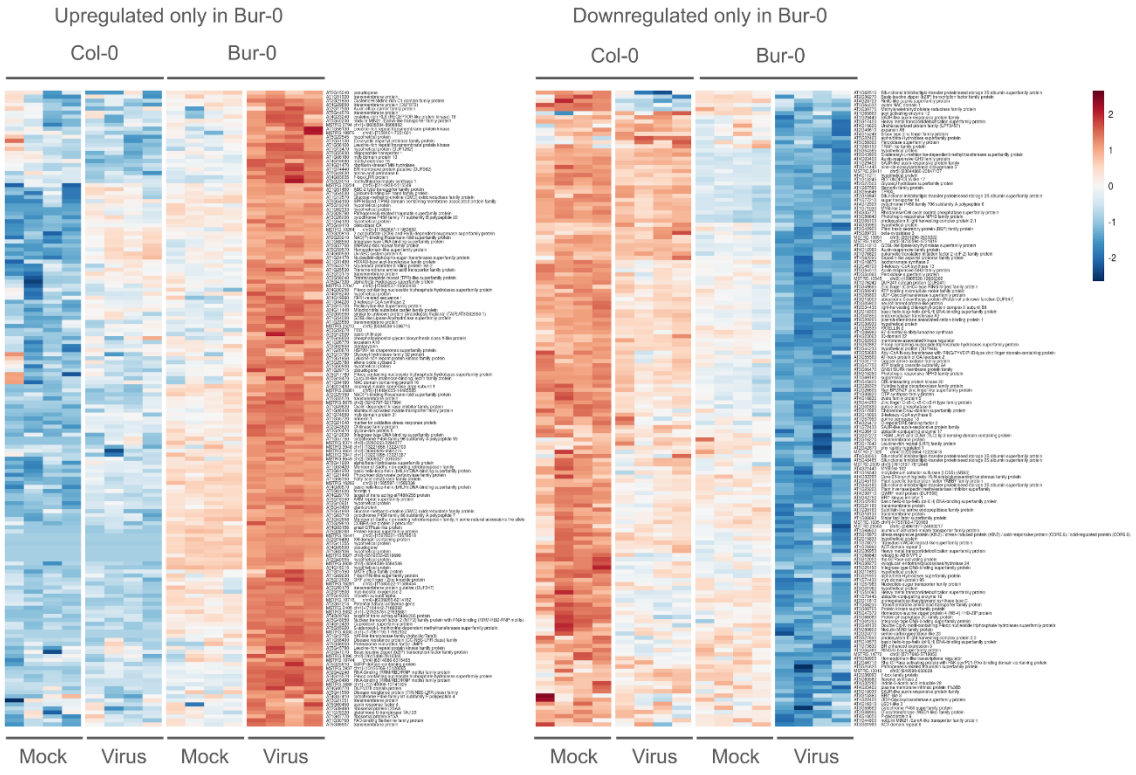

Fig. 7. Differentially expressed mRNAs in TVCV infected plants.

In this way, we got a list of the differentially expressed genes for the Col-0 and Bur-0 samples. We reasoned that the observed symptoms must correlate (or anti-correlate) with the gene expressional changes in the two ecotypes. Since we could observe the leaf polarity deformation only in the Bur ecotype, we expected gene expressional changes in the Bur ecotype only. Therefore, we looked for differentially expressed genes either that upregulated only in the Bur virus while the Bur mock values were equal or higher than in the Col mock and Col virus samples, or that downregulated only in the Bur virus while the Bur mock values were equal or lower than in the Col mock and Col virus samples (Fig 7). We got 152 and 156 genes that were down- or upregulated only in the Bur-0 virus-infected samples, respectively. We performed a GO term enrichment analysis with these gene sets using PlantRegMap server (Tian et al., 2020).

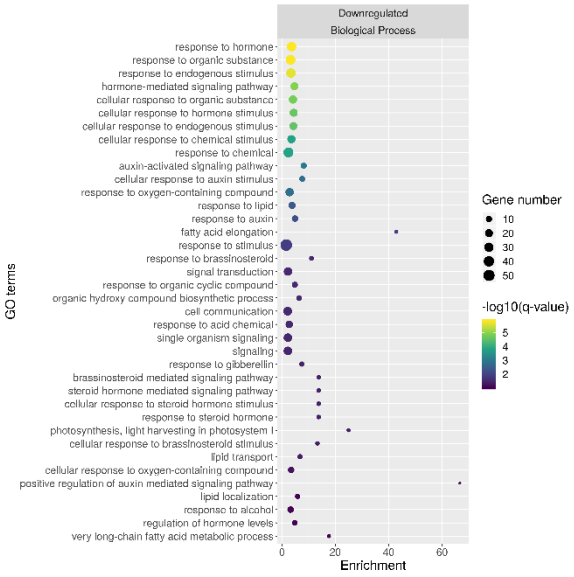

Fig. 8. GO term enrichment analysis of the significantly changed genes in the virus-infected Bur-0 ecotype plants.

Among the downregulated genes, terms related to hormone regulation (response to auxin, brassinosteroid, gibberellic acid, etc) were significantly enriched, while among the upregulated genes, only a few categories were enriched with low significance, all related to plant cell wall functions (Fig 8). Auxin, brassinosteroid, and gibberellic acid signalling are all involved in different stages of leaf development (Ali et al., 2020) and cross-talk in their signaling was observed earlier (Tian et al., 2018). The fact that these hormone pathways are all downregulated during virus-infection suggests that a common key function is blocked by the virus but only in the Bur-0 ecotype. This master regulator could be a transcriptional factor that commonly regulate genes involved in the mentioned hormonal pathways.

To profile the small RNA landscape upon virus infection and match it with the transcriptome profiles, we sequenced the sRNAome in the same samples that were used for the RNA-seq experiments. As expected, the sRNA profiles significantly changed in the virus-infected samples in both ecotypes. Normally, the 24-nt sRNAs dominate the sRNA landscape, while the second more abundant size class is the 21-nt sRNAs. Upon virus infection, this size distribution changes significantly, the 21–22-nt size classes become overwhelmingly dominant (Pitzalis et al., 2020). This is mainly due to the 21- and 22-nt vsiRNAs that are produced by DCL4 and DCL2 enzyme activities, respectively, and amplified by RDR1 or RDR6 in Arabidopsis (Deleris et al., 2006). Analysis of our sRNA libraries showed the same shift in size classes. However, when we filtered out the vsiRNAs by mapping the sRNA sequences to the *Arabidopsis* genome, the remaining host-derived sRNAs also showed a

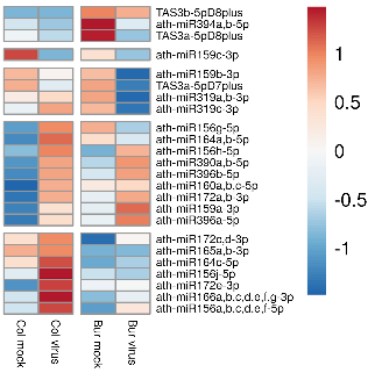

Fig. 9. DE of miRNAs and tasiRNAs in TVCV infected ecotypes.

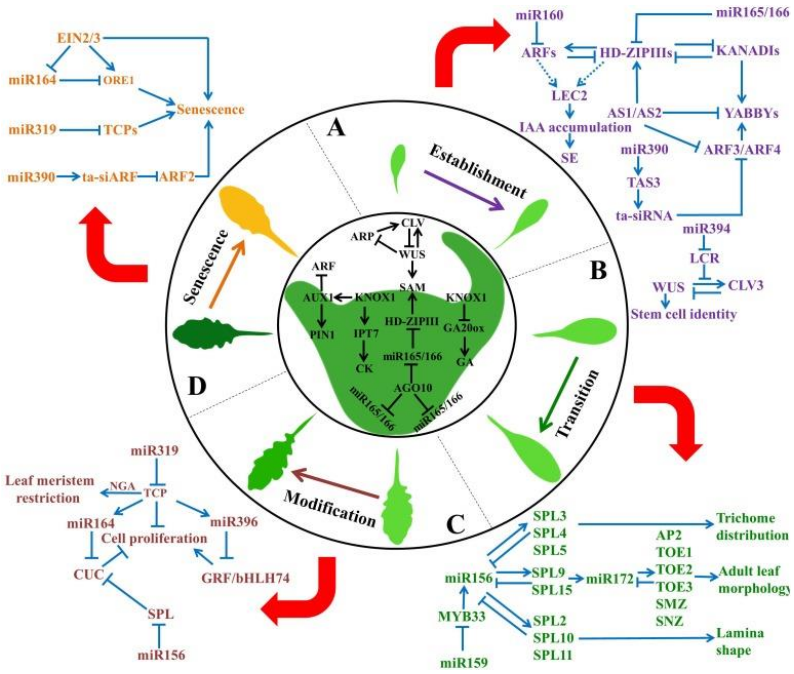

Fig. 9. MiRNAs and tasiRNAs are key regulators of leaf development. Pictures showing the role of sRNAs in leaf development is derived from Yang et al., 2018.

tremendous 21-nt size class enrichment in the virus-infected samples. This can be attributed to the vasiRNAs, which are produced in an RDR1 and DCL4- dependent manner (Cao et al, 2014). To find sRNAs whose expression either increased or decreased only in the Bur-0 virus-infected samples, we performed a differential expression analysis applying the same filtering rules as for the mRNAs. We identified several miRNAs and tasiRNAs, which are known to be involved in the regulation of leaf development (Fig 9), with differential expression upon TVCV infection in the two *Arabidopsis* ecotypes (Fig 10).

We checked the expression profiles of those mRNAs which are targeted by miRNAs that are known to play a role in leaf development. We found that miR159 and miR319 were downregulated only in Bur virus samples, while their target gene expression changed in the opposite direction (Fig 11). Both miRNAs have known roles in leaf polarity determination and are a good candidate with a role in symptom development during virus infection.

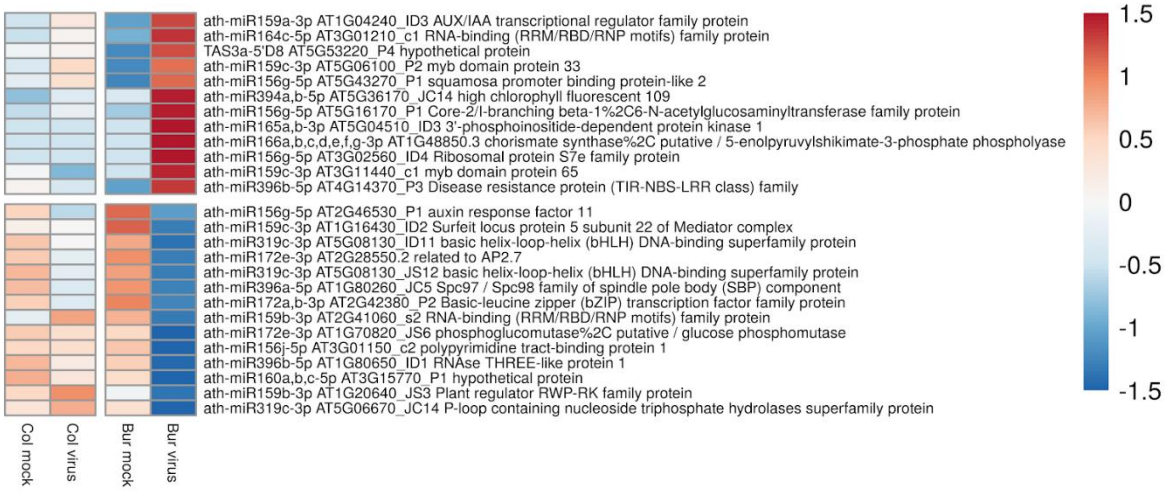

Fig. 11. Expression profile of leaf development-regulating miRNA target transcripts that are changed only in the Bur-0 ecotype only upon TVCV infection.

5.3. Genetic mapping of symptom determinant host loci with the use of Recombinant Inbred Lines (RIL) of the two ecotypes.

To narrow down the number of candidates, a genetic mapping was carried out using the Minimal Set of Recombinant Inbred Lines (RILs) that were obtained from INRA (Versailles, France) along with the genotypes of the RILs that were assessed by genotyping 87 molecular markers evenly distributed along the five chromosomes. Recombinant Inbred Lines (RIL) are useful for preliminary mapping of any trait that differs between parental strains. In plants RILs are generated by crossing two phenotypically and genotypically different parental lines (i.e. ecotypes), obtaining the segregating F2 population and then selfing individual lines descended from a single seed without any selection. The resulting F8 lines are nearly homozygous, carrying homozygous regions from both parental lines. These lines are genotyped by determining the pattern of genetic markers that are known to be polymorphic between the parental lines. The RIL mapping Minimal Set (Bur-0 x Col-0) containing 20 selected lines. This set is not intended for mapping, but rather to give an idea of the variation and transgression of any specific trait in the Bur-0 x Col-0 population using a small number of lines (together with the parents). With 164 lines, the Core-Pop164 is intended for optimized QTL mapping when using all 344 lines is impractical. For our goal to narrow down the list of candidate genes, we used the Minimal Set. Five individuals per line were infected either with TVCV or mock. One line was segregated (and not infected with the virus) and therefore excluded from the further analysis. Most of the RILs had mixed phenotypes that resembled one of the parents regarding some traits while the other parent regarding some other traits. The symptoms on the virus-infected plants were scored after one month (Fig 12) on a scale from 1 (Col-0) to 5 (Bur-0) compared with the corresponding mock-treated RIL (there were lines that developed symptom-like leaf deformation even in the absence of the virus). We focused only on the leaf axial deformation (spooning) either because other symptoms were observable on both ecotypes (i.e. delayed flowering, stunting) or they were difficult to score. A mean score for every line was calculated using the five individual scores per line. We used these scores, together with the genotyping data provided by INRA for an association study using the Windows QTL Cartographer software. A single marker association test revealed that four consecutive markers (c4_02133, c4_03833, c4_04877, and c4_05629) on the top of chromosome 4 were linked to the leaf deformation with a P-value of 0.007 and one more marker (c4_00641) near the terminus of chromosome 4 with a P-value of 0.034 (Fig 13). These markers represent a 3.5 Mb region (with 1% significance level) on chromosome 4 between

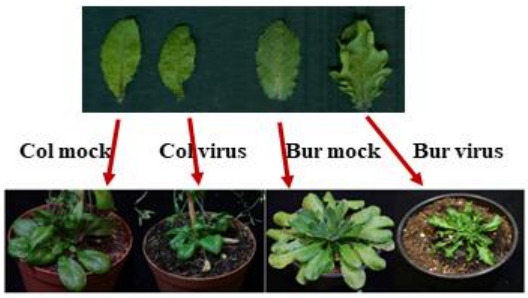

Fig. 12. Virus symptoms on infected Col-0 and Bur-0 plants.

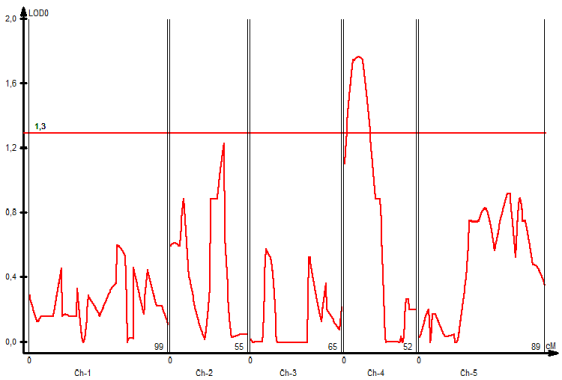

Fig. 13. The single marker association test predicted a major QTL on chromosome 4.

spanning 6.9-13.3 cM. For a more precise mapping position, further RILs showing polymorphism in this region will be tested. We checked the chromosomal location of some known regulators of the leaf axial development (TAS3, ARFs, miR165, HD-Zip factors, AS1,2, etc) and found that miR165B is near to this location, however, it is closer to the end of the chr4 (0.4 Mb). An auxin biosynthetic enzyme, YUC2 is within this region (7.7 Mb) but its role in the leaf axial development has not been studied yet. Of course, some other genes within the predicted region can be responsible for the symptoms, but we only list our best candidate genes in Table 1.

Table 1. Candidate genes that are differentially expressed only in the virus-infected Bur-0 ecotype samples and fall within the mapping region.

Gene ID	Gene name	Protein	Fold-change (virus/mock)	*Q*-value
AT4G03060	*AOP2*	2-oxoglutarate-dependent dioxygenase	1.81	4.23E-07
AT4G03156	*GTPase-like*	Small GTPase-like protein	5.34	1.12E-46
AT4G03400	*DFL2*	Auxin-responsive GH3-like protein 10	0.48	2.79E-08
AT4G03635	*FBL*	F-box/LRR protein	5.16	1.96E-35
AT4G05500	*FBL23*	F-box/LRR protein 23	4.56	2.14E-43

Most of the candidates, except for DFL2, are drastically upregulated upon TVCV infection in Bur-0 (Fig 14). However, none of them change expression upon virus infection or express at all in the Col-0 samples and two of them (AOP2 and FBL23) are annotated as pseudogenes in the Col-0 ecotype.

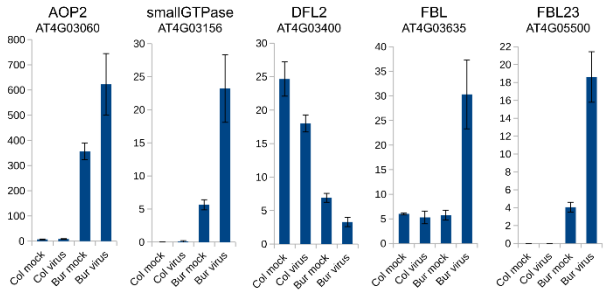

Fig. 14. Expression of symptom determinant candidate genes. Expression values are in tpm (transcript/million)

ALKENYL HYDROXALKYL PRODUCING 2 (AOP2) is a 2-oxoglutarate-dependent dioxygenase which is involved in glucosinolate biosynthesis. This gene is known to be non-functional in the Col-0 ecotype but functional in Cvi, and many other ecotypes. The natural variation in this locus explains the diversification of alkenyl glucosinolate among different ecotypes of *Arabidopsis thaliana*. AOP2 also mediates feed-back regulation of jasmonic acid signaling (Burow et al., 2015) and was shown to be regulated by two young miRNAs, miR826 and miR5090 (He et al., 2014).

The small GTPase-like gene encodes a small, 77 AA protein. Although it is annotated as a GTPase, it lacks some functionally conserved residues suggesting that it has some other functions. This gene does not express in Col-0 either.

DWARF IN LIGHT 2 (DFL2) encodes an early auxin-responsive GH3-type gene involved in red light-specific hypocotyl elongation (Takase et al., 2003). This gene is also downregulated by flg22, a pathogen response elicitor suggesting that downregulation of the auxin pathway is linked to pathogen response in plants (Navarro et al., 2006).

F-BOX-LIKE PROTEIN (FBL) is a member of a protein family called F-box/Leucin-rich repeat containing proteins (Kuroda et al., 2012). The auxin receptors TIR1, AFB2 and AFB3, and the jasmonic acid receptor COI1 also belong to this family. These proteins, upon binding the hormone, direct the ubiquitin-mediated degradation of their target proteins, i.e. AUX/IAA and JAZ proteins, respectively. FBL23 is also similar to these genes, however, it is predicted to be a pseudogene in Col-0. Indeed, it does not express in Col-0 at all, while it is significantly induced in virus-infected Bur-0 plants. When the Bur-0 sequence of the gene was examined, no large ORF could be detected in the publicly available Bur-0 RNA sequence. We assembled the FBL23 transcript *de novo* hoping that a virus-induced sample would result in a more accurate sequence given a higher abundance of the transcript under such condition. We could not detect a large ORF in this transcript either, suggesting that this gene is a noncoding gene in Bur-0 ecotype which is supported by the fact that this gene does not have an intron. The *de novo* assembly resulted in an antisense transcript as well. The TIR1, AFB2, AFB3 auxin receptor genes are known to be regulated by miR393 (Navarro et al., 2006). According to our analysis, neither FBL nor FBL23 is regulated by miR393.

It is worth noting that AUXIN-RESPONSE FACTOR 4 (ARF4), a key determinant of leaf polarity, is also significantly upregulated, while IAA29, an AUX/IAA protein that is regulated by proteolytic degradation by the auxin receptors is downregulated only in the Bur-0 plants. These results suggest that the virus interferes with the auxin signaling at some point, potentially at the proteolytic regulation of signaling components. It was described earlier that different viruses target the same ARF17 in rice (Zhang et al., 2020). Also, viral interference with the F-box mediated proteolytic degradation of host factors is well documented.

5.4. Validation of the symptom determinant host genes.

To confirm that the candidate genes found by analysing the RNA-seq and genetic mapping data indeed play a role in the leaf developmental defect caused by the virus infection, we created deletions in the Bur-0 plants in the above-mentioned genes by the CRISPR/Cas9 system. We expected that the lines carrying mutation in a gene that is upregulated upon virus infection would not display the adaxialized leaf deformation when infected with the virus (mimicking the Col-0 plants), while those ones with a mutation in a gene that is downregulated in Bur-0 virus-infected plants would mimic the virus-induced symptoms even in the absence of the virus.

Fig. 15. fbl23 Bur-0 plant.

We obtained at least one mutant line for each symptom determinant candidate gene. Two candidate gene knock out plants displayed interesting developmental phenotypes in the T0 generation even without virus infection: the fbl23 plants (Fig 15) looked more like Col-0 plants and flowered much earlier than the Bur-0 control plants (Col-0 plants also flower much earlier than Bur-0 plants), while fbl (Fig 16) plants had severe leaf deformations (serrated leaf margins and adaxial leaf blade, with small tentacle-like formations on the abaxial surface). We will collect seeds from all of the CRISPR/Cas9 mutant plant lines to verify these observations and test the plants phenotype upon TVCV infection in the next generations. According to the genotyping analysis, the fbl23 and fbl lines carry large deletions in the target genes in a homozygous form, while the other mutant lines appear to be heterozygous. In these cases, we expect to see phenotypes (if there is any) in the next generation. The effect on viral symptoms upon TVCV infection will be tested on the T2 generation of the mutant Bur-0 plants.

Fig. 16. fbl Bur-0 plant.

6. Identification of plant genes responsible for TuMV symptom development

Similarly, to TVCV, TuMV evoked different symptoms in Col-0 and Bur-0 ecotypes, despite that they belong to completely different virus families (Virgaviridae and Potyviridae, respectively). We have observed the adaxialization of leaves but not the curly leaf margins as in the case of TVCV infection. Furthermore, unlike TVCV infection, TuMV infection also resulted in chlorotic mosaics on the leaves (Fig 17). Adaxialization, as a common response, suggested that the two viruses manipulate the same host factors that determine leaf polarity. To reveal if there are such common host factors, we performed the same analysis as in the case of TVCV. We collected samples for the transcriptome and the sRNAome analysis from mock- and TuMV-infected Col-0 and Bur-0 leaves in four biological replicates, purified total RNAs. The sRNA libraries were prepared in our lab, while the libraries for the RNA-seq were prepared by the sequencing facility (Novogene Ltd). The high-throughput sequencing was carried out with the same sequencing technology and the data analysis was performed in the same way as described for the TVCV samples.

Fig. 17. TuMV symptoms on infected Col-0 and Bur-0 plants.

According to the RNA-seq analysis, the transcriptional response of the TuMV-infected Bur-0 plants were somewhat different from that of the TVCV-infected ones. There were 136 and 94

genes that were down- and upregulated only in the virus-infected Bur-0 plants, respectively (Fig 18). The most significantly enriched GO terms among the downregulated genes were

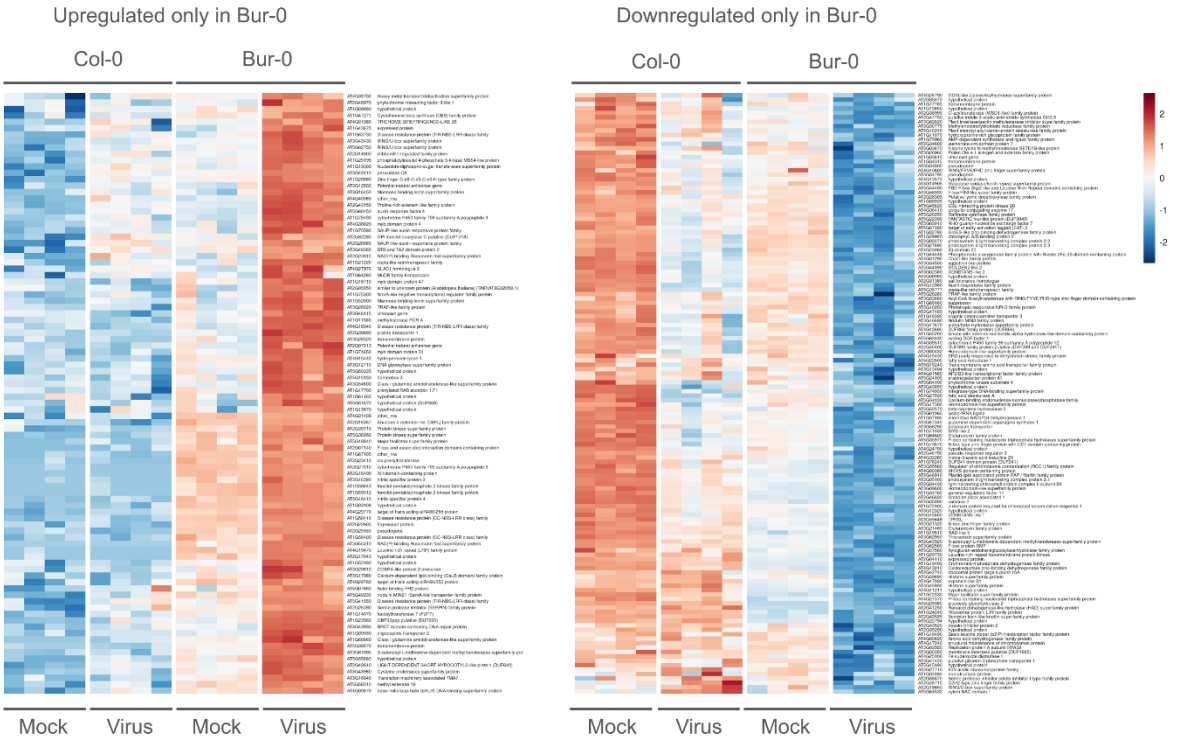

Fig. 18. Differentially expressed mRNAs in TuMV infected plants.

mostly associated with chloroplast functions and organization, photosynthesis, light responses, etc. This is perfectly in line with the observed chlorotic symptoms. There were also many genes associated with circadian rhythm and phytochrome regulation, including the core circadian clock element CCA1 and PRR9. This suggests that some of the TuMV proteins might interfere with a regulator of circadian gene expression and photomorphogenesis. Circadian clock is considered as a determinant of virus-plant interactions (Roden and Ingle, 2009). Jasmonate biosynthesis genes and hormone levels show a circadian pattern which results in a daily change in pathogen sensitivity of the plants (Thines et al., 2019). It was also shown that TuMV infection has a seasonality (Honjo et al., 2020). Beside differences, there were GO terms that were common in TVCV and TuMV infection, like response to auxin and gibberellin hormons. We checked how the expression pattern of the candidates from the TVCV-infected plants look like in the TuMV-infected plants (Fig 19). Some of the candidate genes showed a similar or same pattern as in TVCV infection (i.e. IAA29, DLF2, AOP2) while the FBLs did not really change upon virus infection in the Bur-0 plants. To reveal if one or more candidate genes also play a role in TuMV symptom development, we will infect the CRISPR-mutated plants with TuMV along with TVCV. As in the case of TVCV infection, there were only a few significantly enriched GO terms among the upregulated genes with low significance levels, all related to glycosinolate biosynthesis.

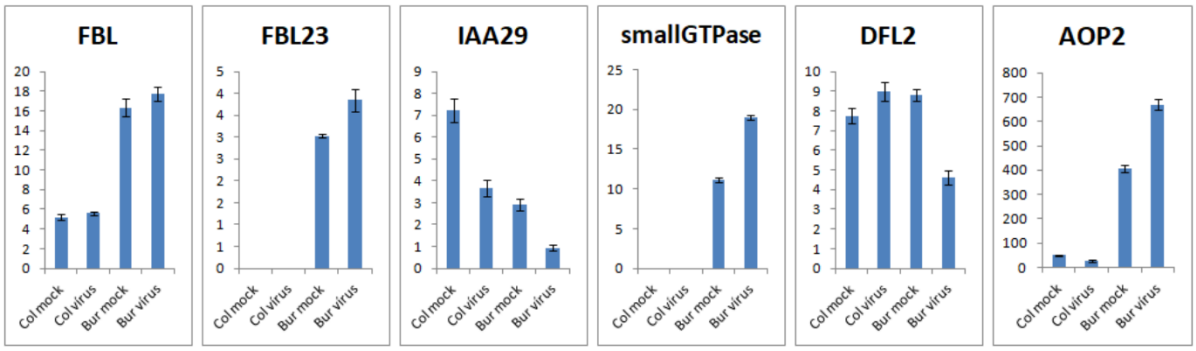

Fig. 19. Expression of symptom determinant candidate genes from the TVCV-infected plants in the TuMV-infected plants. Expression values are in tpm (transcript/million).

The sRNA expressions were similar to the TVCV samples. The 21-nt sRNAs dominated the sRNA landscape in the TuMV-infected plants, as expected. The leaf development regulating miRNAs expressed somewhat differently but the TAS3-ARF4 pathway looked the same: ARF4 was induced only in the TuMV-infected Bur-0 plants. Probably this is behind the adaxialization of the leaves in the virus-infected Bur-0 plants in both virus infections. However, unlike in TVCV-infected Bur-0 plants, the miR319 was not down regulated but somewhat upregulated in the TuMV-infected Bur-0 plants (as well as its target TCP3 but the change was even smaller) explaining the lack of leaf margin defect.

7. References

Ali, S.; Khan, N.; Xie, L. Molecular and Hormonal Regulation of Leaf Morphogenesis in Arabidopsis. Int. J. Mol. Sci. 2020, 21, 5132.

Bartel, D.P. MicroRNAs: Genomics, biogenesis, mechanism and function. Cell, 2004 116: 281-297.

Burgyán J, Havelda Z. Viral suppressors of RNA silencing. Trends Plant Sci. 2011; 16:265–272.

Burow M., Atwell S., Francisco M., Kerwin R.E., Halkier B.A., and Kliebenstein D.J. The glucosinolate biosynthetic gene AOP2 mediates feed-back regulation of jasmonic acidsignaling in Arabidopsis. Mol. Plant., 2015, 8: 1201-1212.

Cao, M. et al. Virus infection triggers widespread silencing of host genes by a distinct class of endogenous siRNAs in Arabidopsis. Proc. Natl Acad. Sci. USA 2014; 111, 14613–14618.

Carrington JC, Kasschau KD, Mahajan SK, Schaad MC. Cell-to-Cell and Long-Distance Transport of Viruses in Plants. Plant Cell 1996; 8:1669–1681.

Chen, X. Small RNAs and Their Roles in Plant Development. Annu. Rev. Cell Dev. Biol., 2009 35: 21-44.

Csorba T., Kontra L., Burgyán J. Viral silencing suppressors: Tools forged to fine-tune host-pathogen coexistence. Virology. 2015, 479-480:85-103.

Deleris A, Gallego-Bartolome J, Bao J, Kasschau KD, Carrington JC, Voinnet O. Hierarchical Action and Inhibition of Plant Dicer-Like Proteins in Antiviral Defense. Science 2006; 313:68–71.

Ding S-W, Voinnet O. Antiviral Immunity Directed by Small RNAs. Cell 2007; 130:413–426.

He H, Liang G, Li Y, Wang F, Yu D. Two young MicroRNAs originating from target duplication mediate nitrogen starvation adaptation via regulation of glucosinolate synthesis in Arabidopsis thaliana. Plant Physiol. 2014; 164:853-865.

Hull R. Chapter 4 - Symptoms and Host Range. In: Hull R, ed. Plant Virology (Fifth Edition). Boston: Academic Press; 2014:145–198.

Jones-Rhoades, M. W., Bartel, D. P. and Bartel, B. MicroRNAs and their regulatory roles in plants. Annu. Rev. Plant Biol., 2006 57: 19-53.

Kis, S., Salamon, P., Kis, V., Szittya, G. Molecular characterization of a beet ringspot nepovirus isolated from Begonia ricinifolia in Hungary. Arch Virol., 2017; 162:3559-3562.

Kuroda H, Yanagawa Y, Takahashi N, Horii Y, Matsui M. A comprehensive analysis of interaction and localization of Arabidopsis SKP1-like (ASK) and F-box (FBX) proteins. PLoS One., 2012; 7:e50009.

Li, F., Pignatta, D., Bendix, C., Brunkard, J. O., Cohn, M. M., Tung, J., Sun, H., Kumar, P., and Baker, B. MicroRNA regulation of plant innate immune receptors. Proc. Natl. Acad. Sci. USA., 2012; 109:1790-1795.

Mallory, A. C. and Vaucheret, H. Functions of microRNAs and related small RNAs in plants. Nat. Genet., 2006 38: S31-S36. doi:10.1038/ng1791.

Medzihradszky, A., Gyula, P., Sós-Hegedűs, A., Szittya, G., Burgyán, J. Transcriptome reprogramming in the shoot apical meristem of CymRSV-infected Nicotiana benthamiana plants associates with viral exclusion and the lack of recovery. Molecular Plant Pathology, 2019; 20 (12), 1748-1758.

Meyers, B.C., Kozik, A., Griego, A., Kuang, H., and Michelmore, R.W. Genome-wide analysis of NBS-LRR-encoding genes in Arabidopsis. Plant Cell, 2003; 15: 809-834.

Molnár A, Csorba T, Lakatos L, Várallyay É, Lacomme C, Burgyán J. Plant Virus-Derived Small Interfering RNAs Originate Predominantly from Highly Structured Single-Stranded Viral RNAs. J. Virol. 2005; 79:7812–7818.

Navarro L, Dunoyer P, Jay F, Arnold B, Dharmasiri N, Estelle M, Voinnet O, Jones JD. A plant miRNA contributes to antibacterial resistance by repressing auxin signaling. Science. 2006; 312:436-439.

Pimentel H, Bray NL, Puente S, Melsted P, Pachter L. Differential analysis of RNA-seq incorporating quantification uncertainty. Nat Methods., 2017; 14:687–690.

Pitzalis, N., Amari, K., Graindorge, S., Pflieger, D., Donaire, L., Wassenegger, M., Llave, C. and Heinlein, M. Turnip mosaic virus in oilseed rape activates networks of sRNA-mediated interactions between viral and host genomes. Communications Biology, 2020, 3:702.

Roden LC, Ingle RA. Lights, rhythms, infection: the role of light and the circadian clock in determining the outcome of plant-pathogen interactions. Plant Cell., 2009; 21:2546-2552.

Salamon P, Sos-Hegedus A, Gyula P, Szittya G. First report of the infection of alfalfa mosaic virus in Salvia sclarea in Hungary. Journal of Plant Pathology; 2018; 100 (3):607.

Shimura H, Pantaleo V, Ishihara T, et al. A Viral Satellite RNA Induces Yellow Symptoms on Tobacco by Targeting a Gene Involved in Chlorophyll Biosynthesis using the RNA Silencing Machinery. PLoS Pathog 2011; 7: e1002021.

Shivaprasad, P. V., Chen, H. M., Patel, K., Bond, D. M., Santos, B. A., and Baulcombe, D. C. A microRNA superfamily regulates nucleotide binding site-leucine-rich repeats and other mRNAs. Plant Cell.,2012; 24:859-874.

Smith NA, Eamens AL, Wang M-B. Viral Small Interfering RNAs Target Host Genes to Mediate Disease Symptoms in Plants. PLoS Pathog 2011;7: e1002022.

Sós-Hegedűs, A., Domonkos, Á., Tóth, T., Gyula, P., Kaló P., Szittya, G. Suppression of NB-LRR Genes by miRNAs Promotes Nitrogen-fixing Nodule Development in Medicago truncatula. Plant Cell Environ., 2020; 43(5):1117-1129.

Szittya, G. and Burgyán, J. RNA interference-mediated intrinsic antiviral immunity in plants. Curr Top Microbiol Immunol. 2013, 371: 153-81.

Takase T, Nakazawa M, Ishikawa A, Manabe K, Matsui M. DFL2, a new member of the Arabidopsis GH3 gene family, is involved in red light-specific hypocotyl elongation. Plant Cell Physiol. 2003; 44:1071-1080.

Tian H, Lv B, Ding T, Bai M, Ding Z. Auxin-BR Interaction Regulates Plant Growth and Development. Front Plant Sci., 2018; 8:2256.

Tian, F., Yang, D.C., Meng, Y.Q., Jin, J. and Gao, G. PlantRegMap: charting functional regulatory maps in plants. Nucleic Acids Res., 2020, 48, D1104-D1113.

Thines B, Parlan EV, Fulton EC. Circadian Network Interactions with Jasmonate Signaling and Defense. Plants (Basel). 2019; 8:252.

Tóth, T., Gyula, P., Salamon, P., Kis, S., Sós-Hegedűs, A., Szittya, G. Molecular characterization and in vitro synthesis of infectious RNA of a Turnip vein-clearing virus isolated from Alliaria petiolata in Hungary. PLoS ONE, 2019; 14: e0224398.

Voinnet O. Induction and suppression of RNA silencing: insights from viral infections. Nat Rev Genet 2005; 6:206–220.

Weaver, L. M., Swiderski, M. R., Li, Y. and Jones, J. D. G. The Arabidopsis thaliana TIR-NB-LRR R-protein, RPP1A; protein localization and constitutive activation of defence by truncated alleles in tobacco and Arabidopsis. Plant J., 2006; 47: 829-840.

Zhai, J., Jeong, D-H., De Paoli, E., Park, S., Rosen, B. D., Li, Y., Gonzales, A. J. Yan, Z., Kitto, L. S., Grusak, M. A., Jackson, S. A., Stacey, G., Cook, D. R., Green, P. J., Sherrier, D. J. and Meyers, B. C. MicroRNAs as master regulators of the plant NB-LRR defense gene family via the production of phased, trans-acting siRNAs. Genes Dev., 2011; 25: 2540-2553.

Zhang H, Li L, He Y, Qin Q, Chen C, Wei Z, Tan X, Xie K, Zhang R, Hong G, Li J, Li J, Yan C, Yan F, Li Y, Chen J, Sun Z. Distinct modes of manipulation of rice auxin response factor OsARF17 by different plant RNA viruses for infection. Proc Natl Acad Sci U S A., 2020; 117:9112-9121.

Yang, T., Wang, Y., Teotia, S., Zhang, Z., and Tang, G. The Making of Leaves: How Small RNA Networks Modulate Leaf Development. Frontiers in Plant Science, 2018; 9: 824.