

Ecotype-specific blockage of tasiARF production by two different RNA viruses in *Arabidopsis*  
 --Manuscript Draft--

<b>Manuscript Number:</b>	PONE-D-22-09203R1
<b>Article Type:</b>	Research Article
<b>Full Title:</b>	Ecotype-specific blockage of tasiARF production by two different RNA viruses in <i>Arabidopsis</i>
<b>Short Title:</b>	Transcriptome analysis of ecotype-dependent viral symptoms in <i>Arabidopsis</i>
<b>Corresponding Author:</b>	Péter Gyula, Ph.D. Hungarian University of Agriculture and Life Sciences Institute of Genetics and Biotechnology: Magyar Agrar- és Éleltudományi Egyetem Genetika és Biotechnológia Intézet Gödöllő, HUNGARY
<b>Keywords:</b>	<i>Arabidopsis thaliana</i> ; ecotypes; TVCV; turnip vein-clearing virus; TuMV; turnip mosaic virus; viral symptom development; leaf deformation; leaf polarity; leaf development; auxin signaling; AGO7; miR390; ARF4; FIL; AFO; YABBY1; TAS3; tasiARF; Transcriptome analysis; antiviral silencing; PTGS; small RNA; tasiRNA; vasiRNA
<b>Abstract:</b>	<i>Arabidopsis thaliana</i> is one of the most studied model organisms of plant biology with hundreds of geographical variants called ecotypes. One might expect that this enormous genetic variety could result in a differential response to pathogens. Indeed, we observed previously that the Bur ecotype develops much more severe symptoms (upward curling leaves and wavy leaf margins) upon infection with two positive strand RNA viruses of different families (turnip vein-clearing virus, TVCV, and turnip mosaic virus, TuMV). To find the genes potentially responsible for the ecotype-specific response, we performed a differential expression analysis of the mRNA and sRNA pools of TVCV and TuMV-infected Bur and Col plants along with the corresponding mock controls. We focused on the genes and sRNAs that showed an induced or reduced expression selectively in the Bur virus samples in both virus series. We found that the two ecotypes respond to the viral infection differently, yet both viruses selectively block the production of the TAS3-derived small RNA specimen called tasiARF only in the virus-infected Bur plants. The tasiARF normally forms a gradient through the adaxial and abaxial part of the leaf (being more abundant in the adaxial part) and post-transcriptionally regulates ARF4, a major leaf polarity determinant in plants. The lack of tasiARF-mediated silencing could lead to an ectopically expressed ARF4 in the adaxial part of the leaf where the misregulation of auxin-dependent signaling would result in an irregular growth of the leaf blade manifesting as upward curling leaf and wavy leaf margin. QTL mapping using Recombinant Inbred Lines (RILs) suggests that the observed symptoms are the result of a multigenic interaction that allows the symptoms to develop only in the Bur ecotype. The particular nature of genetic differences leading to the ecotype-specific symptoms remains obscure and needs further study.
<b>Order of Authors:</b>	Péter Gyula, Ph.D. Tamás Tóth Teréz Gorcsa Tünde Nyikó, Ph.D. Anita Sós-Hegedűs, Ph.D. György Szittya, Ph.D.
<b>Opposed Reviewers:</b>	
<b>Response to Reviewers:</b>	First of all, we appreciate the time and effort the reviewers invested to provide us with their valuable criticism and recommendations. We rewrote the introduction, results, and discussion part according to the suggestions which we believe significantly improved

the manuscript. Please find our detailed reactions below.

Reviewer #1:

#### Aims and methods

Dr Gyula co-authors characterised the sRNA and RNA profiles of two Arabidopsis ecotypes in terms of interaction with two viruses belonging to different taxonomic groups. The aim of the study was to identify patterns/responses that could explain the common phenotype induced by the two viruses in the 'bur' ecotype alone and not in the 'col' ecotype.

#### Relevance of the study

The scientific questions posed by the authors are relevant, all the more so since plant ecotypes differentiate in precise geographic area and stabilise genetic traits whose study is a) in continuity with the establishment of platforms for the conservation and characterisation of plant genetic resources; b) fundamental for investigating and characterising traits of fast-changing environmental adaptation with regard to biotic and abiotic stresses.

#### The main results

The main answer in the study lies in the functionality of the long non-coding RNA called TAS3 and the abnormal production (abolished in the 'bur' ecotype) of tasiARFs. The results are consistent with the known and well-studied function of ARFs in leaf development (also well illustrated in the introduction and discussions). In itself, this finding might be sufficient to support the main conclusion.

In this frame, however, the authors discuss that AGOs involved in the function of tasiARFs could have differences in the two ecotypes, but they do not show it (different amino acid composition SNPs/indels at functional points): the authors are invited to evaluate and show (if present) such differences available in the 1001 genomes project at <https://1001genomes.org/>.

Answer: Thank you for your suggestion. We aligned the amino acid sequences of AGO1, AGO2, AGO7, ARF4, and FIL from 19 different ecotypes, including Col and Bur (Supplementary Fig S2). The analysis revealed that there were 5 amino acid changes in the Bur version of AGO7 compared to the Col version (all in the N-terminal part), while there were 0, 1, 2, and 0 amino acid changes in AGO1, AGO2, ARF4, and FIL, respectively (Fig S2). The Bur version of AGO7 was more different than the versions in the rest of the ecotypes. The 2 amino acid changes in ARF4 were not conserved (E/K and T/I in Col and Bur, respectively), and again, the Bur variant was rare among the investigated ecotypes. The FIL and AGO1 sequences were remarkably conserved among the investigated 19 ecotypes suggesting a strong selection pressure on them. There was some degree of variability in the AGO2 sequence but the Bur was not the most divergent ecotype in this respect.

#### Main criticisms

There are, however, some notable criticisms. The authors adopt a very interesting 'omic' approach to identifying genes involved in the symptomatological expression of the disease without definitely and fully considering (even only in the introduction) what has been discovered so far in terms of viral-associated siRNAs derived from coding genes. In particular, I refer to the work of Cao et al, 2012 (mentioned but not well discussed, being source of valuable information in the supplementary data concerning infection with TuMV, the same virus analysed in the present study), Leonetti et al, 2020 and Pitzalis et al, 2020. In particular, a cursory analysis of the Data\_S1 table reveals that a large proportion of the unique sRNAs are 22-nt long (6,863, as opposed to 16,336 of 21-nt). This is why the authors, should take the opportunity of the knowledge they have gained (see Leonetti et al., 2020 on the functionality of 22-nt vasiRNAs) and possibly extend it.

The authors are therefore invited to consider the issue of vasiRNAs as being responsible, together with tasiARFs. Alternatively they should eliminate the analysis of vasiRNAs from this study.

Answer: We admit that the discussion part could have been more thorough, especially regarding the research on the function vasiRNAs in plants (we fixed this in the Introduction and Discussion part, see the yellow text). It was because we focused on

the differences rather than the similarities between the two ecotypes regarding their antiviral responses. We wanted to avoid discussing the commonly regulated sRNAs that were possibly not responsible for the differential symptoms observed in the two ecotypes. Additionally, we focused on the genes/sRNAs that were similarly differentially expressed only in the Bur ecotype infected by two different viruses and not discussed those genes/sRNAs that were unique for the two viruses or that were changed in the opposite direction. However, we observed that the vast majority of the commonly downregulated genes are potentially regulated by vasiRNAs that are commonly upregulated only in the Bur ecotype infected with both viruses. This is surprising, and therefore, we decided to discuss them in a little more detail, especially because these genes have functions that could be related to leaf development (i.e. auxin metabolism, cell wall expansion, etc).

Regarding the 22-nt-long vasiRNAs, the number of 22-nt sRNAs (6863) is the number of unique sequences in all the samples, including mock samples, therefore, it does not tell us how virus infection affects this size class. Actually, we found no such markedly increased population of this size class as in the case of CaMV infection investigated by Leonetti et al (see our corrected Fig 7). Maybe this is a CaMV-specific feature.

Moreover, with regard to vasiRNAs we note that there are no sRNAs in DATA\_S1 that are derived from FIL and this finding is inconsistent with two of the panels in Figure 10 (i.e. vasiRNAs-FIL).

Answer: We are sorry for being confusing here: the vasiRNA-FIL is derived from a transposon (not from FIL) and potentially targets FIL in trans. It is analogous to tasi-ARF, when the functional siRNA is derived from the TAS3 non-coding RNA and targets ARF3- and ARF4-coding transcripts. We are aware that vasiRNAs could target the gene they derived from (in this case, a transposon). To predict potential vasiRNA/target pairs, we only considered those vasiRNAs and target genes that were differentially expressed only in the Bur virus samples in the opposite direction, i.e. upregulated vasiRNAs vs downregulated target genes. We did not predict all the possible targets of all the vasiRNAs, because we considered it irrelevant to our goal to find the possible cause of the differential leaf deformation.

It is also unclear whether in the histograms of the figures, when referring to sRNAs from specific loci, only those of 21-nt are considered or also those of other lengths: to be clarified.

Answer: When we show the expression of certain sRNA specimens (i.e. miR390a,b-5p, tasiARF a.k.a. TAS3 5'D7(+), TAS1abc-siR255, etc), we considered only the conserved, mature sequence; iso-miRs or other sequence variants (or other, non-functional siRNAs from the same TAS locus) were not included in the analysis.

#### Other criticisms

In the discussions an important concept is reported <>. The first part of the sentence is related to what is requested above (see "The main results" of comments to authors) but is not fully supported in the part concerning the evolution of the VSRs. In this regard, it is well known that VSRs of different viral taxonomic groups target different steps in the silencing pathways with different mechanisms, implying convergent evolution, i.e., having analogous features but non-homologous motifs. Strong VSR evolution/diversification is attributable to episodic selection rather than to pervasive positive selection (Murray et al., 2013). In other words, the diversification of plant virus VSRs is strongly boosted by frequent jumps between host species/genotypes rather than a one-to-one co-evolution. If what shown by Murray et al (2013) is in line with your discussion, please clarify.

Answer: We guess that the incriminated sentence was about the parallel evolution of the VSRs of different viruses targeting rice ARF17. We suppose that in our case the VSRs of the two different viruses interfere with the biosynthesis of the tasi-ARF in different ways. For example, one VSR might bind miR390 while the other might block AGO7 activity. As the reviewer pointed out, this would require non-homologous protein motifs. Both VSRs were reported to bind double-stranded sRNAs and HCPro was reported to bind AGO1. We analyzed the sequence of the miR390 duplex in the two ecotypes because we theorized that the different duplex structure (i.e. because of a possibly different star sequence) could result in a differential binding of miR390 duplex

by the VSR(s). However, we found no evidence for such a difference. AGO1 amino acid sequences of the two ecotypes are also identical. We think that the best target of the VSRs is either the AGO7 or AGO2 proteins, or some unknown component of the tasi-ARF biosynthesis pathway. Murray et al argue that plant VSRs are subject to episodic positive selection due to frequent jumps between different hosts rather than a one-to-one co-evolution with a single host. This is consistent with our findings since both viruses we worked with have a broad host range (i.e. the turnip vein-clearing virus isolate we used, TVCV-ApH, was isolated from *Alliaria petiolata*, a.k.a. garlic mustard).

Minor:  
GWAS is only mentioned once in the discussions: do not use (unnecessary) acronym, which is instead required when the term is recurrent in the text more than 2-3 times.

Answer: The Reviewer is right, we replaced the acronym with the full name.

Reviewer #2:

Although this manuscript did not conclude with genetic examinations, this reviewer thinks it provides helpful suggestions for future studies to identify host factors involved in the symptom expressions the authors focused on, including leaf deformation caused by two distinct virus infections in *Arabidopsis*. The authors explored the host factors by comparison analysis with RNA-seq and small RNA-seq between virus-infected Col and Bur ecotypes. Bur showed leaf deformation with these virus infections, but Col did not. Their RNA-seq showed that two genes, ARF4 and FIL, were specifically induced and reduced in Bur. Then, sRNA-seq consistently revealed that tasiARF that regulates the ARF4 expression was specifically reduced among other tasiRNAs. QTL analysis with RILs between Bur and Col supported their transcriptome analyses; that is, one of the QTL detected was around the locus of ARF4, though the QTL analysis is perhaps considered preliminary level.

Minor points:  
P2, L16, and L17, These ARF4 mean protein and thus should be roman.

Answer: We changed the format to refer to protein products.

P3, L17-22, These descriptions are of the case when plants have appropriate R genes that recognize invaded viruses. Rewrite them.

Answer: Thank you for noticing this, we rephrased the sentence as suggested.

P25, L17, Fig 10 should be Fig 12  
Answer: We fixed it.

**Additional Information:**

Question	Response
<p><b>Financial Disclosure</b></p> <p>Enter a financial disclosure statement that describes the sources of funding for the work included in this submission. Review the <a href="#">submission guidelines</a> for detailed requirements. View published research articles from <a href="#">PLOS ONE</a> for specific examples.</p> <p>This statement is required for submission and <b>will appear in the published article</b> if the submission is accepted. Please make sure it is accurate.</p>	<p>This work was funded by the Hungarian Government organization NRD1 (National Research, Development, and Innovation Office: <a href="https://nkfih.gov.hu/about-the-office">https://nkfih.gov.hu/about-the-office</a>) through the grants K-119701, K-129171, K-134974 awarded to GS, FK-137811 awarded to PG, and PD-129119 awarded to TN. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.</p>

### Unfunded studies

Enter: *The author(s) received no specific funding for this work.*

### Funded studies

Enter a statement with the following details:

- Initials of the authors who received each award
- Grant numbers awarded to each author
- The full name of each funder
- URL of each funder website
- Did the sponsors or funders play any role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript?
- **NO** - Include this sentence at the end of your statement: *The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*
- **YES** - Specify the role(s) played.

\* typeset

### Competing Interests

Use the instructions below to enter a competing interest statement for this submission. On behalf of all authors, disclose any [competing interests](#) that could be perceived to bias this work—acknowledging all financial support and any other relevant financial or non-financial competing interests.

This statement is **required** for submission and **will appear in the published article** if the submission is accepted. Please make sure it is accurate and that any funding sources listed in your Funding Information later in the submission form are also declared in your Financial Disclosure statement.

View published research articles from [PLOS ONE](#) for specific examples.

The authors have declared that no competing interests exist.

**NO authors have competing interests**

Enter: *The authors have declared that no competing interests exist.*

**Authors with competing interests**

Enter competing interest details beginning with this statement:

*I have read the journal's policy and the authors of this manuscript have the following competing interests: [insert competing interests here]*

\* typeset

**Ethics Statement**

N/A

Enter an ethics statement for this submission. This statement is required if the study involved:

- Human participants
- Human specimens or tissue
- Vertebrate animals or cephalopods
- Vertebrate embryos or tissues
- Field research

Write "N/A" if the submission does not require an ethics statement.

General guidance is provided below. Consult the [submission guidelines](#) for detailed instructions. **Make sure that all information entered here is included in the Methods section of the manuscript.**

**Format for specific study types**

**Human Subject Research (involving human participants and/or tissue)**

- Give the name of the institutional review board or ethics committee that approved the study
- Include the approval number and/or a statement indicating approval of this research
- Indicate the form of consent obtained (written/oral) or the reason that consent was not obtained (e.g. the data were analyzed anonymously)

**Animal Research (involving vertebrate animals, embryos or tissues)**

- Provide the name of the Institutional Animal Care and Use Committee (IACUC) or other relevant ethics board that reviewed the study protocol, and indicate whether they approved this research or granted a formal waiver of ethical approval
- Include an approval number if one was obtained
- If the study involved *non-human primates*, add *additional details* about animal welfare and steps taken to ameliorate suffering
- If anesthesia, euthanasia, or any kind of animal sacrifice is part of the study, include briefly which substances and/or methods were applied

**Field Research**

Include the following details if this study involves the collection of plant, animal, or other materials from a natural setting:

- Field permit number
- Name of the institution or relevant body that granted permission

**Data Availability**

Authors are required to make all data underlying the findings described fully available, without restriction, and from the time of publication. PLOS allows rare exceptions to address legal and ethical concerns. See the [PLOS Data Policy](#) and [FAQ](#) for detailed information.

Yes - all data are fully available without restriction

A Data Availability Statement describing where the data can be found is required at submission. Your answers to this question constitute the Data Availability Statement and **will be published in the article**, if accepted.

**Important:** Stating 'data available on request from the author' is not sufficient. If your data are only available upon request, select 'No' for the first question and explain your exceptional situation in the text box.

Do the authors confirm that all data underlying the findings described in their manuscript are fully available without restriction?

**Describe where the data may be found in full sentences. If you are copying our sample text, replace any instances of XXX with the appropriate details.**

- If the data are **held or will be held in a public repository**, include URLs, accession numbers or DOIs. If this information will only be available after acceptance, indicate this by ticking the box below. For example: *All XXX files are available from the XXX database (accession number(s) XXX, XXX).*
- If the data are all contained **within the manuscript and/or Supporting Information files**, enter the following: *All relevant data are within the manuscript and its Supporting Information files.*
- If neither of these applies but you are able to provide **details of access elsewhere**, with or without limitations, please do so. For example:

*Data cannot be shared publicly because of [XXX]. Data are available from the XXX Institutional Data Access / Ethics Committee (contact via XXX) for researchers who meet the criteria for access to confidential data.*

*The data underlying the results presented in the study are available from (include the name of the third party*

The raw sequencing data were deposited in the NCBI SRA database (BioProject: PRJNA788379) under the identifiers SRR17227515–SRR17227578. All the other relevant data are within the manuscript and its Supporting Information files.



*and contact information or URL).*

- This text is appropriate if the data are owned by a third party and authors do not have permission to share the data.

\* typeset

Additional data availability information:

Dear Editor,

We are submitting the revised version (PONE-D-22-09203R1) of the manuscript titled “Ecotype-specific blockage of tasiARF production by two different RNA viruses in *Arabidopsis*”. We addressed the questions the Reviewers raised and rewrote the introduction, results, and discussion part according to the suggestions which we believe significantly improved the manuscript. We also prepared supplementary figures and a new version of Fig 7 and added further references. Every change in the text was highlighted in yellow.

We hope that our answers are satisfactory and the manuscript will finally be accepted.

Yours sincerely,

György Szittyá and Péter Gyula



# Ecotype-specific blockage of tasiARF production by two different RNA viruses in *Arabidopsis*

Péter Gyula<sup>1¶\*</sup>, Tamás Tóth<sup>1,2¶</sup>, Teréz Gorcsa<sup>1</sup>, Tünde Nyikó<sup>1</sup>, Anita Sós-Hegedűs<sup>1</sup>, and György Szittya<sup>1\*</sup>

<sup>1</sup> Department of Plant Biotechnology, Institute of Genetics and Biotechnology, Hungarian University of Agriculture and Life Sciences, Gödöllő, Hungary

<sup>2</sup> Doctoral School of Biology, Institute of Biology, ELTE Eötvös Loránd University, Budapest, Hungary

\* Corresponding author

E-mail: [gyula.peter@uni-mate.hu](mailto:gyula.peter@uni-mate.hu) (GP) and [szittya.gyorgy@uni-mate.hu](mailto:szittya.gyorgy@uni-mate.hu) (GS)

¶ These authors contributed equally to this work.

## Abstract

*Arabidopsis thaliana* is one of the most studied model organisms of plant biology with hundreds of geographical variants called ecotypes. One might expect that this enormous genetic variety could result in differential response to pathogens. Indeed, we observed previously that the Bur ecotype develops much more severe symptoms (upward curling leaves and wavy leaf margins) upon infection with two positive-strand RNA viruses of different families (turnip vein-clearing virus, TVCV, and turnip mosaic virus, TuMV). To find the genes potentially responsible for the ecotype-specific response, we performed a differential expression analysis of the mRNA and sRNA pools of TVCV and TuMV-infected Bur and Col plants along with the corresponding mock controls. We focused on the genes and sRNAs that showed an induced or reduced expression selectively in the Bur virus samples in both virus series. We found that the two ecotypes respond to the viral infection differently, yet both viruses selectively block the production of the *TAS3*-derived small RNA specimen called tasiARF only in the virus-infected Bur plants. The tasiARF normally forms a gradient through the adaxial and abaxial parts of the leaf (being more abundant in the adaxial part) and post-transcriptionally regulates ARF4, a major leaf polarity determinant in plants. The lack of tasiARF-mediated silencing could lead to an ectopically expressed ARF4 in the adaxial part of the leaf where the misregulation of auxin-dependent signaling would result in an irregular growth of the leaf blade manifesting as upward curling leaf and wavy leaf margin. QTL mapping using Recombinant Inbred Lines (RILs) suggests that the observed symptoms are the result of a multigenic interaction that allows the symptoms to develop only in the Bur ecotype. The particular nature of genetic differences leading to the ecotype-specific symptoms remains obscure and needs further study.

## Introduction

Virus-infected plants may develop various symptoms on all or some of their parts. The symptoms may either be nonspecific or quite characteristic of the particular virus. Common symptoms are dwarfing, stunting, various leaf deformations, chlorotic mosaics, vein clearing, etc. These effects may either be 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 environmental conditions, like temperature and light. Finally, plants may show acute or severe symptoms soon after inoculation that may lead to the 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 a gradual decline of the plant [1].

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 virally encoded proteins become active. Movement proteins help the viral nucleic acid move into another cell through plasmodesmata [2]. At the site of entry, plants carrying appropriate R genes may develop a hypersensitive response that involves the overproduction of reactive oxygen species and programmed cell death which prevents further spreading of the virus into the neighboring 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, and necrosis.

To restrict viral accumulation, plants and other eukaryotes utilize a highly adaptive, sequence-specific mechanism called RNA silencing [3]. 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) intermediate product of viral replication trigger antiviral silencing [4]. The dsRNA is randomly cut into 21–22 bp-long fragments called virus-derived short interfering RNAs (vsiRNAs) by specific Dicer-like (DCL) endonucleases [5]. One of the strands of these fragments called the guide strand is then loaded into an RNA-induced silencing complex (RISC). With 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 efficient virus infection, viruses should be able to evade the silencing machinery. In the evolutionary arms race, viruses developed a plethora of different viral silencing suppressor (VSR) molecules that can interfere with various steps of the silencing pathway, including small RNA production, processing, binding, stability, and activity of the RISC [6–8]. As a side effect, VSRs can interfere with endogenous sRNA-regulated processes that result in a corrupted development of the plant.

One major class of regulatory sRNAs is the micro RNAs (miRNAs) which are usually 20–22-nt-long sequences produced by a DCL1-mediated dicing of a hairpin-like, double-stranded precursor with imperfect self-complementarity. Different AGOs preferentially bind miRNAs with a specific 5' nucleotide [9,10] and length while the loading efficiency is controlled by the structure of the double-stranded precursor [11]. Viruses often modify the miRNA landscape of the plant which affects symptom development [12–14].

Many aspects of plant development where a pattern formation is required are regulated by cell-to-cell mobile miRNAs or siRNAs [15–17]. This cell-to-cell movement results in a sharp gradient of the sRNA within a tissue, for example, between the adaxial and abaxial parts of the leaf, establishing patterns of cell specification by creating boundaries that limit

the activity of target genes [17,18]. A conserved regulatory module among the land plants that plays a central role in abaxial-adaxial leaf patterning, juvenile-to-adult phase transition, floral organ development, and lateral root development is the miR390/AGO7/TAS3/tasiARF/ARF3/ARF4 module [19–28]. The tasiARF (*trans*-acting ARF-targeting small interfering RNA) is a special kind of small interfering RNA (siRNA) that is produced from the *TAS3* non-coding RNA (ncRNA) precursor. During the production of tasiARF, the *TAS3* transcript is first cleaved by a miR390-loaded AGO7 [29], then, a double-stranded intermediate product is produced by the RNA-dependent RNA polymerase RDR6/SGS3/DRB4 complex. This dsRNA is then diced in a phased manner, mostly by DCL4, producing a set of 21-nt-long siRNAs. From this siRNA population, only a few sequences are conserved evolutionary. The conserved tasiARF is loaded into AGO1 and post-transcriptionally regulates the transcripts of *ARF3* and *ARF4* transcription factor genes, the protein product of which negatively regulates auxin-dependent gene expression [20,23,30]. The tasiARF is cell-to-cell mobile and forms a gradient across the leaf blade being more abundant on the adaxial side [18,31]. This results in the confinement of *ARF4* in the abaxial side of the leaf and contributes to the determination of abaxial cell identity.

Another sRNA-mediated signaling pathway that is conserved among land plants is the miRNA-regulation of the large family of *NUCLEOTIDE-BINDING LEUCINE-RICH REPEAT CONTAINING RECEPTOR* genes (NB-LRRs) which are important regulators of innate immunity. They recognize specific pathogen effectors and trigger resistance responses [32,33]. The members of this gene family are targeted by various, unrelated, lineage-specific 22-nt-long miRNAs, some of which are more widespread (i.e. the miR482/2118 superfamily) [34], while others can be found only in certain plant groups [35]. These genes are constantly repressed by the above-mentioned miRNAs but relieved from repression when the level of the regulatory miRNA is lowered upon pathogen infection. Interestingly, in the presence of a symbiont, the miRNA level is changed in the opposite direction leading to an even lower

level of NB-LRR genes and enhanced nodulation in *Medicago truncatula* [36]. The miRNA cleavage of the transcripts can initiate the production of secondary phasiRNAs: dsRNA intermediates are produced by the RDR6/SGS3/DRB4 complex and then diced by DCL4/DCL2, similarly to the case of *TAS3*. These secondary siRNAs can also target related genes amplifying the immune response [37].

There are other tasiRNA- or phasiRNA-producing loci in plants that are less conserved as the *TAS3* regulatory module [30]. In *Arabidopsis thaliana*, *TAS1a,b,c*, and *TAS2* noncoding RNA are cleaved by miR173-loaded AGO1 [38], while *TAS4* is cleaved by miR828 [39] and secondary siRNAs are produced as described above. Some siRNAs emerging from these loci target host genes: *TAS1*-derived siRNAs can target *HEAT-INDUCED TAS1 TARGET 1* and *2 (HTT1/2)* [40], *TAS2*-derived siRNAs can target several *PENTATRICOPEPTIDE REPEAT-CONTAINING PROTEIN (PPR)* genes [30], while a *TAS4*-derived siRNA can target a set of MYB transcription factor genes [39]. Other secondary phasiRNA-producing loci are mostly protein coding genes and include the miR393-initialized auxin receptor genes *TRANSPORT INHIBITOR RESPONSE 1 (TIR1)* and *AUXIN SIGNALING F-BOX 1–3 (AFBI–3)* that regulate auxin signaling [41].

The presence of the virus also triggers the synthesis of a different class of siRNAs (virus-activated siRNAs, vasiRNAs) that regulate other host genes and modulate the antiviral response of the plant [42], for example, by downregulating genes coding for different functions of the photosynthetic apparatus. Lowering the energy production rate within the virus-infected plant cell would restrict the replication of the virus which requires a lot of energy [43]. This conserved regulatory action of vasiRNAs against the host metabolism could explain the observation that almost all virus-infected plant show some degree of retarded growth. The production of vasiRNAs was shown to be dependent on RNA-DEPENDENT RNA POLYMERASE 1 (RDR1) and DICER-LIKE 4 (DCL4) [42] making them unique among other endogenous siRNAs like trans-acting siRNAs (tasiRNAs) or phased siRNAs



(phasiRNAs) which depend on the action of RDR6 and DCL4/DCL2. Unlike tasiRNAs, vasiRNAs do not require a phase initiating miRNA- or siRNA-mediated cleavage and, therefore, they are not produced in a phased manner. Pitzalis et al [44] proposed a mechanism for the triggering of vasiRNA production: virus infection could induce a disruption or overload of RNA decay pathways, thereby leading to the accumulation of aberrant transcripts that are prone to dsRNA formation and processing into siRNAs. Most of the vasiRNAs are 21-nt-long [42] but in some cases, for example, during infection by the pararetrovirus CaMV, a significant amount of 22-nt-long vasiRNAs are also produced in three different species of the *Brassicaceae* family, probably due to the activity of DCL2 [43]. The vasiRNAs can target the cleavage of the same mRNA that they were derived from or other genes *in trans*, similarly to tasiRNAs.

Viral infections often lead to severe diseases in plants that result in significant economic loss. Tobamoviruses are among the most studied plant viruses because they are widespread and infect economically important plants, including the large family of *Solanaceae* (i.e., potato, tomato, pepper, eggplant, etc.) and *Brassicaceae* (i.e., cruciferous vegetables like cabbage, turnip, mustard, rapeseed, etc.). The model plant of *Brassicaceae* is *Arabidopsis thaliana* (thale cress) which has several geographical varieties called ecotypes. We previously characterized a new isolate of *Turnip vein-clearing virus* (TVCV-ApH) [45], a crucifer-infecting tobamovirus [46,47], that caused severe leaf deformation (upward curling leaves) on the Bur ecotype of *Arabidopsis thaliana*, but not on Col ecotype. The 6.3 kb genome of TVCV-ApH contains four ORFs. ORF1 codes for the 125 kDa small replicase subunit, while the suppression of an amber termination codon of ORF1 results in the 182 kDa large replicase subunit (ORF2). The ORF3 and ORF4 translate from 3' subgenomic RNAs and result in a 30 kDa movement protein, and a 17 kDa coat protein, respectively [45]. Several reports demonstrated that besides its role in replication, the p125 protein acts as a viral silencing suppressor through the binding of vasiRNAs [48–50]. We also observed this

ecotype-specific leaf deformation in the case of turnip mosaic virus (TuMV) infection. TuMV is a member of the *Potyviridae* family and it is the most important virus infecting cruciferous crops [51]. The TuMV genome is around 9.8 kb and the genomic RNA is translated into a large polyprotein and a frame-shift protein. The large polyprotein is subsequently processed by the action of viral-encoded HCPro protease into ten mature functional products [52,53]. A frame-shift protein, P3N-PIPO, was reported to be involved in the pathogenesis and movement of TuMV [54,55]. The TuMV HCPro protein has RNA silencing suppressor (RSS) activity [56] through the binding of vsRNAs [57,58]. The ecotype-specific leaf deformation phenotype observations with two different RNA virus infections suggested that there are genetic differences in the hosts that determine this response. We also expected that this differential regulation will manifest in the transcriptome or sRNAome profiles and by comparing these profiles we will be able to find the host factors responsible for the ecotype-specific symptom development. Indeed, the comparative transcriptome analysis of the TVCV-, TuMV-, and mock-infected Col and Bur plants resulted in a few candidate genes that could be responsible for the leaf deformation. Some of these genes are known regulators of leaf polarity while others may represent novel, ecotype-specific functions in leaf development or other biological processes that are commonly disturbed by TVCV and TuMV infection.

## **Materials and methods**

### **Plant growth conditions and viral infection**

Seeds of *Arabidopsis thaliana* Col and Bur ecotypes were sown on soil in pots, watered, wrapped with cling film, and kept in the dark for two days. After that, pots were moved into a growth chamber (SANYO/Panasonic MLR-352-PE) where plants grew under long-day conditions (16 h light, 8 h dark, constant 21 °C) for two weeks. Then, individual seedlings

were transferred into Jiffies ([www.jiffypot.com](http://www.jiffypot.com)) and grown for further two weeks in the same chamber. The four-week-old seedlings were transferred into pots filled with soil and moved into a lightroom where they grew for one day under long-day conditions. The next day, in the case of TVCV-ApH, plants were inoculated as described before [45]. Briefly, 2  $\mu\text{L}$  (1  $\mu\text{g}$ ) of either total RNA purified from systemic leaves of *Nicotiana benthamiana* infected with an infectious *in vitro* transcript of TVCV-ApH or water (mock) was combined with 5.5  $\mu\text{L}$  water and 7.5  $\mu\text{L}$  2 $\times$  inoculation buffer and was rub-inoculated into the true leaves of four-week-old *Arabidopsis* plants. In the case of TuMV, *Arabidopsis* plants were infected with extracts of *N. benthamiana* leaves infected with an infectious *in vitro* transcript of TuMV. For both viruses, three leaves were inoculated per plant and 60 plants were inoculated per sample (Col mock, Col virus, Bur mock, and Bur virus). After inoculation, the plants were sprayed with water and grown under long-day conditions in the same lightroom.

## **Purification of RNA samples for high-throughput sequencing**

For total RNA purification, two symptomatic systemic leaves were collected from every infected plant 14 day-post-inoculation (dpi) between ZT4 and ZT6. For the mock samples, leaves of similar age as the infected ones were collected. Total RNA was purified from individual leaves (480 samples altogether) using a phenol-chloroform extraction method [59]. The purified RNA samples were quantified by Nanodrop photometer and 1  $\mu\text{g}$  of total RNA per sample was run on a denaturing agarose gel. The gel was stained with ethidium bromide and photographed with ChemiDoc™ (Bio-Rad, Hercules, CA, USA). The relative amount of the viral RNAs that were visible over the ribosomal RNA bands was quantified with Image Lab v5.1 software (Bio-Rad, Hercules, CA, USA) using the 18S rRNA bands as the reference. Seven samples with similar levels of viral genomic RNAs were pooled together to form a biological replicate. In this way, four replicates were created.

## **High-throughput sequencing of mRNA and sRNA populations**

The libraries for the RNA-seq and sRNA-seq from the TVCV-infected samples were prepared and sequencing reactions were carried out by LC Sciences LLC (Houston, TX, USA). The polyA-selected RNA-seq libraries were prepared using the TruSeq Stranded mRNA Library Prep Kit (Illumina, San Diego, USA) and sequenced with a paired-end 2×150 bp chemistry on Illumina HiSeq X platform. The sRNA-seq libraries were prepared using the TruSeq Small RNA Sample Prep Kit (Illumina, San Diego, USA) and sequenced on an Illumina HiSeq X platform with a single-end 50 bp chemistry.

PolyA-selected, stranded RNA-seq libraries for the TuMV samples were prepared and sequenced by Novogene Co., Ltd. (Cambridge, UK) on a NovaSeq 6000 sequencing platform with a paired-end 2×150 bp chemistry. The sRNA libraries (including a size-separation by PAGE) were prepared in our laboratory using the TruSeq Small RNA Sample Prep Kit (Illumina, San Diego, USA) and were sequenced by Novogene Co., Ltd (Cambridge, UK) on a NovaSeq 6000 platform with a single-end 50 bp chemistry.

The raw sequencing data were deposited in the NCBI SRA database (BioProject: PRJNA788379) under the identifiers SRR17227515–SRR17227578.

## **Bioinformatic analysis of high-throughput sequencing data**

### **RNA-seq analysis**

The quality of the raw reads was checked with FastQC v0.11.8 [60]. Trimming of adapter sequences and filtering of low-quality reads were carried out with cutadapt v2.8 [61]. The clean reads were aligned to the AtRTD2 reference transcriptome [62] and the normalized expression values (transcript per million, TPM) were calculated with kallisto v0.44.0 [63] with the following parameters: -b 10 --bias --rf-stranded. The TPM values calculated for the alternative transcript isoforms were aggregated to represent gene-level expression and were

normalized between samples with sleuth v0.29.0 [63]. To test whether there is a significant difference between the means of the Bur (or Col) virus-infected and all the other samples, a Wald test was applied. The calculated  $P$ -values were corrected for multiple testing using the Benjamini-Hochberg method [64]. Genes with  $qval$  (corrected  $P$ -value, at 1% false discovery rate)  $< 0.01$ ,  $b$  (estimated fold-change)  $> \log(2)$ , and  $mean\_obs$  (mean raw expression level)  $> 1$  were considered as differentially expressed genes (DEGs).

The DEG expressions were visualized in a heatmap, for which  $Z$ -scores were calculated in the following way: for every gene, the mean and standard deviation of the expression across samples were calculated, then from every individual value the mean was subtracted and the resulting value was divided by the standard deviation. The  $Z$ -score tells how many standard deviations an expression value is from the mean expression of an individual gene. In this way, genes with different mean expression levels can be compared. The heatmaps were generated with the `geom_tile` function of the `ggplot2` R package [65].

Principal component analysis of the RNA-seq data was performed using the `prcomp` R package. Normalized gene expression values (transcript per million, TPM) were scaled ( $\log_2$ -transformed) and centered before the analysis. The plots were generated with the `autoplot` function of the `ggfortify` R package [66].

Comparison of the DEG sets from different conditions was analyzed with the `UpSet` R package [67].

The Gene Ontology (GO) term analysis of the DEGs was performed using the PlantRegMap server [68] with the default settings. The enrichment factor was calculated by dividing the observed number of genes in a particular category by the expected number of genes in that category. The plot was generated with the `geom_point` function of the `ggplot2` R package [65].

## Small RNA-seq analysis

The raw sequences were processed and filtered using cutadapt v2.8 [61] with the following parameters: `-a TGG AATTCTCGGGTGCCAAGG -m 20 -M 25 -q 20 --max-n=0 --discard-untrimmed`. The trimmed sequences were further filtered to remove tRNA- and rRNA-derived sequences using the short read aligner bowtie2 v2.3.5.1 with the following parameters: `-k 1 -D 20 -R 3 -N 1 -L 10 -i S,1,0.50` and a reference sequence set containing tRNA, rRNA, snRNA, and snoRNA sequences from *Viridiplantae*. The unaligned sequences were collected and aligned to the *Arabidopsis thaliana* TAIR10 reference genome with ShortStack v3.8.5 [69] with the following parameters: `--align_only --bowtie_m all --ranmax none --keep_qual`. The raw count table of the genome-mapped sRNA sequences was produced using the `fasta` and `view` functions of the samtools suite v1.9 [70] and custom shell scripts. The sRNS-producing loci were predicted using ShortStack v3.8.5 with the following parameters: `--dicermin 20 --dicermax 25`. The normalization and differential expression analysis of the sequences or the sRNA loci were performed with the DESeq2 R package [71]. Only those sequences/loci were accepted that passed through the following filter: *padj* (corrected *P*-value, at 1% false discovery rate) < 0.05, *log2FoldChange* > log(1.5), and *baseMean* (mean raw expression level) > 1. The DESeq2 normalization factors were used for the calculation of the sequence length distribution plot.

The sRNA target analysis was performed using the psRNATarget server [72] with the default settings except that the target accessibility calculation was allowed. We provided the 114 Bur virus upregulated sRNAs and the 88 Bur virus downregulated transcripts that were extracted from the AtRTD2 transcripts [62] using the `notseq` function of the EMBOSS package v6.6 [73]. Only the predicted cleavage hits were considered for further analysis.

## RT-qPCR analysis

The primers to measure *ARF4* (AT5G60450) expression (*ARF4\_qPCR\_F*: 5'- GCT CCT CTT GAC TAC AAA CAA CAG -3', *ARF4\_qPCR\_R*: 5'- GGC GAA ACT TCC ACT CTA CTC C -3') were designed with PerlPrimer v1.2.4 [74] in a way that at least one primer must span an intron to prevent amplification from genomic DNA. The *FIL* (AT2G45190) was amplified using the primers described before (*FIL\_qPCR\_F*: 5'- TGG TAC AGC AAC CAC ATC GGA CAG -3', *FIL\_qPCR\_R*: 5'- GCC AAA CCA TCC TTG CGG TTA ATG -3') [75]. The *PDF2* gene (AT1G13320) was used as an internal reference, primers to measure *PDF2* were published earlier (*PDF2\_qPCR\_F*: 5'- TCA TTC CGA TAG TCG ACC AAG -3', *PDF2\_qPCR\_R*: 5'- TTG ATT TGC GAA ATA CCG AAC -3') [76].

For the first-strand cDNA synthesis, 1 µg of total RNA was treated with DNase I, then reverse transcribed with RevertAid H Minus Reverse Transcriptase (Thermo Fisher Scientific, Waltham, MA, USA) and a random primer according to the manufacturer's instructions. The cDNAs were diluted ten times, and 1 µL was used in a 10 µL qPCR. The qPCRs were performed using SYBR™ Select Master Mix (Thermo Fisher Scientific, Waltham, MA, USA), following the manufacturer's instructions. The reactions were run in a LightCycler® 96 Real-Time PCR machine (Roche, Basel, Switzerland). The reaction profile was the following: preincubation at 95 °C for 1 min; amplification at 95 °C for 15 s and 60 °C for 30 s, repeated 45 times; melting at 95 °C for 10 s, then 65 °C for 1 min and continuous heating to 97 °C; cooling at 37 °C for 30 s. Four biological and two technical replicates were measured per sample. The values for every biological replicate were calculated as the mean of the two technical replicates. The relative expression values were calculated by the LightCycler96® v1.1 software (Roche, Basel, Switzerland). Unpaired, two-tailed Student's *t*-tests (at 0.95 confidence level) were performed using the *t.test* R package to assess if there

is a significant difference between the means of the mock- and virus-infected samples in the Col and Bur plants either in the TVCV or TuMV series.

## **QTL mapping**

To obtain a coarse genetic map, a publicly available Minimal Recombinant Inbred Line set (INRA, Versailles, France) derived from a cross of *Arabidopsis thaliana* Bur and Col ecotypes was amended with 23 more lines selected randomly from the remaining lines [77]. Altogether, 43 lines and the two parental lines were used in the experiment. Five plants per line were infected either with TVCV-ApH or mock as described above and the leaf deformation (upward curling) was scored from 0 to 5 (0 = infected Col, 5 = infected Bur). The average score of the five plants was recorded and used for assessing correlation with the genotypes of the RILs that were obtained from INRA (<http://publiclines.versailles.inrae.fr/page/20>). The analysis was carried out with Windows QTL Cartographer v2.5\_011 [78] using the single marker association test.

## **Results**

### **TVCV and TuMV infection causes leaf deformation only on Bur ecotype**

We observed earlier that upon infection with turnip vein-clearing virus (TVCV, a tobamovirus), the Bur ecotype of *Arabidopsis thaliana* displayed much more severe symptoms than Col [45]. The most obvious symptom was the leaf deformation which included polarity defect (upward curling leaves) and wavy leaf margin that could not be observed on Col plants under the conditions tested (Fig 1). We were curious whether this difference in symptom development is specific to this virus-host interaction or can be



observed with other viruses as well. Therefore, we infected the above-mentioned two *Arabidopsis* ecotypes with turnip mosaic virus (TuMV, a potyvirus). We observed the same difference in symptom development between the two ecotypes, namely, the Bur plants showed upward curling leaves while the Col plants did not have leaf deformity (Fig 1).

**Fig 1. Viral symptoms of TVCV- and TuMV-infected *Arabidopsis thaliana* Col and Bur ecotypes.**

Four-week-old plants grown under long-day conditions at 21°C were infected either with virus or mock solutions. After two weeks, plants were photographed and samples for RNA purifications were collected from mock- and virus-infected plants. For the RNA purification, samples were taken from symptomatic leaves of Bur or leaves of similar developmental stages from the Col and mock samples. To obtain one replicate, seven leaves from different plants were pooled (one leaf per plant). Four replicates were prepared in such a way for every mock and virus-infected sample.

Both viruses replicated and reached similar viral RNA levels in each ecotype, suggesting that viral replication, short-distance, and long-distance viral movement are not involved in the observed symptom differences. The differential leaf development defect in the two ecotypes upon both virus infections suggests that there are host factors whose ecotype-specific misregulation leads to an altered leaf developmental program only in the Bur plants.

**TVCV and TuMV infection elicit different transcriptome responses in Col and Bur ecotypes**

Upon viral infection, the transcriptome of the infected host changes significantly [79–83]. The virus modulates the host environment in favor 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 virus similarly in many ways. However, since there were significant differences in the observable symptoms, we anticipated that there would be gene expression changes that reflect these differences. In order to reveal these differential gene expression profiles, we collected leaves from TVCV- and TuMV-infected *Arabidopsis Col* and *Bur* plants along with mock-infected ones in four biological replicates for RNA-seq experiments and purified total RNA from them. Next, the polyA-tail mRNA fraction was sequenced either on an Illumina HiSeq X (TVCV) or a NovaSeq 6000 (TuMV) platform. The clean reads were mapped to the AtRTD2 *Arabidopsis thaliana* reference transcriptome [62] with kallisto [63] and the gene expression values were normalized between samples with sleuth [63] (Table S1). We performed a principal component analysis to reveal which experimental variables are responsible for the most variances between samples (Fig 2).

**Fig 2. Principal Component Analysis of the RNA-seq data.**

PCA was performed using the prcomp R package. Normalized gene expressions (Transcript Per Million, TPM) were scaled and centered before the analysis. (A) The first component (PC1) separates the samples by ecotype and explains 36.85% of the total variances between the samples while the second component (PC2) separates the samples by virus series and explains further 23.35% of the variances. (B) The third component (PC3) separates samples that are either mock or virus-infected, which is responsible for a further 8.06% of the variances. The four biological replicates are marked with the same color for easier identification.

According to the analysis, the most important factor that separates samples and is responsible for 38.23% of the total variances (PC1) is the ecotype, while the second most important factor (PC2) separates the TVCV and TuMV series (Fig 2A). This latter is either

because of the different technology that was used for sequencing the two series in two different laboratories or could reflect a true biological difference (or both). The third most important factor (PC3) that separates the mock- and virus-infected samples is responsible only for 8.06% of the total variances observed between samples (Fig 2B). This figure also shows that the mock- and virus-infected samples in the Bur ecotype are more separated than in the Col ecotype (Fig 2B). We interpret these results that there is a substantial difference between the two ecotypes in their responses to viral infection.

## **Finding the genes responsible for the common leaf polarity**

### **defect in the TVCV- and TuMV-infected Bur plants**

To find out which genes are responsible for the upward curling leaves in the TVCV- or TuMV-infected Bur ecotype, we performed a differential gene expression analysis using sleuth [63], separately for the two virus series. We expected that genes responsible for the leaf deformation that was observed only in the virus-infected Bur ecotype will either be induced or reduced only in the Bur virus samples both in the TVCV and TuMV series compared to all the other samples in the series, namely, Bur mock, Col mock and Col virus samples. Alternatively, it is possible that genes are preferentially induced or reduced in the Col virus samples that prevent symptom development. We considered genes differentially expressed only if the  $Q$ -value from the statistical test was lower than 0.01 (1% false discovery rate), the absolute value of the effect ( $b$  parameter, or estimated fold-change, log-transformed) was higher than  $\log(2)$ , and the mean raw expression ( $mean\_obs$ ) was higher than 1. In this way, we got a list of the differentially expressed genes for the Bur virus or Col virus samples for the TVCV and TuMV series (Fig 3).

**Fig 3. Expression pattern of the differentially expressed genes between Bur and Col plants infected either with TVCV or TuMV.**

To get a list of differentially expressed genes (DEGs) that significantly change only in the Bur (or Col) virus-infected samples, a Wald-test was applied (i.e., Bur virus samples against all the others). Only genes with  $qval < 0.01$ ,  $b$  (estimated fold-change)  $> \log(2)$  and  $mean\_obs$  (mean raw expression level)  $> 1$  were accepted as DEGs. The number of DEGs is noted in the title. Normalized expression values (transcript per million, TPM) of individual, differentially expressed genes in (A) TVCV and (B) TuMV-infected samples were centered by calculating Z-scores. Z-score tells how many Standard Deviations an expression value is from the mean expression of an individual gene. In this way, genes with different mean expression levels can be compared. The expressions of DEGs in samples are represented in a heatmap. The expressions of the four biological replicates are shown for every sample.

There were 278 genes that were up- and 548 that were downregulated in the Bur virus samples, while 386 and 162 genes were up- and downregulated in the Col virus samples in the TVCV series (Fig 3A, Table S2). These numbers were 184, 317, 456, and 115 in the TuMV series, respectively (Fig 3B, Table S2). To filter for the genes that were regulated similarly in the two virus series, we performed a set analysis with the UpSet R package (Fig 4).

**Fig 4. Commonly regulated genes in TVCV and TuMV-infected plants.**

Sets of DEGs in eight different conditions were compared and visualized using the UpSet R package. The vast majority of the DEGs were specific to one condition (represented by single dots), while the number of common elements between the different conditions (connected dots) was much smaller. The most abundant common elements were observed between the sets of downregulated genes in Bur virus samples both in TVCV and TuMV-infected plants

(blue) and the upregulated genes in Bur virus samples both in TVCV and TuMV-infected plants (orange). All the other common elements were much smaller in size, sets with zero elements are not shown.

According to this analysis, the majority of the differentially expressed genes were unique to a certain set and probably not responsible for the commonly observed leaf polarity misregulation. The most abundant intersection with 88 genes was the one that contained downregulated genes in Bur virus samples both in the TVCV and TuMV series (Table S3b), while the second most abundant with 36 genes was the one containing the upregulated genes in Bur virus samples in both virus series (Table S3a). All the other intersections were much smaller and not relevant for our investigation. We performed a GO term enrichment analysis with the above-mentioned two gene sets using the PlantRegMap server [68]. Among the upregulated genes, categories were enriched with moderate or low significance, related to plant cell wall functions (GO:0071554, GO:0071555, GO:0071669, etc), cellular developmental process (GO:0048869), response to hormone (GO:0009725), cell differentiation (GO:0030154) while among the downregulated genes, several terms related to photosynthesis and light regulation (GO:0009768, GO:0009765, GO:0019684, etc), hormone metabolic process (GO:0042445), regulation of hormone level (GO:0010817), regionalization (GO:0003002), pattern specification process (GO:0007389), etc., were significantly enriched (Fig 5, Table S3c–d).

**Fig 5. GO term enrichment analysis of common DEGs of TVCV and TuMV-infected Bur plants.**

The common DEGs of the TVCV and TuMV-infected Bur plants (orange and blue colored sets in Fig 4) were subjected to GO-term enrichment analysis using the PlantRegMap server with the default settings. Only terms in the Biological Process category are shown, the results

of the full analysis can be found in Table S3c–d. Enrichment was calculated by dividing the observed number of genes in a particular category by the expected number of genes in that category.

Interestingly, the latter two categories (GO:0003002: regionalization and GO:0007389: pattern specification process) in the downregulated genes contained *FIL* (*FILAMENTOUS FLOWER*, AT2G45190) that encodes a member of the YABBY family of transcriptional regulators involved in abaxial cell type specification in leaves [84]. Actually, *FIL* is the most significantly differentially down-regulated (in the Bur virus samples) gene in the common set of the TVCV and TuMV series (Table S3b). *FIL* directly regulates *ARF4* (*AUXIN RESPONSE FACTOR 4*, AT5G60450) [85] that was found in the response to hormone and the cellular developmental process GO categories in the enriched sets of upregulated genes. Furthermore, the important leaf abaxial identity determinant *ARF4* is known to be post-transcriptionally regulated by tasiARF-guided AGO1, as described in the Introduction part.

## **Virus-responsive sRNA profiles of the ecotypes are different**

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. We performed a Principal Component Analysis of the normalized sRNA expression values to check the quality of our dataset (Fig 6).

### **Fig 6. Principal Component Analysis of the sRNA-seq data.**

PCA was performed using the prcomp R package. DESeq2-normalized sRNA expressions were scaled and centered before the analysis. (A) The first component (PC1) separates the samples by virus series and explains 29.1% of the total variances between the samples while the second component (PC2) separates the samples by ecotype and explains further 19.88%

of the variances. (B) The third component (PC3) separates samples that are either mock or virus-infected, which is responsible for a further 10.46% of the variances. The four biological replicates are marked with the same color for easier identification.

According to the PCA, the first principal component separates samples by the virus series which could either mean that the two viruses elicit quite different responses or could reflect the different sample preparation and sequencing platforms that were used for the generation of data (or both). The second principal component separates the samples by ecotype, while the third one by treatment. The substantial differences in the responses of the two ecotypes to the two viruses suggested that there would be a limited number of commonly regulated sRNAs that could be responsible for the observed common symptoms.

Next, we analyzed the length distribution of the sRNAs in the samples. 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 switches: the 21–22-nt size classes become overwhelmingly dominant [13,86,87,44]. This is mainly due to the 21- and 22-nt viral siRNAs that are produced by DCL4 and DCL2 enzyme activities, respectively, and amplified by RDR1 or RDR6 in *Arabidopsis* [5,88–90]. Analysis of our sRNA libraries showed the same shift in size classes. However, when we filtered out the virus-derived vsRNAs by mapping the sRNA sequences to the *Arabidopsis* genome, the remaining host-derived sRNAs also showed a tremendous 21-nt size class enrichment in the virus-infected samples (Fig 7).

**Fig 7. Sequence-length distribution of the host-derived sRNAs in Bur and Col plants infected either with TVCV or TuMV.**

The raw sRNA sequences were first adapter-trimmed, quality- and length-filtered, then rRNA and tRNA sequences were removed. The remaining sRNA sequences were mapped to the

*Arabidopsis thaliana* TAIR10 reference genome with ShortStack. The raw abundances of the genome-mapped sequences were normalized with DESeq2. The normalization factors were used to scale the raw abundances of the size classes measured by FastQC. Error bars represent the standard error of the mean of four biological replicates.

This can either be attributed to the mostly 21-nt-long miRNAs or host-derived, also vastly 21-nt-long virus-activated siRNAs (vasiRNAs), which are produced in an RDR1 and DCL4-dependent manner from plant precursor transcripts [42]. The vasiRNAs regulate host genes that can modulate the antiviral response of the plant. We hypothesized that some vasiRNAs may be produced in an ecotype-specific manner and can regulate the genes that we have found in the comparative transcriptome analysis. To find sRNAs whose expression either increased or decreased only in the Bur virus samples, we performed a differential expression analysis of the genome-mapped, clean sRNAs which had a mean read count  $> 1$  (102195 unique, 20–25-nt-long sequences) using DESeq2. We applied a filtering rule that was less strict than in the case of RNA-seq analysis ( $\log_2\text{FoldChange} > \log_2(1.5)$ ,  $\text{padj} < 0.05$ ,  $\text{baseMean} > 0$ ) because according to our experience, sRNA expressional changes are usually smaller and the data are noisier. In this way, we got sets of differentially expressed sRNAs in the Bur (or Col) virus samples in the TVCV and TuMV series (Fig 8).

**Fig 8. Expression pattern of the differentially expressed sRNA sequences between Bur and Col plants infected either with TVCV or TuMV.**

To get a list of differentially expressed sRNAs (DESS) that significantly change only in the Bur (or Col) virus-infected samples, a Wald-test was applied (i.e., Bur virus samples against all the others). Only sequences with  $\text{padj} < 0.05$ ,  $\text{Log}_2\text{FoldChange} > \log(1.5)$ , and  $\text{baseMean}$  (mean raw expression level)  $> 0$  were accepted as DESS. The number of DESS is noted in the title. Normalized expression values of the individual, differentially expressed sequences in



(A) TVCV and (B) TuMV-infected samples were centered by calculating Z-scores. Z-score tells how many Standard Deviations an expression value is from the mean expression of an individual sequence. In this way, sequences with different mean expression levels can be compared. The expressions of DESs in samples are represented in a heatmap. The expressions of the four biological replicates are shown for every sample.

The analysis revealed that sRNAs tended to be upregulated upon virus infection rather than downregulated: there were 2266 sequences that were selectively upregulated and 38 downregulated in the Bur virus samples in the TVCV series, and 521 and 48 in the TuMV series. These numbers were similar regarding the selectively up- and downregulated sRNAs in the Col virus samples in both virus series (Fig 8). The majority of upregulated sRNAs in both virus series were 21–22-nt-long while the downregulated sRNAs were mainly 23–24-nt-long (Table S4).

To find commonly regulated sRNAs between the TVCV and TuMV series, we analyzed the differentially expressed sRNAs with the UpSet R package like in the case of RNA-seq data (Fig 9).

**Fig 9. Commonly regulated sRNA sequences in TVCV and TuMV-infected plants.**

Sets of differentially expressed sRNA sequences (DES) in eight different conditions were compared and visualized using the UpSet R package. The vast majority of the DESs were specific to one condition (represented by single dots), while the number of common elements between the different conditions (connected dots) was much smaller. Blue color marks the sRNA sequences that are downregulated in Bur virus samples both in TVCV and TuMV-infected plants while orange color marks the sRNA sequences that are upregulated in Bur virus samples both in TVCV and TuMV-infected plants. Sets with zero elements are not shown.

We focused on the sets that have common sequences between the TVCV and TuMV Bur virus samples, either up- or downregulated to find potential candidates that can regulate the DEGs identified in the RNA-seq analysis. We found 133 upregulated and one downregulated sRNA sequence that was common between the TVCV and TuMV Bur virus samples. Annotation of these sequences revealed that most of the 133 upregulated sequences are 21-nt long and associated with intergenic regions, transposable elements, *TAS1*, and *TAS2* genes, some protein-coding genes and pseudogenes, and some miRNA genes (*MIR158B*, *MIR162B*, *MIR390A*, *MIR396A*, *MIR835A*, *MIR841A*, *MIR846A*), although only one mature miRNA sequence was found (ath-miR846-3p), all the other MIR-associated sequences are iso-miRs or siRNAs. It is worth mentioning that 41% of the sequences contain one mismatch (including the *TAS1C*- and *TAS2*-derived, and miR390, miR396, miR835a, miR841a sequences) which suggests that they are Bur-specific sequences. We checked if the 133 Bur virus upregulated sequences can be found in the Col virus upregulated sequences allowing 2 mismatches and we found that 19 sequences were also present in the Col set, so we removed them. We were curious if the remaining 114 Bur virus upregulated sRNAs can target the 88 Bur virus downregulated genes identified in the RNA-seq analysis, therefore, we performed a target prediction using the psRNATarget server [72]. According to the prediction, 86 out of the 114 sRNAs can potentially regulate 69 out of the 88 genes (Table S5). When we performed a GO term enrichment analysis of the potentially vasiRNA-regulated genes, the regionalization and pattern specification categories remained significantly enriched (Fig S1). We found that a transposon-derived 21-nt-long siRNA (5'- AUU GCA GAG AUG GAU GUA CAA -3') could potentially target *FIL in trans*. The *FIL*-targeting sequence is identical in the two ecotypes but induced only in the Bur virus sample. One other gene in the GO categories regionalization (GO:0003002) and pattern specification process (GO:0007389) is also potentially regulated by a vasiRNA: *DOT3 (DEFECTIVELY ORGANIZED*

*TRIBUTARIES 3*, AT5G10250) is targeted by an intergenic siRNA that contains one mismatch compared to the Col reference sequence, but the Col version is not induced in the Col virus sample. The most significantly enriched categories are related to photosynthesis and chloroplast functions. This is consistent with the previous observations that vasiRNAs target genes coding for different members of the photosynthetic apparatus to inhibit the replication of viruses by tuning down the energy production of the plant cell [43]. Apparently, there is an ecotype difference in this response that could contribute to the symptom severity.

The miR396 has been described to target the GRF family of transcription factors that are required for coordination of cell division and differentiation during leaf development in *Arabidopsis* [91]. The GRFs are not among the 88 downregulated genes we found. We checked their expression pattern in our data but they do not show any consistent pattern. There might be other genes related to leaf development directly or indirectly that could be regulated post-transcriptionally by vasiRNAs.

The one commonly downregulated sequence is the tasiARF that regulates *ARF4* which was found in the 36 Bur virus upregulated genes (Table S4).

## **tasiARF expression anti-correlates with the expression of its target *ARF4***

The finding that tasiARF is the only commonly downregulated sRNA between the TVCV and TuMV Bur virus samples prompted us to look up the expression profile of the tasiARF-related pathway components, namely, the phase-initiating miR390 that targets *TAS3* (the precursor of tasiARF), *AGO7*, the protein product of which binds miR390, *RDR6*, *SGS3*, *DRB4*, and *DCL4*, the protein products of which are involved in the synthesis and dicing of the dsRNA precursor, the tasiARF, and its target *ARF4*, the identified vasiRNA that

potentially targets *FIL*, and *FIL* itself that together with *ARF4* are involved in the leaf polarity determination (Fig 10).

**Fig 10. Expression pattern of the *TAS3* regulatory module components.**

The normalized expression values of the sRNAs and genes involved in the *TAS3* regulatory module are shown in the TVCV and TuMV series. Error bars represent the standard error of the mean of four biological replicates. The actual data points are also shown.

Analyzing Fig 10, we can observe the following: miR390 is virus-induced in both ecotypes, especially by TVCV infection. The *AGO7*, *RDR6*, and *DRB4* mRNA levels did not change significantly, while the *SGS3* was induced only in the Bur virus samples in both virus series. *DCL4* was induced in the Bur virus sample but the induced level was the same as the levels in the Col mock and virus samples. The tasiARF levels significantly drop in the Bur virus samples both in the TVCV and TuMV series. The expression profile of *ARF4* is anti-correlated with the tasiARF and the *FIL* profile in both virus series. The newly identified vasiRNA-FIL is dramatically induced in the Bur virus samples both the TVCV and TuMV series and anti-correlates with the *FIL* expression. The *ARF4* and *FIL* expression levels were validated by RT-qPCR (Fig 11).

**Fig 11. Validation of the *ARF4* and *FIL* expression patterns with RT-qPCR.**

(A) RT-qPCRs were performed on the samples used for the high-throughput sequencing. Four biological and two technical replicates were measured per sample. The values for every biological replicate were calculated as the mean of the two technical replicates and shown as dots in the figure. The error bars represent the standard error of the mean of four biological replicates. The *PDF2* gene (AT1G13320) was used as an internal reference [76]. (B) Unpaired, two-tailed Student's *t*-tests at 95% confidence intervals were performed to assess

if there is a significant difference between the means of the mock and virus-infected samples in the Col and Bur plants either in the TVCV or TuMV series. The resulting *P*-values for every pair tested are indicated in the table. Values less than 0.05 indicate a significant difference. (C) Expression pattern of the unspliced pre-mRNA transcript of *ARF4* (genomic or *gARF4*) in the RNA-seq data. The error bars represent the standard error of the mean of four biological replicates. The individual data points are also shown.

To decide whether the observed elevated level of *ARF4* is due to a transcriptional (i.e., induction by a transcription factor like *FIL*) or post-transcriptional regulation (i.e., by tasiARF), we checked the expression level of the unspliced pre-mRNA transcript of *ARF4* by amending the reference transcriptome with the genomic sequence of the *ARF4* gene (*gARF4*) and performed an RNA-seq analysis as described before. According to this analysis, the *gARF4* transcript levels were very low in all samples and no significant change was observed that could explain the elevated level of *ARF4* in the Bur virus sample (Fig 11C). This suggests that the observed *ARF4* induction is a result of decreased post-transcriptional regulation by tasiARF.

Taken together, the induction of miR390 and *SGS3* would suggest a higher level of *TAS3* processing and eventually a higher level of tasiARF, but the case is the opposite: the tasiARF production is compromised during virus infection, especially in the Bur ecotype. This observation suggests that the virus blocks one or more key steps of phased siRNA production. This can either be the miR390-mediated cleavage of the *TAS3* ncRNA or the RDR6/*SGS3*-mediated dsRNA precursor synthesis, or the phased cleavage of the dsRNA precursor by DCL4. The miR390-mediated cleavage is specific to the *TAS3* pathway, while the other components are shared by other tasiRNA-producing processes.

## **Viral infection induces the production of tasiRNAs except for the tasiARF**

In *Arabidopsis thaliana*, there are eight *TAS* loci of four families and many other phased siRNA-producing loci. Most of the sRNAs emerging from these loci are not AGO-competent and have no known biological role. However, some of them have evolved to target host genes. Only the *TAS3*-derived tasiARF is conserved among the land plants [24–26]. We wondered if the virus also blocks the production of other tasiRNAs or if this phenomenon is specific to the *TAS3* pathway, therefore, we analyzed their expression profiles. Surprisingly, unlike the tasiARF, all the investigated functional tasiRNAs were virus-induced (Fig 12), except for the phosphate starvation-induced *TAS4*-siR81 [39,92] which is scarcely expressed under our circumstances.

### **Fig 12. Expression pattern of functional tasiRNAs other than the *TAS3*-derived tasiARF.**

The normalized expression values of the functional siRNAs are shown in the TVCV and TuMV series. We considered only the mature, functional sequence; other sequence variants or non-functional siRNAs from the same *TAS* locus were not included in the analysis. Error bars represent the standard error of the mean of four biological replicates. The actual data points are also shown.

These results suggest that the *TAS3* regulatory module is blocked at the *TAS3*-specific miR390 cleavage since all the other steps are carried out by shared components including the Bur virus-induced *SGS3* that might be responsible for the enhanced production of tasiRNAs from other loci. This blockage could happen either through ecotype-specific sequestering of miR390 or the AGO7 protein (or some unknown component of the tasiARF biogenesis

pathway) by the viral silencing suppressor. Alternatively, the AGO1-loading of tasiARF could be blocked which might lead to the destabilization of this siRNA specimen.

## **QTL mapping suggests that the observed leaf deformation is a result of multigenic interaction**

To support our findings, we performed a limited quantitative trait locus (QTL) mapping experiment using 43 selected recombinant inbred lines (RILs) between Bur and Col ecotypes that were established by INRA [77]. For the experiments, five plants per line were infected either with TVCV-ApH or mock and the leaf deformation was scored by inspection. The genotype data for the markers were obtained from INRA and along with the phenotype data, were used as input for the mapping software Windows QTL Cartographer [78]. A Single Marker Association test identified several markers on different chromosomes associated with the leaf deformation (Doc S1). The strongest association was observed at the bottom of chromosome 5, where the *ARF4* (AT5G60450) is located (among others). However, the presence of other possible QTLs suggests that the leaf deformations are not caused by the erroneous expression of one gene (*ARF4*), rather, it might be the result of the interaction of multiple genes.

## **Ecotype-specific genetic differences in the tasiARF biogenesis pathway components**

We hypothesized that the VSRs could interfere with tasiARF biogenesis either by sequestering miR390 or by blocking some functions of AGO7, or some unknown component of the biogenesis pathway in an ecotype-specific manner. One possibility was that the structure of the miR390 sRNA duplex was different in the two ecotypes because of a possibly different sequence of the star strand, resulting in a differential binding by the VSRs. The

sequence and structure of the miR390 sRNA duplex were shown to be important for selective binding by AGO7 [29,93] and possibly affect its binding to VSRs as well. We investigated this possibility by predicting the miR390 loci (*MIR390A* on chromosome 2, and *MIR390B* on chromosome 5) in the Col and Bur ecotypes with ShortStack using the appropriate sRNA and genome sequences of the individual ecotypes separately. The analysis revealed no difference in the sequence of miR390 precursors or miRNA duplex structures.

Alternatively, VSRs could bind and block one of the components of the tasiARF biogenesis pathway in an ecotype-specific manner. To investigate this possibility, we compared the amino acid sequences of the proteins involved in the tasiARF biogenesis. We downloaded the protein sequences of AGO1, AGO2, AGO7, ARF4, and FIL from 19 different *Arabidopsis thaliana* ecotypes (including Bur and Col) from the homepage of the 1001 genome project (<http://mtweb.cs.ucl.ac.uk/mus/www/19genomes/index.html>), and aligned the sequence variants with Clustal omega [94]. According to the alignment, there were five amino acid changes in the Bur version of AGO7 protein compared to the Col version (all in the N-terminal part), while there were two and one amino acid changes in ARF4 and AGO2 proteins, respectively. There was no difference in the sequence of AGO1 and FIL proteins in the Col and Bur ecotypes (Fig S2). The Bur version of AGO7 was more different than the versions in the rest of the ecotypes. The two amino acid changes in ARF4 were not conserved (E/K and T/I in Col and Bur, respectively), and again, the Bur variant was rare among the investigated ecotypes. The FIL and AGO1 sequences were remarkably conserved among the investigated 19 ecotypes suggesting a strong selection pressure on them. There was some degree of variability in the AGO2 sequence but the Bur was not the most divergent ecotype in this respect.

Taken together, the best candidate for the ecotype-specific blockage of tasiARF production is AGO7 (or an unknown component of the TAS3 module) but the exact mechanism of blockage remains to be elucidated.



# Discussion

Viral symptom development is regulated by both viral and host factors. On the virus side, viral silencing suppressors were found to be the major symptom determinants. However, much less data is available for such symptom determinants on the plant side. During an infection, the virus modulates the molecular environment of the host cell in a way that favors the replication and spread of the virus. Since the viral genome has a limited coding capacity, the virus utilizes host factors for its replication, cell-to-cell movement, virion assembly, and release. Genetic variations in these host factors could affect the severity of viral symptoms. Such variations occur naturally and determine the host range in which the virus can replicate efficiently. Host factors can also be manipulated artificially which may result in a restricted viral replication or enhanced immune response [95,96].

*Arabidopsis thaliana* is one of the most studied model organisms of plant biology. There are hundreds of ecotypes of this species that differ significantly in their geno- and phenotypes and therefore, are susceptible to viral infection differently. Indeed, in a genome-wide association study under field conditions, 317 *Arabidopsis* accessions were phenotyped for TuMV viral RNA accumulation and found significant differences among the susceptibility of the different ecotypes [97]. We also screened *Arabidopsis* ecotypes that display more severe symptoms upon virus infection (unpublished results) and we found that the Bur ecotype displays more severe leaf developmental defects than the Col ecotype when infected either with TVCV or TuMV. Since leaf development is regulated by sRNAs at many points [15,16,98,99], we expected that the mRNA and sRNA profiles of the Col and Bur ecotypes would be significantly different when infected with viruses. Namely, we expected that genes and sRNAs that were responsible for the observed severe symptoms only in the Bur ecotype would change significantly (in the Bur virus sample only compared to the Col mock, Col

virus, or Bur mock samples). To find such genes and sRNAs, we performed transcriptome and sRNAome analyses. Our analysis revealed that the levels of *FIL* (*YABBY1*, *AFO*) and *ARF4*, two transcription factors that are major abaxial polarity determinants in *Arabidopsis* leaf development [99], and the small RNA specimens tasiARF and vasiRNA-FIL that negatively regulates *ARF4* [30] and *FIL*, respectively, change significantly in the Bur ecotype when infected either with TVCV or TuMV, while such a change cannot be observed in the Col background. This led us to assume that the corruption of the *TAS3*-mediated regulatory module might be responsible for the observed leaf deformation in the virus-infected Bur plants. According to our analysis, miR390 is virus-induced in both ecotypes, especially upon TVCV infection but the tasiARF production is blocked only in the Bur virus sample. We observed that the production of other tasiRNAs is not blocked, rather, they are induced upon virus infection, especially in the Bur virus samples. This suggested that both viruses likely specifically block the *TAS3* pathway at the phase initiating AGO7-mediated cleavage. The Bur-specific depletion of tasiARF might lead to an ectopic expression of *ARF4* in the adaxial part of the leaf which could result in a misregulated auxin-dependent cell expansion in this domain, manifested as an upward curling leaf and wavy leaf margin.

*ARF4* is called a repressive, class B ARF [100–102] because it lacks the glutamine-rich transactivator domain that other activating ARFs have. Repressive ARFs might act by forming a dysfunctional heterodimer with other ARFs, and can form repressive chromatin in the promoter of the regulated gene [103]. In a previous study, *ARF8* was shown to be the major symptom determinant in transgenic plants expressing three different viral silencing suppressors [104]. Unlike *ARF4*, *ARF8* is an activating ARF and regulated post-transcriptionally by miR167 rather than tasiARF. In the *Arabidopsis* plants (Col background) overexpressing the TuMV silencing suppressor HCPro, *ARF4* was found to be upregulated that coincided with upward curling leaves, just like in our experiment. However, in the *arf4* mutant, the symptoms do not disappear, unlike in the *arf8* mutant. This might be because of

the different background (the *arf4* mutant is in Col) that suggests a multigenic interaction that allows the symptoms to develop only in the Bur background.

In a different work that utilized a dexamethasone-inducible *FIL:FIL-GR* construct, researchers showed that *FIL* positively regulates *ARF4* and *KANI* but not *ARF3* [85]. In our viral-infected samples, we observed the downregulation of *FIL* and upregulation of *ARF4* which seemingly contradicts the above-mentioned finding. However, the drop of the tasiARF level in our Bur virus sample in both the TVCV and TuMV series anti-correlates well with the *ARF4* level which might underline the importance of the post-transcriptional regulation of *ARF4*. It is also important to note that in the above-mentioned study, *FIL* was selectively induced in the abaxial part of the leaf where the *ARF4* is normally confined by the tasiARF gradient, while in our study, the *ARF4* expression was probably increased in the adaxial part due to the absence of inhibitory tasiARF. Also, the newly identified virus-induced vasiRNA-FIL might interfere with the delicate regulation of leaf polarity in an ecotype-specific manner.

Understanding the host-pathogen interaction network to diminish viral symptom development is one of the central goals of virological research. In this study, we have shown that two viruses having very different replication strategies, (TVCV and TuMV which are genuine representatives of *Tobamovirus* and *Potyvirus* families, respectively) cause very similar upward curling leaf and wavy leaf margin symptoms on the Bur but not on Col ecotype plants. The molecular basis for the ecotype difference therefore likely lies within the presence or absence of specific host factors that react in a very similar way to the infection of by the two pathogens. By studying the transcriptomes and sRNAomes of the two sets of infections we hypothesize that the basis of the symptom development in Bur plants may be the enhanced impairment of the *TAS3* pathway. Both viruses possess a VSR that binds sRNAs, constituting a common target of action, for example, miR390 duplex binding by the two VSRs would impair *TAS3*-derived tasiARF biogenesis. An ecotype-specific structural difference in the miR390 duplex could result in selective binding by the VSRs. However, we

could not find evidence for a different duplex structure of miR390a or miR390b in the two ecotypes. The RISC-loading efficiency of miR390 is very low and the overexpression of miR390 does not improve it, because the bottleneck is the availability of AGO7 [11]. It is possible that the miR390 loading into AGO7 is differentially affected by the presence of the virus either because the VSR directly manipulates the loading process or indirectly by relocating the AGO7 within the cell to a compartment that lacks miR390. Since some VSRs can bind ARGONAUTE proteins to block their antiviral action, we compared the amino acid sequences of the AGO proteins that could be involved in the production of tasiARF. There was no difference in the AGO1 sequence but AGO2 and especially AGO7 were more variable between the investigated ecotypes. We think that the best explanation is that the VSR binds and blocks some functions of the AGO7 protein selectively in the Bur ecotype but that would require a parallel evolution of the VSRs of different viruses to target the same AGO7 protein. Such parallel evolution of viral factors to target the same host factor is not unprecedented [105]. A study investigating the evolution of plant VSRs argues that plant VSRs are subject to episodic positive selection due to frequent jumps between different hosts rather than a one-to-one co-evolution with a single host [106]. This is consistent with our findings since both viruses we worked with have a broad host range (the turnip vein-clearing virus isolate we used, TVCV-ApH was isolated from *Alliaria petiolata*, a.k.a. garlic mustard [45]). It is unclear why the virus would favor blocking AGO7, although it was described that AGO7 has some antiviral function [107,108]. Another interesting possibility is that since the virus replication is often associated with an endomembrane system and the AGO7-mediated tasiRNA production is also endomembrane-associated [109], therefore, the viral replication could interfere with the AGO7-dependent tasiARF production. However, this interference should not be a passive process since it should happen only in the Bur ecotype.

Our QTL mapping results suggest that a multigenic interaction is behind the observed leaf deformations. Many of the genes that were selectively up- and downregulated in the Bur

virus samples could be related to leaf development: many of the upregulated genes regulate cell wall structure while many downregulated genes are involved in auxin biosynthesis (i.e. *TAR2*) or signaling (i.e. *IAA29*), and could be directly involved in ARF4-mediated auxin signaling. It is worth mentioning that there are numerous transcription factors in both sets (see Table S3) that could form an intricate regulatory network with or without ARF4 the net effect of which would lead to the observed phenotype. It is difficult to tell which players are the most upstream in the signaling chain that leads to the observed ecotype-specific leaf deformation, although we believe that the disrupted abaxial-adaxial gradient of tasiARF could be a major contributor. To reveal the exact nature of the selective blockage of tasiARF production in the virus-infected Bur ecotype, further research is needed in the future.

## **Supporting information**

**Fig S1. GO term enrichment analysis of the downregulated, potential vasiRNA target genes.**

**Fig S2. Amino acid alignment of ARF4, FIL, AGO1, AGO2, and AGO7 in 19 different ecotypes.**

**Table S1. Sequencing statistics of the RNA-seq and sRNA-seq data.**

**Table S2. Normalized expression values of all *Arabidopsis* genes; Wald test results and normalized expression values of differentially expressed genes.**

**Table S3. Commonly regulated genes between the TVCV and TuMV series and their GO term enrichment analysis.** Content of the two sets (Bur virus up- and downregulated genes in both virus series), and the full result of their GO term enrichment analysis.

**Table S4. Wald test results and normalized expression values of differentially expressed sRNAs; commonly regulated sRNAs between the TVCV and TuMV series.**

**Table S5. Target prediction of the Bur virus upregulated sRNAs among the Bur virus downregulated genes.** Result of the target prediction performed using the psRNATarget

server. Only the predicted Cleavage hits are shown. The sRNAs and the target genes are annotated.

**Data S1. Normalized expression values of all sRNAs in the samples.** Annotated expression table of sRNA sequences having a mean expression  $> 1$ . The table contains 102195 unique, 20–25-nt-long sequences ordered by mean expression value (highly expressed first).

**Doc S1. Result of the QTL mapping.** The result of a single marker analysis performed by WinQTL Cartographer.

## Acknowledgments

We would like to thank Tibor Csorba for his valuable suggestions to improve the manuscript.

## References

1. Hull R. Chapter 4 - Symptoms and Host Range. In: Hull R, editor. *Plant Virology* (Fifth Edition). Boston: Academic Press; 2014. pp. 145–198. Available: <http://www.sciencedirect.com/science/article/pii/B9780123848710000042>
2. Carrington JC, Kasschau KD, Mahajan SK, Schaad MC. Cell-to-Cell and Long-Distance Transport of Viruses in Plants. *Plant Cell Online*. 1996;8: 1669–1681. doi:10.1105/tpc.8.10.1669
3. Ding S-W, Voinnet O. Antiviral immunity directed by small RNAs. *Cell*. 2007;130: 413–426. doi:10.1016/j.cell.2007.07.039
4. 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. doi:10.1128/JVI.79.12.7812-7818.2005

5. 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. doi:10.1126/science.1128214
6. Voinnet O. Induction and suppression of RNA silencing: insights from viral infections. *Nat Rev Genet*. 2005;6: 206–220. doi:10.1038/nrg1555
7. Burgyán J, Havelda Z. Viral suppressors of RNA silencing. *Trends Plant Sci*. 2011;16: 265–272. doi:10.1016/j.tplants.2011.02.010
8. Csorba T, Kontra L, Burgyán J. Viral silencing suppressors: Tools forged to fine-tune host-pathogen coexistence. *Virology*. 2015;479–480: 85–103. doi:10.1016/j.virol.2015.02.028
9. Mi S, Cai T, Hu Y, Chen Y, Hodges E, Ni F, et al. Sorting of small RNAs into Arabidopsis Argonaute complexes is directed by the 5' terminal nucleotide. *Cell*. 2008;133: 116–127. doi:10.1016/j.cell.2008.02.034
10. Takeda A, Iwasaki S, Watanabe T, Utsumi M, Watanabe Y. The Mechanism Selecting the Guide Strand from Small RNA Duplexes is Different Among Argonaute Proteins. *Plant Cell Physiol*. 2008;49: 493–500. doi:10.1093/pcp/pcn043
11. Dalmadi Á, Gyula P, Bálint J, Szittyá G, Havelda Z. AGO-unbound cytosolic pool of mature miRNAs in plant cells reveals a novel regulatory step at AGO1 loading. *Nucleic Acids Res*. 2019;47: 9803–9817. doi:10.1093/nar/gkz690
12. Satish D, Mukherjee SK, Gupta D. The landscape of microRNAs in plant viral infections. *Plant Gene*. 2021; 100293. doi:10.1016/j.plgene.2021.100293
13. Zhang B, Li W, Zhang J, Wang L, Wu J. Roles of Small RNAs in Virus-Plant Interactions. *Viruses*. 2019;11: 827. doi:10.3390/v11090827
14. Bazzini AA, Hopp HE, Beachy RN, Asurmendi S. Infection and coaccumulation of tobacco mosaic virus proteins alter microRNA levels, correlating with symptom and plant development. *Proc Natl Acad Sci*. 2007;104: 12157–12162.

doi:10.1073/pnas.0705114104

15. Yang T, Wang Y, Teotia S, Zhang Z, Tang G. The Making of Leaves: How Small RNA Networks Modulate Leaf Development. *Front Plant Sci.* 2018;9.  
doi:10.3389/fpls.2018.00824
16. Pulido A, Laufs P. Co-ordination of developmental processes by small RNAs during leaf development. *J Exp Bot.* 2010;61: 1277–1291. doi:10.1093/jxb/erp397
17. Marín-González E, Suárez-López P. “And yet it moves”: Cell-to-cell and long-distance signaling by plant microRNAs. *Plant Sci.* 2012;196: 18–30.  
doi:10.1016/j.plantsci.2012.07.009
18. Chitwood DH, Nogueira FTS, Howell MD, Montgomery TA, Carrington JC, Timmermans MCP. Pattern formation via small RNA mobility. *Genes Dev.* 2009;23: 549–554. doi:10.1101/gad.1770009
19. Xia R, Xu J, Meyers BC. The Emergence, Evolution, and Diversification of the miR390-TAS3-ARF Pathway in Land Plants. *Plant Cell.* 2017;29: 1232–1247.  
doi:10.1105/tpc.17.00185
20. Fahlgren N, Montgomery TA, Howell MD, Allen E, Dvorak SK, Alexander AL, et al. Regulation of AUXIN RESPONSE FACTOR3 by TAS3 ta-siRNA Affects Developmental Timing and Patterning in Arabidopsis. *Curr Biol.* 2006;16: 939–944.  
doi:10.1016/j.cub.2006.03.065
21. Garcia D, Collier SA, Byrne ME, Martienssen RA. Specification of Leaf Polarity in Arabidopsis via the trans-Acting siRNA Pathway. *Curr Biol.* 2006;16: 933–938.  
doi:10.1016/j.cub.2006.03.064
22. Marin E, Jouannet V, Herz A, Lokerse AS, Weijers D, Vaucheret H, et al. miR390, Arabidopsis TAS3 tasiRNAs, and Their AUXIN RESPONSE FACTOR Targets Define an Autoregulatory Network Quantitatively Regulating Lateral Root Growth. *Plant Cell.* 2010;22: 1104–1117. doi:10.1105/tpc.109.072553



23. Hunter C, Willmann MR, Wu G, Yoshikawa M, Gutiérrez-Nava M de la L, Poethig SR. Trans-acting siRNA-mediated repression of ETTIN and ARF4 regulates heteroblasty in Arabidopsis. *Development*. 2006;133: 2973–2981.  
doi:10.1242/dev.02491
24. Axtell MJ, Jan C, Rajagopalan R, Bartel DP. A Two-Hit Trigger for siRNA Biogenesis in Plants. *Cell*. 2006;127: 565–577. doi:10.1016/j.cell.2006.09.032
25. Nagasaki H, Itoh J, Hayashi K, Hibara K, Satoh-Nagasawa N, Nosaka M, et al. The small interfering RNA production pathway is required for shoot meristem initiation in rice. *Proc Natl Acad Sci*. 2007;104: 14867–14871. doi:10.1073/pnas.0704339104
26. Douglas RN, Wiley D, Sarkar A, Springer N, Timmermans MCP, Scanlon MJ. ragged seedling2 Encodes an ARGONAUTE7-Like Protein Required for Mediolateral Expansion, but Not Dorsiventrality, of Maize Leaves. *Plant Cell*. 2010;22: 1441–1451. doi:10.1105/tpc.109.071613
27. Adenot X, Elmayan T, Laressergues D, Boutet S, Bouché N, Gascioli V, et al. DRB4-Dependent TAS3 trans-Acting siRNAs Control Leaf Morphology through AGO7. *Curr Biol*. 2006;16: 927–932. doi:10.1016/j.cub.2006.03.035
28. Pekker I, Alvarez JP, Eshed Y. Auxin Response Factors Mediate Arabidopsis Organ Asymmetry via Modulation of KANADI Activity. *Plant Cell*. 2005;17: 2899–2910.  
doi:10.1105/tpc.105.034876
29. Montgomery TA, Howell MD, Cuperus JT, Li D, Hansen JE, Alexander AL, et al. Specificity of ARGONAUTE7-miR390 interaction and dual functionality in TAS3 trans-acting siRNA formation. *Cell*. 2008;133: 128–141.  
doi:10.1016/j.cell.2008.02.033
30. Allen E, Xie Z, Gustafson AM, Carrington JC. microRNA-Directed Phasing during Trans-Acting siRNA Biogenesis in Plants. *Cell*. 2005;121: 207–221.  
doi:10.1016/j.cell.2005.04.004

31. Schwab R, Maizel A, Ruiz-Ferrer V, Garcia D, Bayer M, Crespi M, et al. Endogenous TasiRNAs Mediate Non-Cell Autonomous Effects on Gene Regulation in *Arabidopsis thaliana*. *PLOS ONE*. 2009;4: e5980. doi:10.1371/journal.pone.0005980
32. Meyers BC, Kaushik S, Nandety RS. Evolving disease resistance genes. *Curr Opin Plant Biol*. 2005;8: 129–134. doi:10.1016/j.pbi.2005.01.002
33. Leonetti P, Stuttmann J, Pantaleo V. Regulation of plant antiviral defense genes via host RNA-silencing mechanisms. *Virol J*. 2021;18: 194. doi:10.1186/s12985-021-01664-3
34. Shivaprasad PV, Chen H-M, Patel K, Bond DM, Santos BACM, Baulcombe DC. A microRNA superfamily regulates nucleotide binding site-leucine-rich repeats and other mRNAs. *Plant Cell*. 2012;24: 859–874. doi:10.1105/tpc.111.095380
35. Zhang Y, Xia R, Kuang H, Meyers BC. The Diversification of Plant NBS-LRR Defense Genes Directs the Evolution of MicroRNAs That Target Them. *Mol Biol Evol*. 2016;33: 2692–2705. doi:10.1093/molbev/msw154
36. Sós- Hegedűs A, Domonkos Á, Tóth T, Gyula P, Kaló P, Szittyá G. Suppression of NB-LRR genes by miRNAs promotes nitrogen-fixing nodule development in *Medicago truncatula*. *Plant Cell Environ*. 2020;43: 1117–1129. doi:10.1111/pce.13698
37. Zhai J, Jeong D-H, De Paoli E, Park S, Rosen BD, Li Y, et al. 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. doi:10.1101/gad.177527.111
38. Felippes FF, Weigel D. Triggering the formation of tasiRNAs in *Arabidopsis thaliana*: the role of microRNA miR173. *EMBO Rep*. 2009;10: 264–270. doi:10.1038/embor.2008.247
39. Rajagopalan R, Vaucheret H, Trejo J, Bartel DP. A diverse and evolutionarily fluid

- set of microRNAs in *Arabidopsis thaliana*. *Genes Dev.* 2006;20: 3407–3425.  
doi:10.1101/gad.1476406
40. Li S, Liu J, Liu Z, Li X, Wu F, He Y. HEAT-INDUCED TAS1 TARGET1 Mediates Thermotolerance via HEAT STRESS TRANSCRIPTION FACTOR A1a-Directed Pathways in *Arabidopsis*. *Plant Cell.* 2014;26: 1764–1780.  
doi:10.1105/tpc.114.124883
41. Si-Ammour A, Windels D, Arn-Boulidoires E, Kutter C, Ailhas J, Meins F, et al. miR393 and Secondary siRNAs Regulate Expression of the TIR1/AFB2 Auxin Receptor Clade and Auxin-Related Development of *Arabidopsis* Leaves. *Plant Physiol.* 2011;157: 683–691. doi:10.1104/pp.111.180083
42. Cao M, Du P, Wang X, Yu Y-Q, Qiu Y-H, Li W, et al. Virus infection triggers widespread silencing of host genes by a distinct class of endogenous siRNAs in *Arabidopsis*. *Proc Natl Acad Sci.* 2014;111: 14613–14618.  
doi:10.1073/pnas.1407131111
43. Leonetti P, Ghasemzadeh A, Consiglio A, Gursinsky T, Behrens S, Pantaleo V. Endogenous activated small interfering RNAs in virus- infected Brassicaceae crops show a common host gene- silencing pattern affecting photosynthesis and stress response. *New Phytol.* 2021;229: 1650–1664. doi:10.1111/nph.16932
44. Pitzalis N, Amari K, Graindorge S, Pflieger D, Donaire L, Wassenegger M, et al. Turnip mosaic virus in oilseed rape activates networks of sRNA-mediated interactions between viral and host genomes. *Commun Biol.* 2020;3: 1–16.  
doi:10.1038/s42003-020-01425-y
45. Tóth T, Gyula P, Salamon P, Kis S, Sós-Hegedűs A, Szittyá 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.  
doi:10.1371/journal.pone.0224398

46. Lartey RT. Occurrence of a vein-clearing tobamovirus in turnip. *Plant Dis.* 1993;77: 21. doi:10.1094/PD-77-0021
47. Melcher U. Turnip vein-clearing virus, from pathogen to host expression profile. *Mol Plant Pathol.* 2003;4: 133–140. doi:10.1046/j.1364-3703.2003.00159.x
48. Csorba T, Bovi A, Dalmay T, Burgyán J. The p122 subunit of Tobacco mosaic virus replicase is a potent silencing suppressor and compromises both small interfering RNA- and microRNA-mediated pathways. *J Virol.* 2007;81: 11768–11780. doi:10.1128/JVI.01230-07
49. Ding XS, Liu J, Cheng N-H, Folimonov A, Hou Y-M, Bao Y, et al. The Tobacco mosaic virus 126-kDa protein associated with virus replication and movement suppresses RNA silencing. *Mol Plant-Microbe Interact MPMI.* 2004;17: 583–592. doi:10.1094/MPMI.2004.17.6.583
50. Vogler H, Akbergenov R, Shivaprasad PV, Dang V, Fasler M, Kwon M-O, et al. Modification of small RNAs associated with suppression of RNA silencing by tobamovirus replicase protein. *J Virol.* 2007;81: 10379–10388. doi:10.1128/JVI.00727-07
51. Shattuck VI. *The Biology, Epidemiology, and Control of Turnip Mosaic Virus.* Horticultural Reviews. John Wiley & Sons, Ltd; 1992. pp. 199–238. doi:10.1002/9780470650523.ch4
52. Riechmann JL, Laín S, García JA 1992. Highlights and prospects of potyvirus molecular biology. *J Gen Virol.* 1992;73: 1–16. doi:10.1099/0022-1317-73-1-1
53. Urcuqui-Inchima S, Haenni A-L, Bernardi F. Potyvirus proteins: a wealth of functions. *Virus Res.* 2001;74: 157–175. doi:10.1016/S0168-1702(01)00220-9
54. Chung BY-W, Miller WA, Atkins JF, Firth AE. An overlapping essential gene in the Potyviridae. *Proc Natl Acad Sci.* 2008;105: 5897–5902. doi:10.1073/pnas.0800468105

55. Wei T, Zhang C, Hong J, Xiong R, Kasschau KD, Zhou X, et al. Formation of Complexes at Plasmodesmata for Potyvirus Intercellular Movement Is Mediated by the Viral Protein P3N-PIPO. *PLOS Pathog.* 2010;6: e1000962.  
doi:10.1371/journal.ppat.1000962
56. Anandalakshmi R, Pruss GJ, Ge X, Marathe R, Mallory AC, Smith TH, et al. A viral suppressor of gene silencing in plants. *Proc Natl Acad Sci.* 1998;95: 13079–13084.  
doi:10.1073/pnas.95.22.13079
57. Garcia-Ruiz H, Carbonell A, Hoyer JS, Fahlgren N, Gilbert KB, Takeda A, et al. Roles and Programming of Arabidopsis ARGONAUTE Proteins during Turnip Mosaic Virus Infection. *PLoS Pathog.* 2015;11: e1004755.  
doi:10.1371/journal.ppat.1004755
58. Lakatos L, Csorba T, Pantaleo V, Chapman EJ, Carrington JC, Liu Y-P, et al. Small RNA binding is a common strategy to suppress RNA silencing by several viral suppressors. *EMBO J.* 2006;25: 2768–2780. doi:10.1038/sj.emboj.7601164
59. Szittyá G, Salamon P, Burgyán J. The complete nucleotide sequence and synthesis of infectious RNA of genomic and defective interfering RNAs of TBSV-P. *Virus Res.* 2000;69: 131–136. doi:10.1016/S0168-1702(00)00178-7
60. Andrews S. FastQC: a quality control tool for high throughput sequence data. In: Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data [Internet]. 2010. Available:  
<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
61. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal.* 2011;17: 10–12. doi:10.14806/ej.17.1.200
62. Zhang R, Calixto CPG, Marquez Y, Venhuizen P, Tzioutziou NA, Guo W, et al. A high quality Arabidopsis transcriptome for accurate transcript-level analysis of alternative splicing. *Nucleic Acids Res.* 2017;45: 5061–5073.

- doi:10.1093/nar/gkx267
63. Pimentel H, Bray NL, Puente S, Melsted P, Pachter L. Differential analysis of RNA-seq incorporating quantification uncertainty. *Nat Methods*. 2017;14: 687–690.  
doi:10.1038/nmeth.4324
  64. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Stat Soc Ser B Methodol*. 1995;57: 289–300.
  65. Wickham H. *Ggplot2: elegant graphics for data analysis*. New York: Springer; 2009.
  66. Tang Y, Horikoshi M, Li W. ggfortify: unified interface to visualize statistical results of popular R packages. *R J*. 2016;8: 474–485.
  67. Conway JR, Lex A, Gehlenborg N. UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics*. 2017;33: 2938–2940.  
doi:10.1093/bioinformatics/btx364
  68. Tian F, Yang D-C, Meng Y-Q, Jin J, Gao G. PlantRegMap: charting functional regulatory maps in plants. *Nucleic Acids Res*. 2020;48: D1104–D1113.  
doi:10.1093/nar/gkz1020
  69. Axtell MJ. ShortStack: Comprehensive annotation and quantification of small RNA genes. *RNA*. 2013 [cited 29 Mar 2016]. doi:10.1261/rna.035279.112
  70. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25: 2078–2079.  
doi:10.1093/bioinformatics/btp352
  71. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15: 550. doi:10.1186/s13059-014-0550-8
  72. Dai X, Zhao PX. psRNATarget: a plant small RNA target analysis server. *Nucleic Acids Res*. 2011;39: W155–W159. doi:10.1093/nar/gkr319

73. Rice P, Longden I, Bleasby A. EMBOSS: The European Molecular Biology Open Software Suite. *Trends Genet.* 2000;16: 276–277. doi:10.1016/S0168-9525(00)02024-2
74. Marshall OJ. PerlPrimer: cross-platform, graphical primer design for standard, bisulphite and real-time PCR. *Bioinformatics.* 2004;20: 2471–2472. doi:10.1093/bioinformatics/bth254
75. Czechowski T, Bari RP, Stitt M, Scheible W-R, Udvardi MK. Real-time RT-PCR profiling of over 1400 Arabidopsis transcription factors: unprecedented sensitivity reveals novel root- and shoot-specific genes. *Plant J.* 2004;38: 366–379. doi:10.1111/j.1365-313X.2004.02051.x
76. Lilly ST, Drummond RSM, Pearson MN, MacDiarmid RM. Identification and Validation of Reference Genes for Normalization of Transcripts from Virus-Infected Arabidopsis thaliana. *Mol Plant Microbe Interact.* 2011;24: 294–304. doi:10.1094/MPMI-10-10-0236
77. Simon M, Loudet O, Durand S, Bérard A, Brunel D, Sennesal F-X, et al. Quantitative Trait Loci Mapping in Five New Large Recombinant Inbred Line Populations of Arabidopsis thaliana Genotyped With Consensus Single-Nucleotide Polymorphism Markers. *Genetics.* 2008;178: 2253–2264. doi:10.1534/genetics.107.083899
78. Wang S, Basten CJ, Zeng Z-B. Windows QTL Cartographer 2.5. Department of Statistics, North Carolina State University, Raleigh, NC.; 2012. Available: <http://statgen.ncsu.edu/qtlcart/WQTLCart.htm>
79. Whitham SA, Quan S, Chang H-S, Cooper B, Estes B, Zhu T, et al. Diverse RNA viruses elicit the expression of common sets of genes in susceptible Arabidopsis thaliana plants. *Plant J.* 2003;33: 271–283. doi:10.1046/j.1365-313X.2003.01625.x
80. Whitham SA, Yang C, Goodin MM. Global Impact: Elucidating Plant Responses to

- Viral Infection. *Mol Plant Microbe Interact.* 2006;19: 1207–1215.  
doi:10.1094/MPMI-19-1207
81. Bengyella L, Waikhom SD, Allie F, Rey C. Virus tolerance and recovery from viral induced-symptoms in plants are associated with transcriptome reprogramming. *Plant Mol Biol.* 2015;89: 243–252. doi:10.1007/s11103-015-0362-6
  82. Postnikova OA, Nemchinov LG. Comparative analysis of microarray data in *Arabidopsis* transcriptome during compatible interactions with plant viruses. *Virol J.* 2012;9: 101. doi:10.1186/1743-422X-9-101
  83. Medzihradzky 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. *Mol Plant Pathol.* 2019;20: 1748–1758. doi:10.1111/mpp.12875
  84. Siegfried KR, Eshed Y, Baum SF, Otsuga D, Drews GN, Bowman JL. Members of the YABBY gene family specify abaxial cell fate in *Arabidopsis*. *Dev Camb Engl.* 1999;126: 4117–4128.
  85. Bonaccorso O, Lee JE, Puah L, Scutt CP, Golz JF. FILAMENTOUS FLOWER controls lateral organ development by acting as both an activator and a repressor. *BMC Plant Biol.* 2012;12: 176. doi:10.1186/1471-2229-12-176
  86. Hu Q, Hollunder J, Niehl A, Kørner CJ, Gereige D, Windels D, et al. Specific Impact of Tobamovirus Infection on the *Arabidopsis* Small RNA Profile. *PLoS ONE.* 2011;6: e19549. doi:10.1371/journal.pone.0019549
  87. Tatineni S, Riethoven J-JM, Graybosch RA, French R, Mitra A. Dynamics of Small RNA Profiles of Virus and Host Origin in Wheat Cultivars Synergistically Infected by Wheat Streak Mosaic Virus and Triticum Mosaic Virus: Virus Infection Caused a Drastic Shift in the Endogenous Small RNA Profile. *PLOS ONE.* 2014;9: e111577. doi:10.1371/journal.pone.0111577



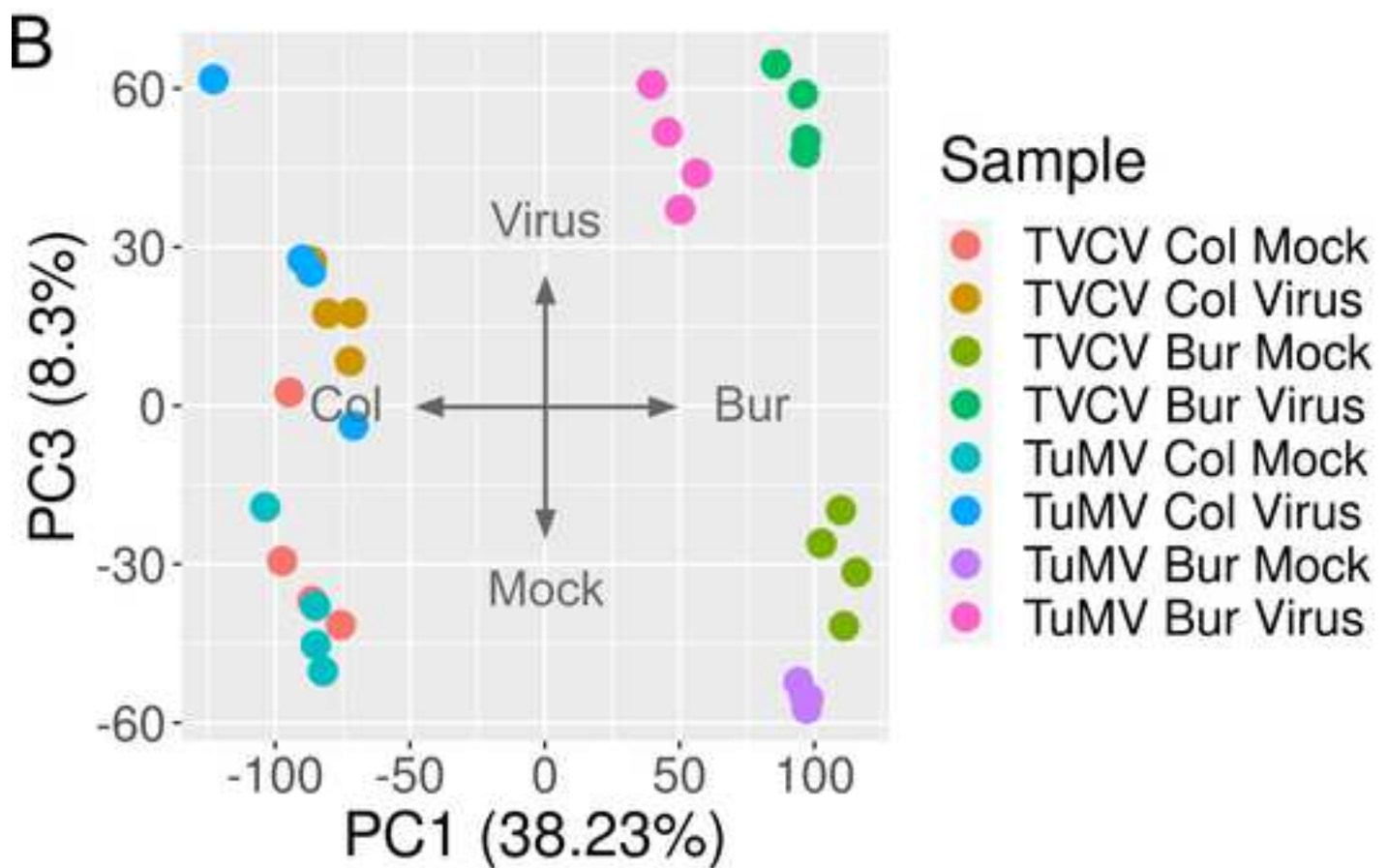
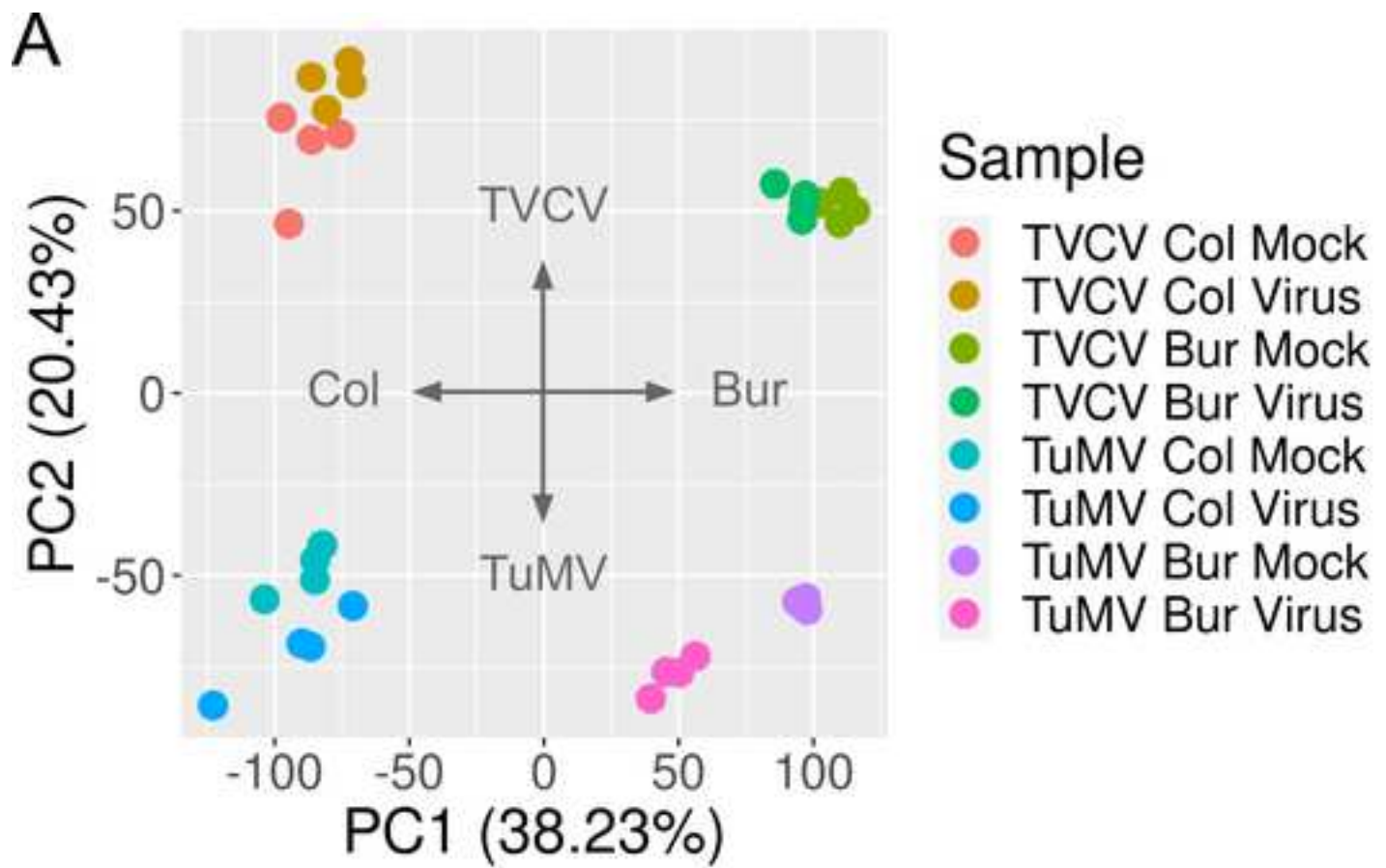
88. Blevins T, Rajeswaran R, Shivaprasad PV, Beknazariants D, Si-Ammour A, Park H-S, et al. Four plant Dicers mediate viral small RNA biogenesis and DNA virus induced silencing. *Nucleic Acids Res.* 2006;34: 6233–6246. doi:10.1093/nar/gkl886
89. Bouché N, Laressergues D, Gascioli V, Vaucheret H. An antagonistic function for *Arabidopsis* DCL2 in development and a new function for DCL4 in generating viral siRNAs. *EMBO J.* 2006;25: 3347–3356. doi:10.1038/sj.emboj.7601217
90. Wang X-B, Wu Q, Ito T, Cillo F, Li W-X, Chen X, et al. RNAi-mediated viral immunity requires amplification of virus-derived siRNAs in *Arabidopsis thaliana*. *Proc Natl Acad Sci.* 2010;107: 484–489. doi:10.1073/pnas.0904086107
91. Wang L, Gu X, Xu D, Wang W, Wang H, Zeng M, et al. miR396-targeted AtGRF transcription factors are required for coordination of cell division and differentiation during leaf development in *Arabidopsis*. *J Exp Bot.* 2011;62: 761–773. doi:10.1093/jxb/erq307
92. Howell MD, Fahlgren N, Chapman EJ, Cumbie JS, Sullivan CM, Givan SA, et al. Genome-Wide Analysis of the RNA-DEPENDENT RNA POLYMERASE6/DICER-LIKE4 Pathway in *Arabidopsis* Reveals Dependency on miRNA- and tasiRNA-Directed Targeting. *Plant Cell.* 2007;19: 926–942. doi:10.1105/tpc.107.050062
93. Endo Y, Iwakawa H, Tomari Y. *Arabidopsis* ARGONAUTE7 selects miR390 through multiple checkpoints during RISC assembly. *EMBO Rep.* 2013;14: 652–658. doi:10.1038/embor.2013.73
94. Madeira F, Pearce M, Tivey ARN, Basutkar P, Lee J, Edbali O, et al. Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Res.* 2022; gkac240. doi:10.1093/nar/gkac240
95. Zaidi SS-A, Tashkandi M, Mansoor S, Mahfouz MM. Engineering Plant Immunity: Using CRISPR/Cas9 to Generate Virus Resistance. *Front Plant Sci.* 2016;7: 1673. doi:10.3389/fpls.2016.01673

96. Borrelli VMG, Brambilla V, Rogowsky P, Marocco A, Lanubile A. The Enhancement of Plant Disease Resistance Using CRISPR/Cas9 Technology. *Front Plant Sci.* 2018;9: 1245. doi:10.3389/fpls.2018.01245
97. Rubio B, Cosson P, Caballero M, Revers F, Bergelson J, Roux F, et al. Genome-wide association study reveals new loci involved in *Arabidopsis thaliana* and Turnip mosaic virus (TuMV) interactions in the field. *New Phytol.* 2019;221: 2026–2038. doi:10.1111/nph.15507
98. D’Ario M, Griffiths-Jones S, Kim M. Small RNAs: Big Impact on Plant Development. *Trends Plant Sci.* 2017;22: 1056–1068. doi:10.1016/j.tplants.2017.09.009
99. Yamaguchi T, Nukazuka A, Tsukaya H. Leaf adaxial–abaxial polarity specification and lamina outgrowth: evolution and development. *Plant Cell Physiol.* 2012;53: 1180–1194. doi:10.1093/pcp/pcs074
100. Ulmasov T, Hagen G, Guilfoyle TJ. Activation and repression of transcription by auxin-response factors. *Proc Natl Acad Sci.* 1999;96: 5844–5849. doi:10.1073/pnas.96.10.5844
101. Ulmasov T, Hagen G, Guilfoyle TJ. Dimerization and DNA binding of auxin response factors. *Plant J.* 1999;19: 309–319. doi:https://doi.org/10.1046/j.1365-313X.1999.00538.x
102. Weijers D, Wagner D. Transcriptional Responses to the Auxin Hormone. *Annu Rev Plant Biol.* 2016;67: 539–574. doi:10.1146/annurev-arplant-043015-112122
103. Chung Y, Zhu Y, Wu M-F, Simonini S, Kuhn A, Armenta-Medina A, et al. Auxin Response Factors promote organogenesis by chromatin-mediated repression of the pluripotency gene SHOOTMERISTEMLESS. *Nat Commun.* 2019;10: 886. doi:10.1038/s41467-019-08861-3
104. Jay F, Wang Y, Yu A, Taconnat L, Pelletier S, Colot V, et al. Misregulation of

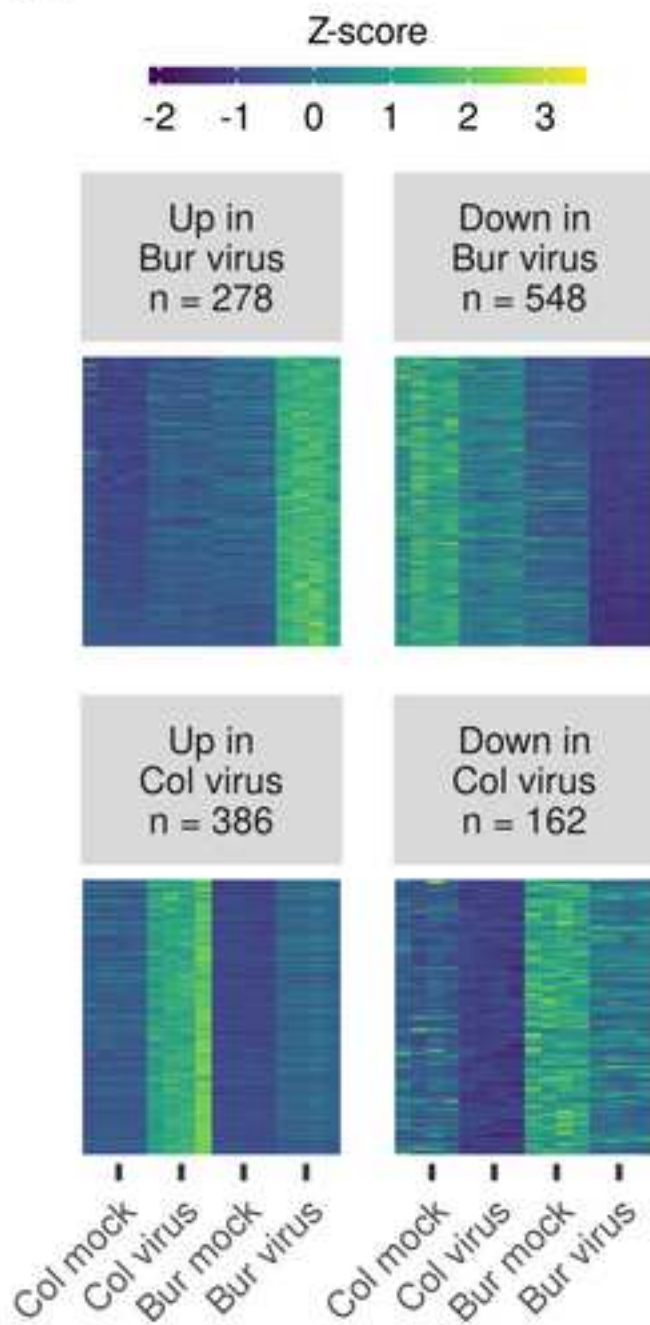
- AUXIN RESPONSE FACTOR 8 Underlies the Developmental Abnormalities Caused by Three Distinct Viral Silencing Suppressors in Arabidopsis. *PLOS Pathog.* 2011;7: e1002035. doi:10.1371/journal.ppat.1002035
105. Zhang H, Li L, He Y, Qin Q, Chen C, Wei Z, et al. Distinct modes of manipulation of rice auxin response factor OsARF17 by different plant RNA viruses for infection. *Proc Natl Acad Sci.* 2020;117: 9112–9121. doi:10.1073/pnas.1918254117
106. Murray GGR, Kosakovskiy SL, Obbard DJ. Suppressors of RNAi from plant viruses are subject to episodic positive selection. *Proc R Soc B Biol Sci.* 2013;280: 20130965. doi:10.1098/rspb.2013.0965
107. Qu F, Ye X, Morris TJ. Arabidopsis DRB4, AGO1, AGO7, and RDR6 participate in a DCL4-initiated antiviral RNA silencing pathway negatively regulated by DCL1. *Proc Natl Acad Sci.* 2008;105: 14732–14737. doi:10.1073/pnas.0805760105
108. Carbonell A, Carrington JC. Antiviral roles of plant ARGONAUTES. *Curr Opin Plant Biol.* 2015;27: 111–117. doi:10.1016/j.pbi.2015.06.013
109. Jouannet V, Moreno AB, Elmayan T, Vaucheret H, Crespi MD, Maizel A. Cytoplasmic Arabidopsis AGO7 accumulates in membrane-associated siRNA bodies and is required for ta-siRNA biogenesis. *EMBO J.* 2012;31: 1704–1713. doi:10.1038/emboj.2012.20







A



B

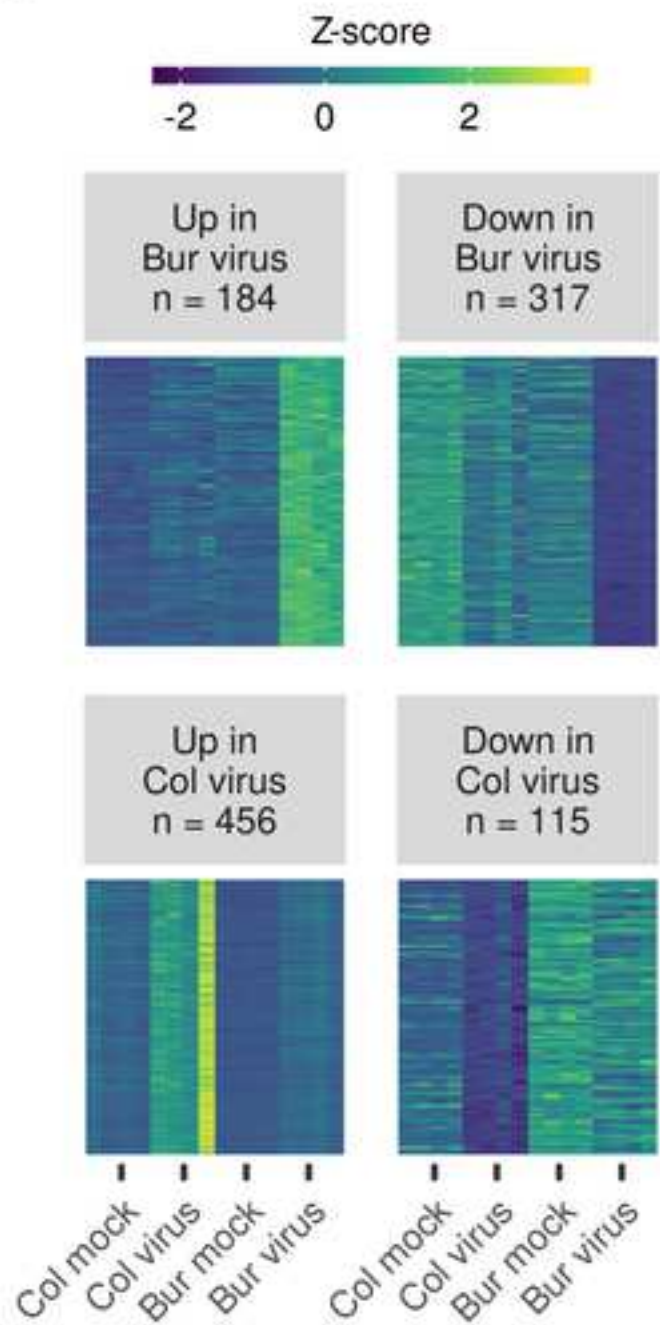


Figure 4

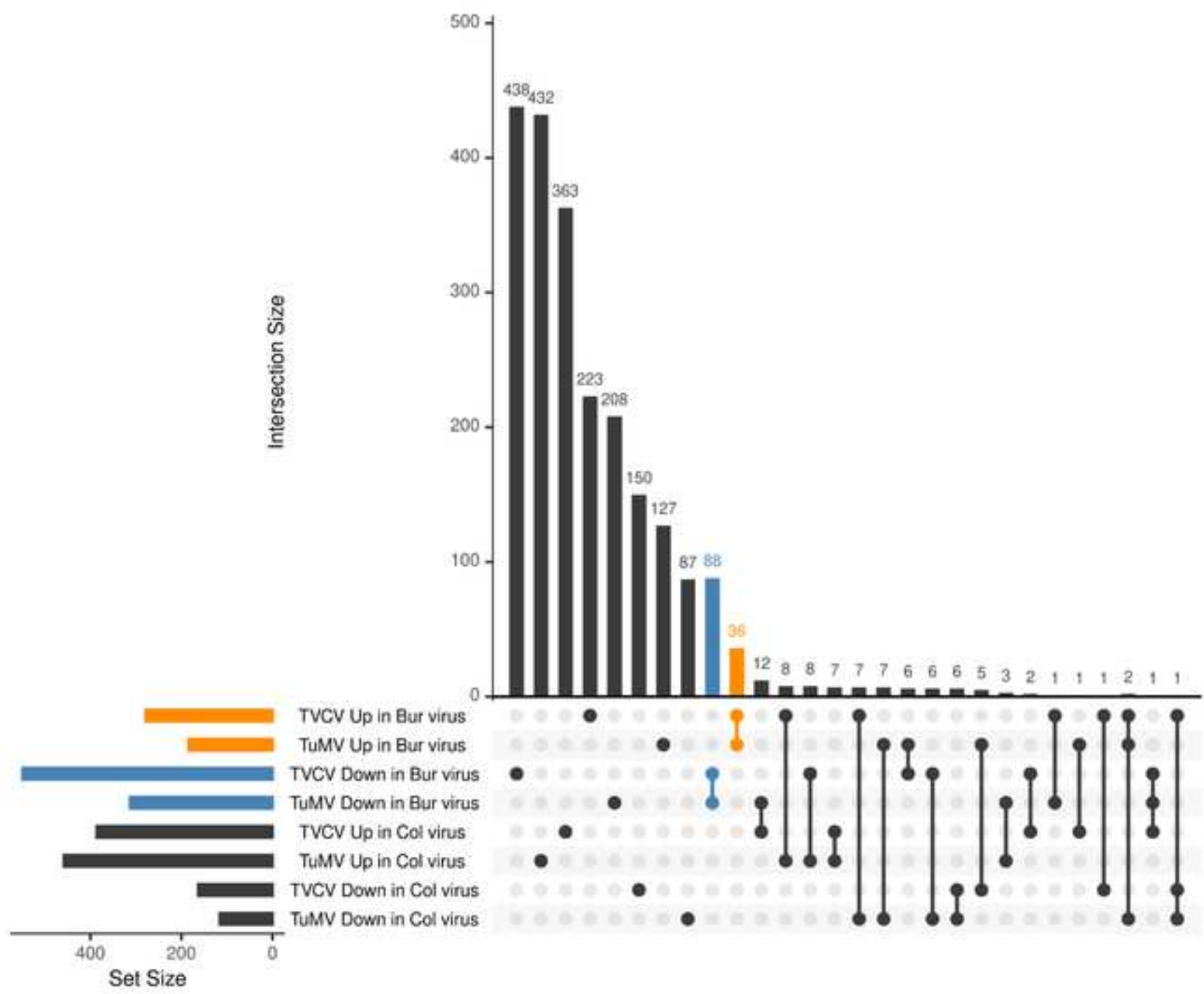
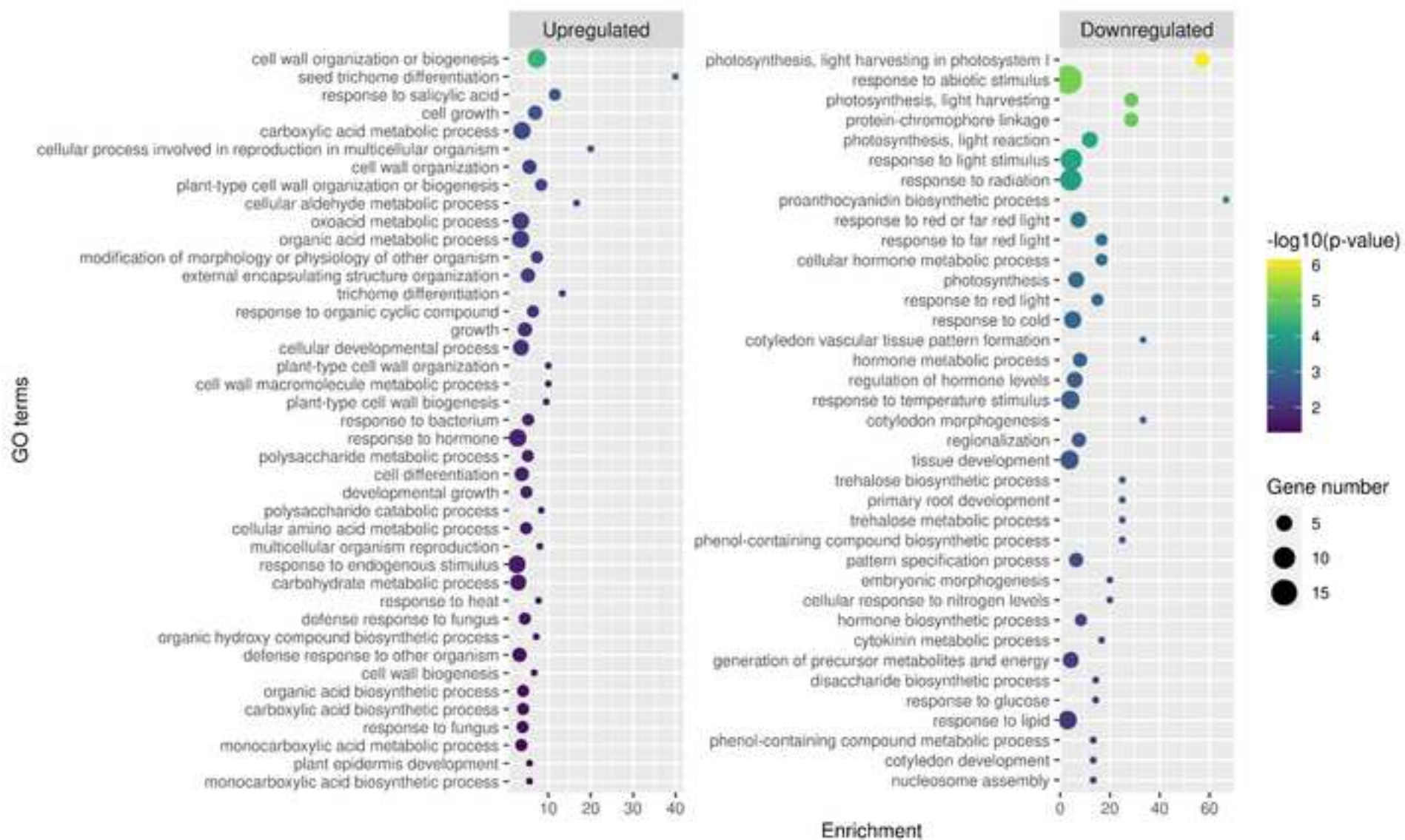
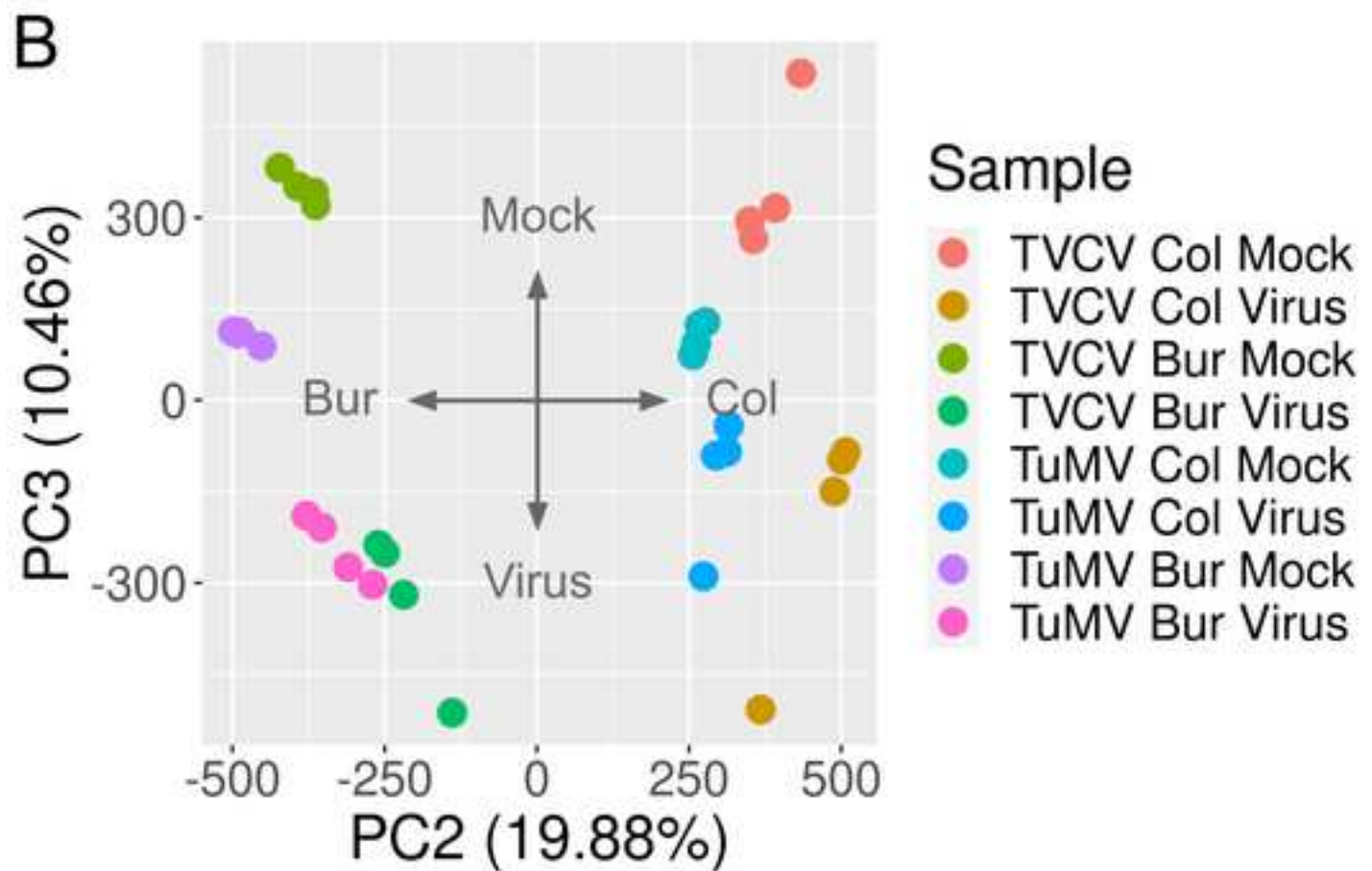
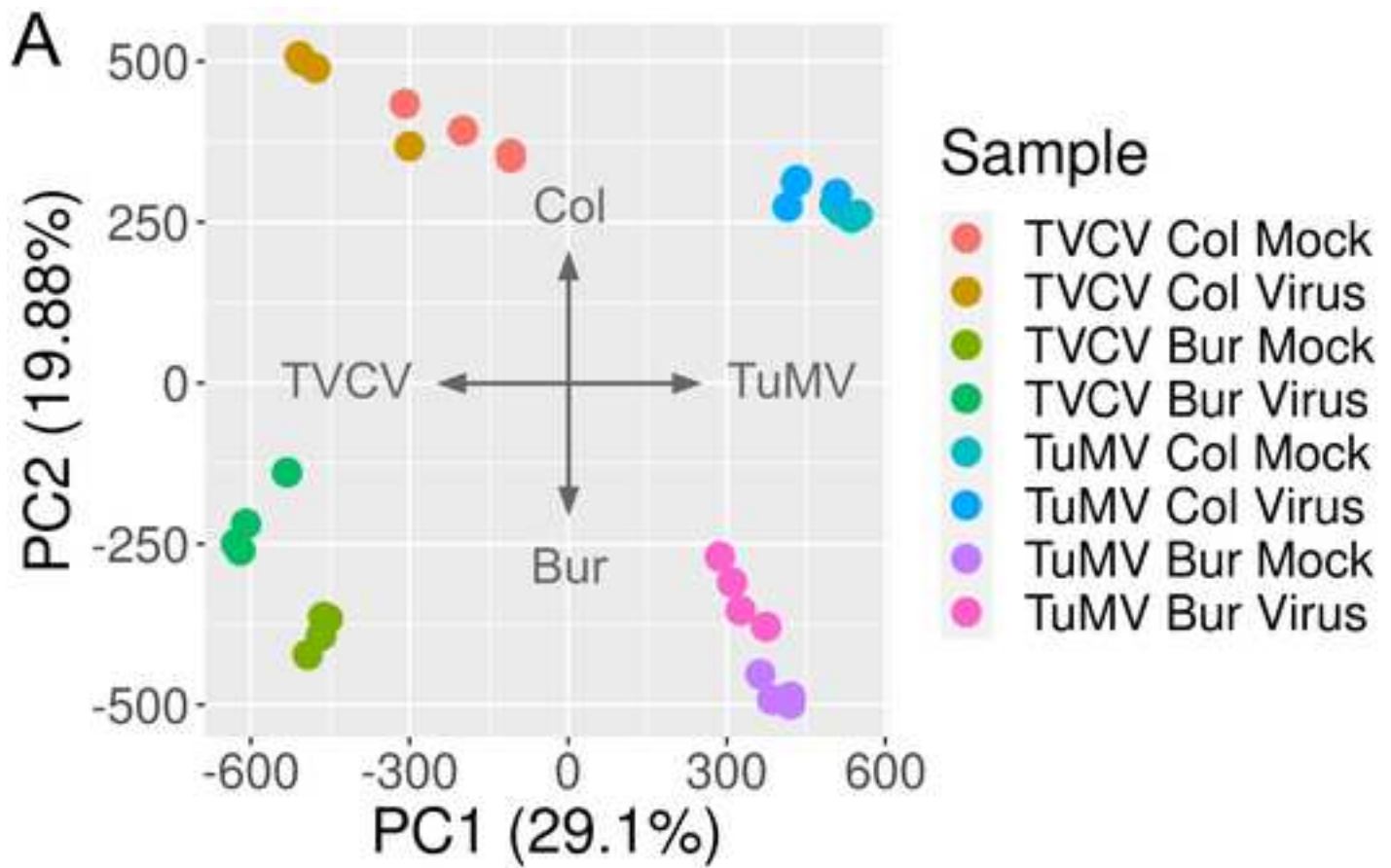


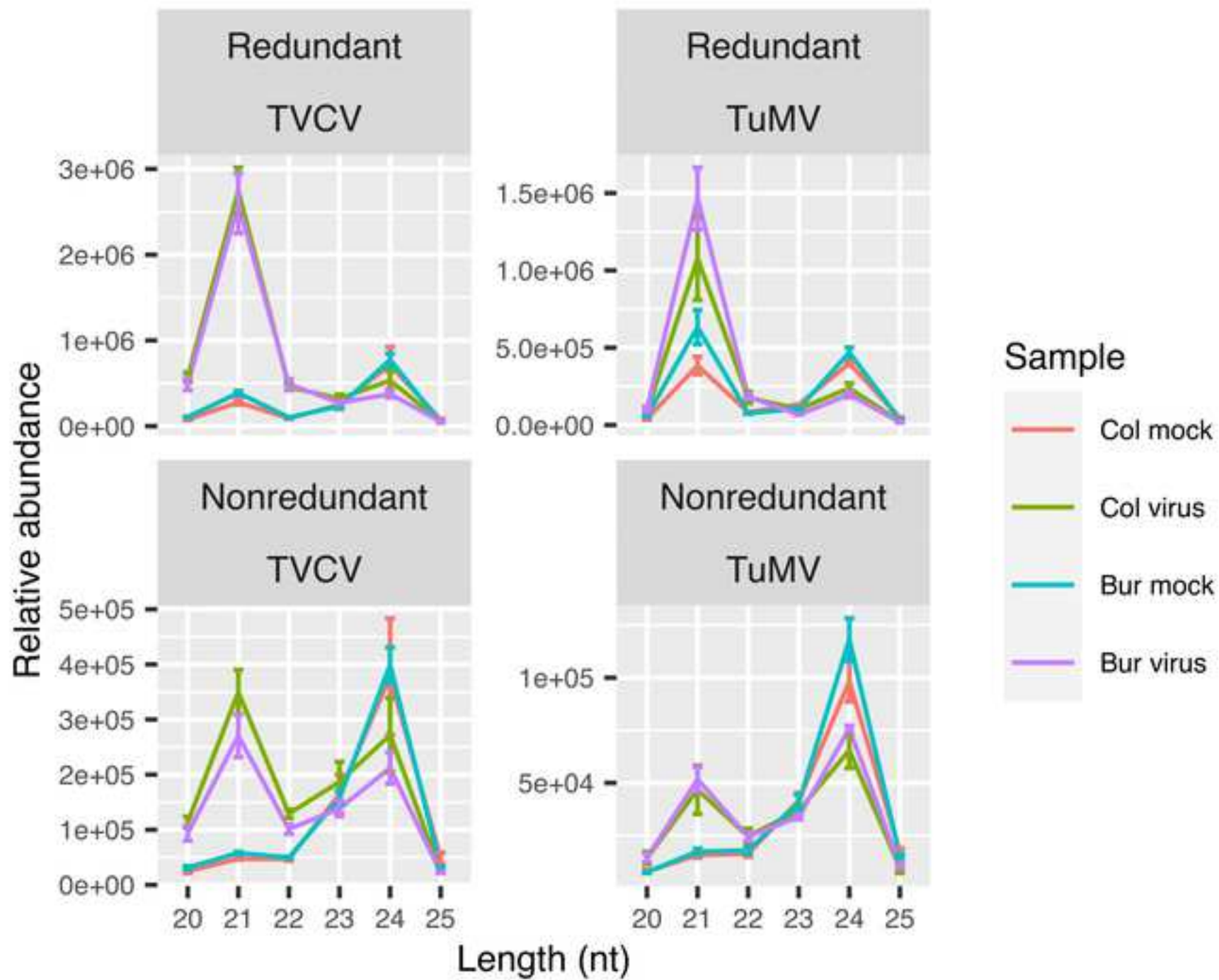
Figure 5

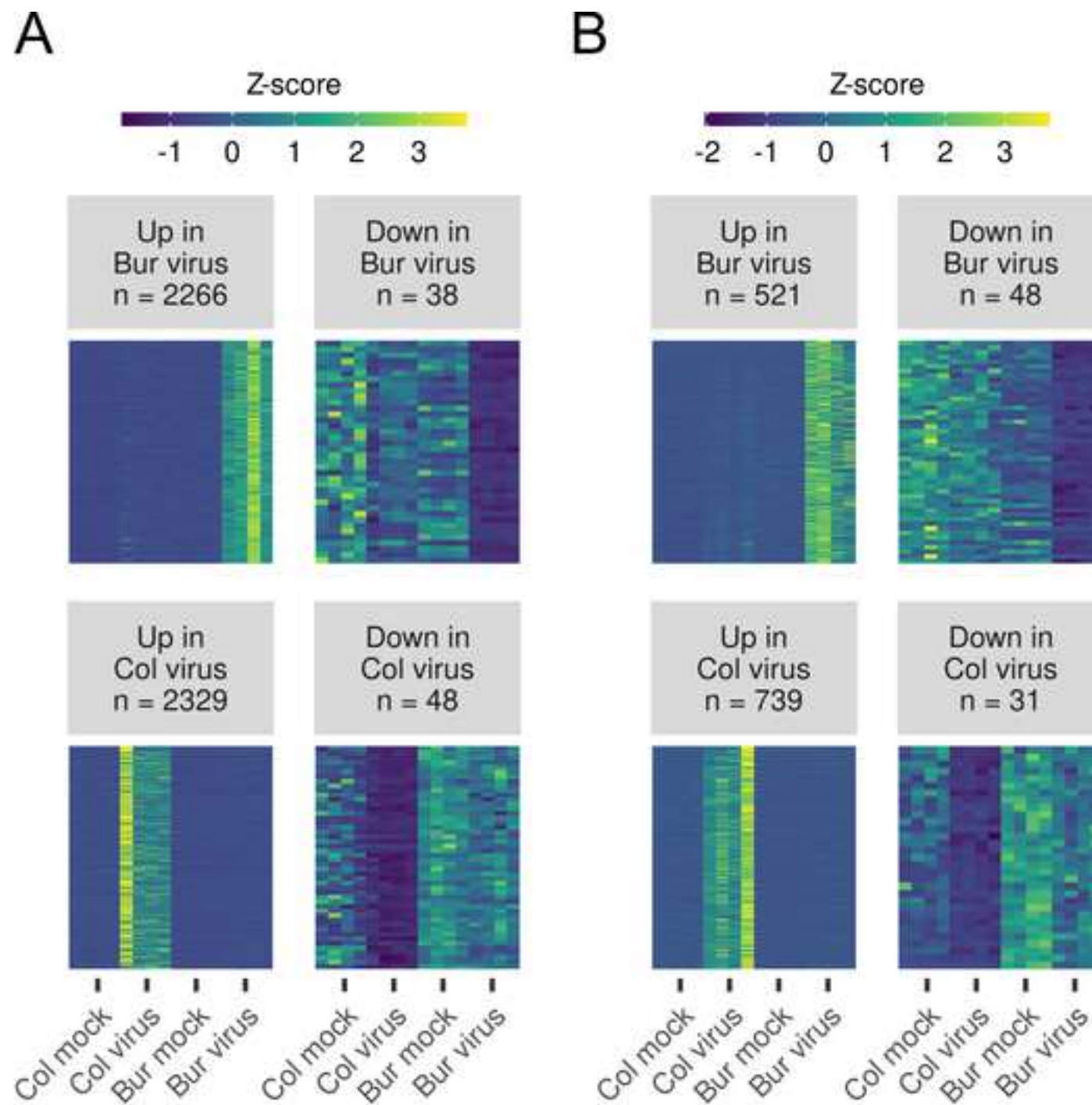
[Click here to access/download;Figure;Fig5.tif](#)

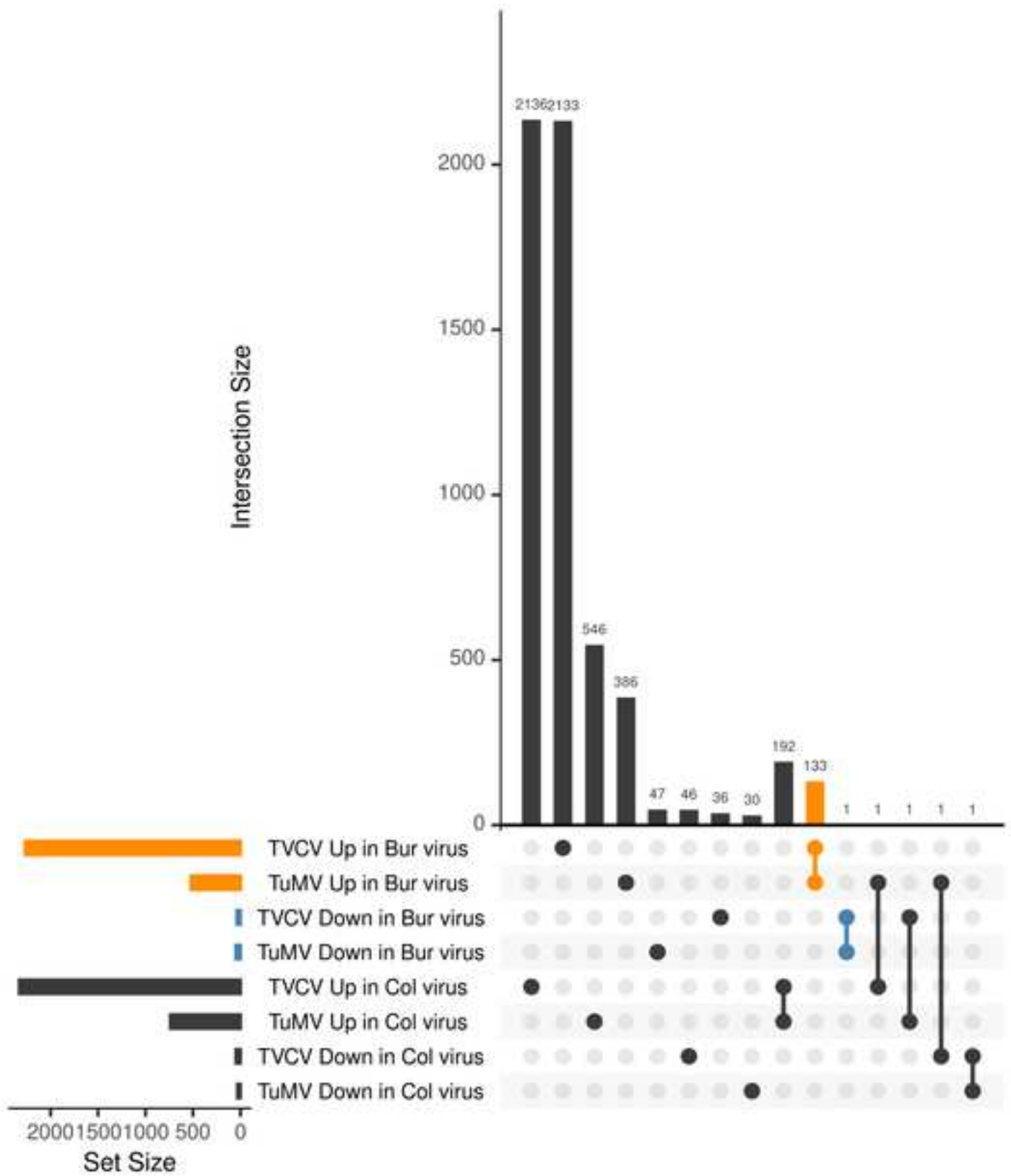




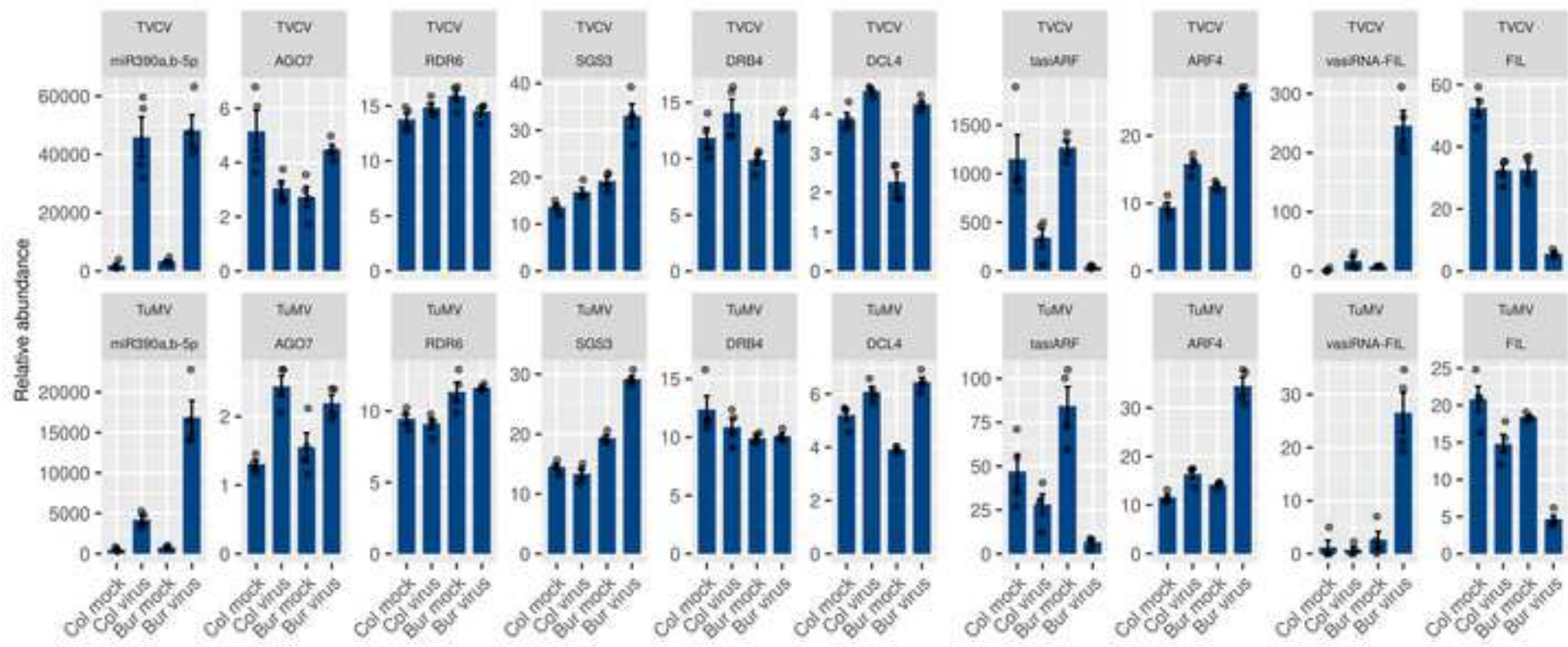




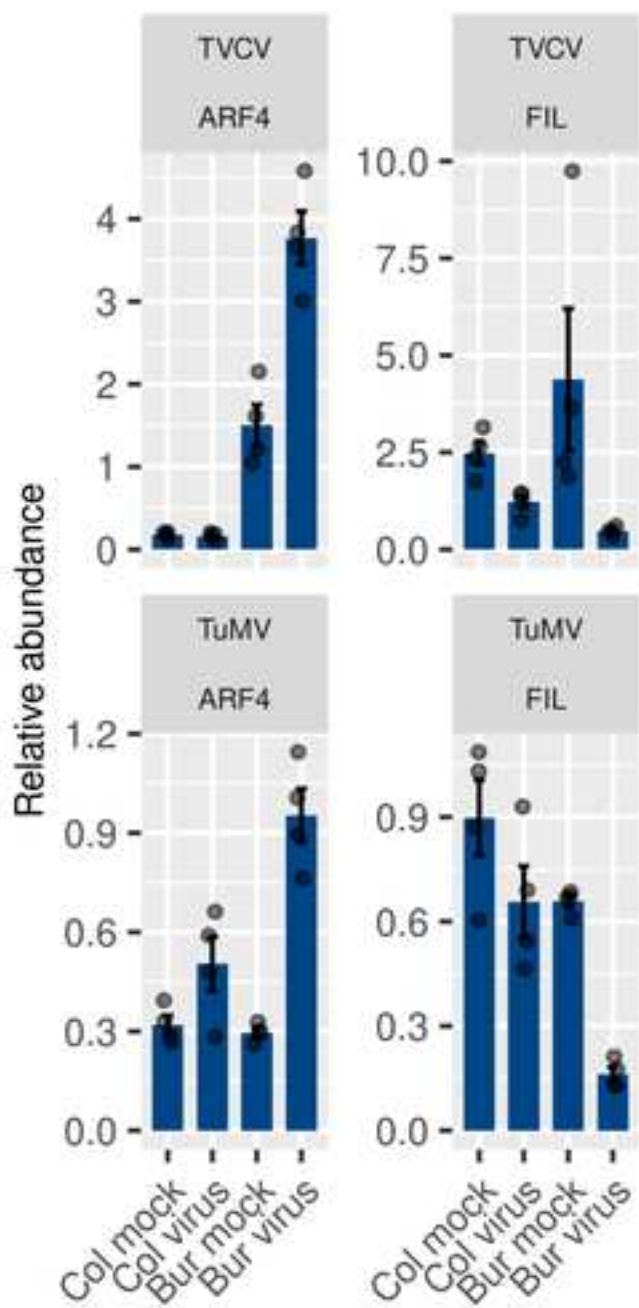








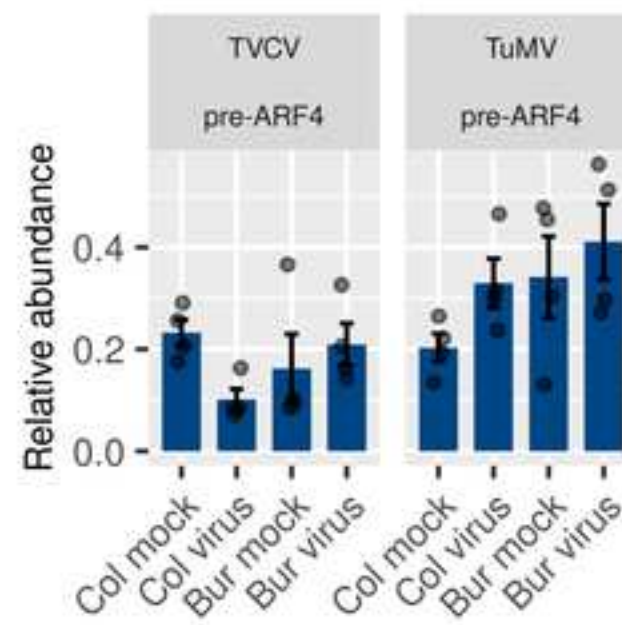
A

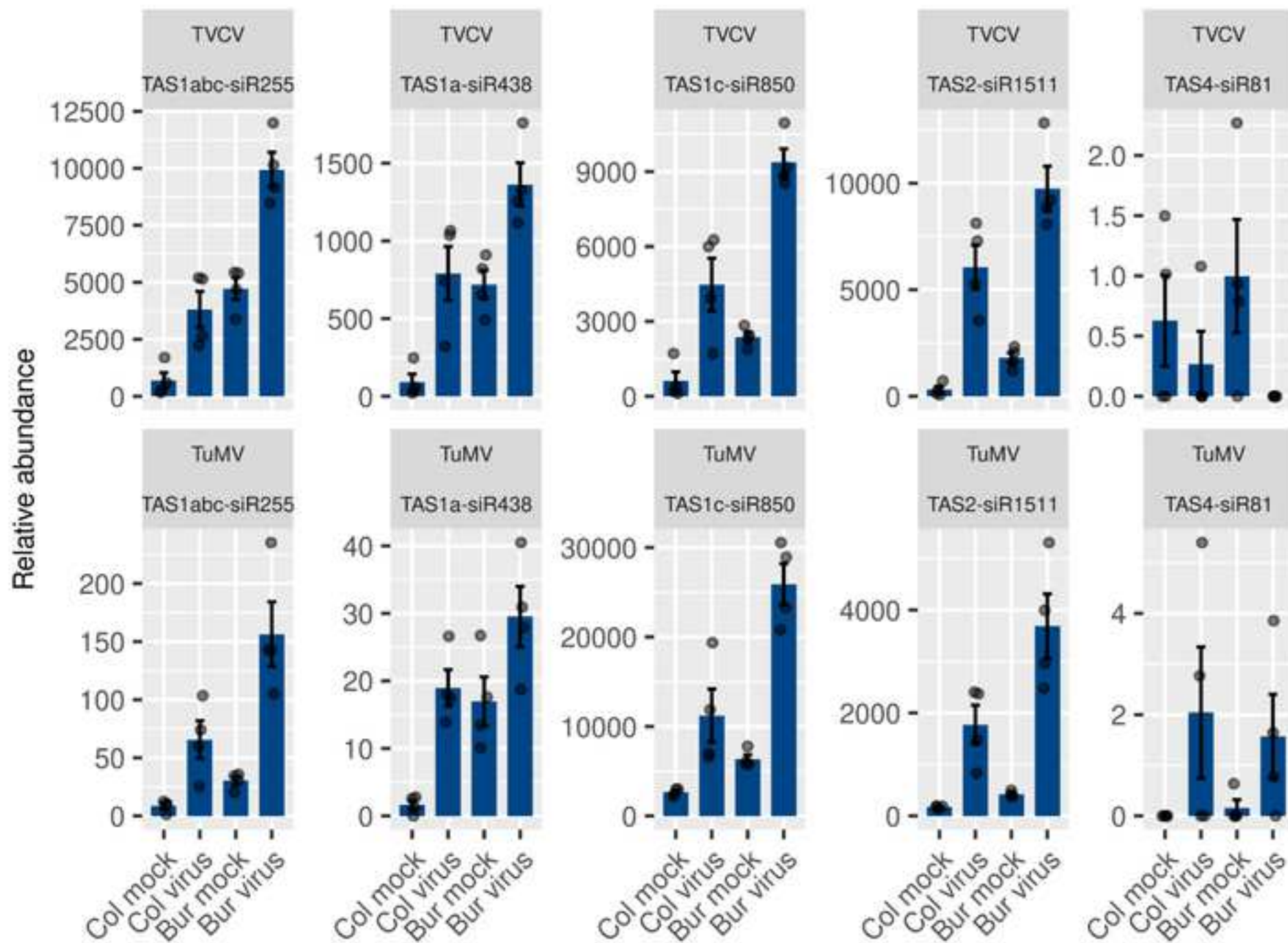


B

mock vs virus	<i>ARF4</i>	<i>FIL</i>
TV CV Col	0.583160	0.009227
TV CV Bur	0.001403	0.077041
TuMV Col	0.078988	0.157117
TuMV Bur	0.000208	0.000001

C











Click here to access/download  
**Supporting Information**  
Table\_S1.xlsx



Click here to access/download  
**Supporting Information**  
Table\_S2.xlsx



Click here to access/download  
**Supporting Information**  
Table\_S3.xlsx



Click here to access/download  
**Supporting Information**  
Table\_S4.xlsx

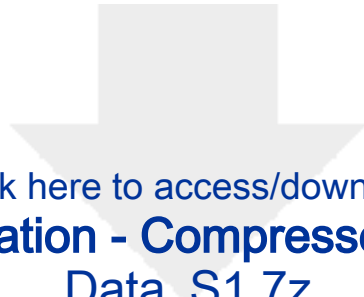


Click here to access/download  
**Supporting Information**  
Table\_S5.xlsx



Click here to access/download  
**Supporting Information**  
Doc\_S1.pdf





[Click here to access/download](#)

**Supporting Information - Compressed/ZIP File Archive**  
**Data\_S1.7z**



# Ecotype-specific blockage of tasiARF production by two different RNA viruses in *Arabidopsis*

Péter Gyula<sup>1¶\*</sup>, Tamás Tóth<sup>1,2¶</sup>, Teréz Gorcsa<sup>1</sup>, Tünde Nyikó<sup>1</sup>, Anita Sós-Hegedűs<sup>1</sup>, and  
György Szittya<sup>1\*</sup>

<sup>1</sup> Department of Plant Biotechnology, Institute of Genetics and Biotechnology, Hungarian  
University of Agriculture and Life Sciences, Gödöllő, Hungary

<sup>2</sup> Doctoral School of Biology, Institute of Biology, ELTE Eötvös Loránd University,  
Budapest, Hungary

\* Corresponding author

E-mail: gyula.peter@uni-mate.hu (GP) and szittya.gyorgy@uni-mate.hu (GS)

¶ These authors contributed equally to this work.



## Abstract

*Arabidopsis thaliana* is one of the most studied model organisms of plant biology with hundreds of geographical variants called ecotypes. One might expect that this enormous genetic variety could result in differential response to pathogens. Indeed, we observed previously that the Bur ecotype develops much more severe symptoms (upward curling leaves and wavy leaf margins) upon infection with two positive-strand RNA viruses of different families (turnip vein-clearing virus, TVCV, and turnip mosaic virus, TuMV). To find the genes potentially responsible for the ecotype-specific response, we performed a differential expression analysis of the mRNA and sRNA pools of TVCV and TuMV-infected Bur and Col plants along with the corresponding mock controls. We focused on the genes and sRNAs that showed an induced or reduced expression selectively in the Bur virus samples in both virus series. We found that the two ecotypes respond to the viral infection differently, yet both viruses selectively block the production of the *TAS3*-derived small RNA specimen called tasiARF only in the virus-infected Bur plants. The tasiARF normally forms a gradient through the adaxial and abaxial parts of the leaf (being more abundant in the adaxial part) and post-transcriptionally regulates **ARF4**, a major leaf polarity determinant in plants. The lack of tasiARF-mediated silencing could lead to an ectopically expressed **ARF4** in the adaxial part of the leaf where the misregulation of auxin-dependent signaling would result in an irregular growth of the leaf blade manifesting as upward curling leaf and wavy leaf margin. QTL mapping using Recombinant Inbred Lines (RILs) suggests that the observed symptoms are the result of a multigenic interaction that allows the symptoms to develop only in the Bur ecotype. The particular nature of genetic differences leading to the ecotype-specific symptoms remains obscure and needs further study.

## Introduction

Virus-infected plants may develop various symptoms on all or some of their parts. The symptoms may either be nonspecific or quite characteristic of the particular virus. Common symptoms are dwarfing, stunting, various leaf deformations, chlorotic mosaics, vein clearing, etc. These effects may either be 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 environmental conditions, like temperature and light. Finally, plants may show acute or severe symptoms soon after inoculation that may lead to the 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 a gradual decline of the plant [1].

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 virally encoded proteins become active. Movement proteins help the viral nucleic acid move into another cell through plasmodesmata [2]. At the site of entry, plants **carrying appropriate R genes may** develop a hypersensitive response that involves the overproduction of reactive oxygen species and programmed cell death which prevents further spreading of the virus into the neighboring 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, and necrosis.

To restrict viral accumulation, plants and other eukaryotes utilize a highly adaptive, sequence-specific mechanism called RNA silencing [3]. 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) intermediate product of viral replication trigger antiviral silencing [4]. The dsRNA is randomly cut into 21–22 bp-long fragments called virus-derived short interfering RNAs (vsiRNAs) by specific Dicer-like (DCL) endonucleases [5]. One of the strands of these fragments called the guide strand is then loaded into an RNA-induced silencing complex (RISC). With 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 efficient virus infection, viruses should be able to evade the silencing machinery. In the evolutionary arms race, viruses developed a plethora of different viral silencing suppressor (VSR) molecules that can interfere with various steps of the silencing pathway, including small RNA production, processing, binding, stability, and activity of the RISC [6–8]. As a side effect, VSRs can interfere with endogenous sRNA-regulated processes that result in a corrupted development of the plant.

One major class of regulatory sRNAs is the micro RNAs (miRNAs) which are usually 20–22-nt-long sequences produced by a DCL1-mediated dicing of a hairpin-like, double-stranded precursor with imperfect self-complementarity. Different AGOs preferentially bind miRNAs with a specific 5' nucleotide [9,10] and length while the loading efficiency is controlled by the structure of the double-stranded precursor [11]. Viruses often modify the miRNA landscape of the plant which affects symptom development [12–14].

Many aspects of plant development where a pattern formation is required are regulated by cell-to-cell mobile miRNAs or siRNAs [15–17]. This cell-to-cell movement results in a sharp gradient of the sRNA within a tissue, for example, between the adaxial and abaxial parts of the leaf, establishing patterns of cell specification by creating boundaries that limit

the activity of target genes [17,18]. A conserved regulatory module among the land plants that plays a central role in abaxial-adaxial leaf patterning, juvenile-to-adult phase transition, floral organ development, and lateral root development is the miR390/AGO7/TAS3/tasiARF/ARF3/ARF4 module [19–28]. The tasiARF (*trans*-acting ARF-targeting small interfering RNA) is a special kind of small interfering RNA (siRNA) that is produced from the *TAS3* non-coding RNA (ncRNA) precursor. During the production of tasiARF, the *TAS3* transcript is first cleaved by a miR390-loaded AGO7 [29], then, a double-stranded intermediate product is produced by the RNA-dependent RNA polymerase RDR6/SGS3/DRB4 complex. This dsRNA is then diced in a phased manner, mostly by DCL4, producing a set of 21-nt-long siRNAs. From this siRNA population, only a few sequences are conserved evolutionary. The conserved tasiARF is loaded into AGO1 and post-transcriptionally regulates the transcripts of *ARF3* and *ARF4* transcription factor genes, the protein product of which negatively regulates auxin-dependent gene expression [20,23,30]. The tasiARF is cell-to-cell mobile and forms a gradient across the leaf blade being more abundant on the adaxial side [18,31]. This results in the confinement of *ARF4* in the abaxial side of the leaf and contributes to the determination of abaxial cell identity.

Another sRNA-mediated signaling pathway that is conserved among land plants is the miRNA-regulation of the large family of *NUCLEOTIDE-BINDING LEUCINE-RICH REPEAT CONTAINING RECEPTOR* genes (NB-LRRs) which are important regulators of innate immunity. They recognize specific pathogen effectors and trigger resistance responses [32,33]. The members of this gene family are targeted by various, unrelated, lineage-specific 22-nt-long miRNAs, some of which are more widespread (i.e. the miR482/2118 superfamily) [34], while others can be found only in certain plant groups [35]. These genes are constantly repressed by the above-mentioned miRNAs but relieved from repression when the level of the regulatory miRNA is lowered upon pathogen infection. Interestingly, in the presence of a symbiont, the miRNA level is changed in the opposite direction leading to an even lower

level of NB-LRR genes and enhanced nodulation in *Medicago truncatula* [36]. The miRNA cleavage of the transcripts can initiate the production of secondary phased siRNAs: dsRNA intermediates are produced by the RDR6/SGS3/DRB4 complex and then diced by DCL4/DCL2, similarly to the case of *TAS3*. These secondary siRNAs can also target related genes amplifying the immune response [37].

There are other tasiRNA- or phased siRNA-producing loci in plants that are less conserved as the *TAS3* regulatory module [30]. In *Arabidopsis thaliana*, *TAS1a,b,c*, and *TAS2* noncoding RNA are cleaved by miR173-loaded AGO1 [38], while *TAS4* is cleaved by miR828 [39] and secondary siRNAs are produced as described above. Some siRNAs emerging from these loci target host genes: *TAS1*-derived siRNAs can target *HEAT-INDUCED TAS1 TARGET 1* and *2 (HTT1/2)* [40], *TAS2*-derived siRNAs can target several *PENTATRICOPEPTIDE REPEAT-CONTAINING PROTEIN (PPR)* genes [30], while a *TAS4*-derived siRNA can target a set of MYB transcription factor genes [39]. Other secondary phased siRNA-producing loci are mostly protein coding genes and include the miR393-initialized auxin receptor genes *TRANSPORT INHIBITOR RESPONSE 1 (TIR1)* and *AUXIN SIGNALING F-BOX 1–3 (AFBI–3)* that regulate auxin signaling [41].

The presence of the virus also triggers the synthesis of a different class of siRNAs (virus-activated siRNAs, vasiRNAs) that regulate other host genes and modulate the antiviral response of the plant [42], for example, by downregulating genes coding for different functions of the photosynthetic apparatus. Lowering the energy production rate within the virus-infected plant cell would restrict the replication of the virus which requires a lot of energy [43]. This conserved regulatory action of vasiRNAs against the host metabolism could explain the observation that almost all virus-infected plants show some degree of retarded growth. The production of vasiRNAs was shown to be dependent on RNA-DEPENDENT RNA POLYMERASE 1 (RDR1) and DICER-LIKE 4 (DCL4) [42] making them unique among other endogenous siRNAs like trans-acting siRNAs (tasiRNAs) or phased siRNAs

(phasiRNAs) which depend on the action of RDR6 and DCL4/DCL2. Unlike tasiRNAs, vasiRNAs do not require a phase initiating miRNA- or siRNA-mediated cleavage and, therefore, they are not produced in a phased manner. Pitzalis et al [44] proposed a mechanism for the triggering of vasiRNA production: virus infection could induce a disruption or overload of RNA decay pathways, thereby leading to the accumulation of aberrant transcripts that are prone to dsRNA formation and processing into siRNAs. Most of the vasiRNAs are 21-nt-long [42] but in some cases, for example, during infection by the pararetrovirus CaMV, a significant amount of 22-nt-long vasiRNAs are also produced in three different species of the *Brassicaceae* family, probably due to the activity of DCL2 [43]. The vasiRNAs can target the cleavage of the same mRNA that they were derived from or other genes *in trans*, similarly to tasiRNAs.

Viral infections often lead to severe diseases in plants that result in significant economic loss. Tobamoviruses are among the most studied plant viruses because they are widespread and infect economically important plants, including the large family of *Solanaceae* (i.e., potato, tomato, pepper, eggplant, etc.) and *Brassicaceae* (i.e., cruciferous vegetables like cabbage, turnip, mustard, rapeseed, etc.). The model plant of *Brassicaceae* is *Arabidopsis thaliana* (thale cress) which has several geographical varieties called ecotypes. We previously characterized a new isolate of *Turnip vein-clearing virus* (TVCV-ApH) [45], a crucifer-infecting tobamovirus [46,47], that caused severe leaf deformation (upward curling leaves) on the Bur ecotype of *Arabidopsis thaliana*, but not on Col ecotype. The 6.3 kb genome of TVCV-ApH contains four ORFs. ORF1 codes for the 125 kDa small replicase subunit, while the suppression of an amber termination codon of ORF1 results in the 182 kDa large replicase subunit (ORF2). The ORF3 and ORF4 translate from 3' subgenomic RNAs and result in a 30 kDa movement protein, and a 17 kDa coat protein, respectively [45]. Several reports demonstrated that besides its role in replication, the p125 protein acts as a viral silencing suppressor through the binding of vasiRNAs [48–50]. We also observed this

ecotype-specific leaf deformation in the case of turnip mosaic virus (TuMV) infection. TuMV is a member of the *Potyviridae* family and it is the most important virus infecting cruciferous crops [51]. The TuMV genome is around 9.8 kb and the genomic RNA is translated into a large polyprotein and a frame-shift protein. The large polyprotein is subsequently processed by the action of viral-encoded HCPro protease into ten mature functional products [52,53]. A frame-shift protein, P3N-PIPO, was reported to be involved in the pathogenesis and movement of TuMV [54,55]. The TuMV HCPro protein has RNA silencing suppressor (RSS) activity [56] through the binding of vsRNAs [57,58]. The ecotype-specific leaf deformation phenotype observations with two different RNA virus infections suggested that there are genetic differences in the hosts that determine this response. We also expected that this differential regulation will manifest in the transcriptome or sRNAome profiles and by comparing these profiles we will be able to find the host factors responsible for the ecotype-specific symptom development. Indeed, the comparative transcriptome analysis of the TVCV-, TuMV-, and mock-infected Col and Bur plants resulted in a few candidate genes that could be responsible for the leaf deformation. Some of these genes are known regulators of leaf polarity while others may represent novel, ecotype-specific functions in leaf development or other biological processes that are commonly disturbed by TVCV and TuMV infection.

## **Materials and methods**

### **Plant growth conditions and viral infection**

Seeds of *Arabidopsis thaliana* Col and Bur ecotypes were sown on soil in pots, watered, wrapped with cling film, and kept in the dark for two days. After that, pots were moved into a growth chamber (SANYO/Panasonic MLR-352-PE) where plants grew under long-day conditions (16 h light, 8 h dark, constant 21 °C) for two weeks. Then, individual seedlings

were transferred into Jiffies ([www.jiffypot.com](http://www.jiffypot.com)) and grown for further two weeks in the same chamber. The four-week-old seedlings were transferred into pots filled with soil and moved into a lightroom where they grew for one day under long-day conditions. The next day, in the case of TVCV-ApH, plants were inoculated as described before [45]. Briefly, 2  $\mu\text{L}$  (1  $\mu\text{g}$ ) of either total RNA purified from systemic leaves of *Nicotiana benthamiana* infected with an infectious *in vitro* transcript of TVCV-ApH or water (mock) was combined with 5.5  $\mu\text{L}$  water and 7.5  $\mu\text{L}$  2 $\times$  inoculation buffer and was rub-inoculated into the true leaves of four-week-old *Arabidopsis* plants. In the case of TuMV, *Arabidopsis* plants were infected with extracts of *N. benthamiana* leaves infected with an infectious *in vitro* transcript of TuMV. For both viruses, three leaves were inoculated per plant and 60 plants were inoculated per sample (Col mock, Col virus, Bur mock, and Bur virus). After inoculation, the plants were sprayed with water and grown under long-day conditions in the same lightroom.

## **Purification of RNA samples for high-throughput sequencing**

For total RNA purification, two symptomatic systemic leaves were collected from every infected plant 14 day-post-inoculation (dpi) between ZT4 and ZT6. For the mock samples, leaves of similar age as the infected ones were collected. Total RNA was purified from individual leaves (480 samples altogether) using a phenol-chloroform extraction method [59]. The purified RNA samples were quantified by Nanodrop photometer and 1  $\mu\text{g}$  of total RNA per sample was run on a denaturing agarose gel. The gel was stained with ethidium bromide and photographed with ChemiDoc™ (Bio-Rad, Hercules, CA, USA). The relative amount of the viral RNAs that were visible over the ribosomal RNA bands was quantified with Image Lab v5.1 software (Bio-Rad, Hercules, CA, USA) using the 18S rRNA bands as the reference. Seven samples with similar levels of viral genomic RNAs were pooled together to form a biological replicate. In this way, four replicates were created.



## **High-throughput sequencing of mRNA and sRNA populations**

The libraries for the RNA-seq and sRNA-seq from the TVCV-infected samples were prepared and sequencing reactions were carried out by LC Sciences LLC (Houston, TX, USA). The polyA-selected RNA-seq libraries were prepared using the TruSeq Stranded mRNA Library Prep Kit (Illumina, San Diego, USA) and sequenced with a paired-end 2×150 bp chemistry on Illumina HiSeq X platform. The sRNA-seq libraries were prepared using the TruSeq Small RNA Sample Prep Kit (Illumina, San Diego, USA) and sequenced on an Illumina HiSeq X platform with a single-end 50 bp chemistry.

PolyA-selected, stranded RNA-seq libraries for the TuMV samples were prepared and sequenced by Novogene Co., Ltd. (Cambridge, UK) on a NovaSeq 6000 sequencing platform with a paired-end 2×150 bp chemistry. The sRNA libraries (including a size-separation by PAGE) were prepared in our laboratory using the TruSeq Small RNA Sample Prep Kit (Illumina, San Diego, USA) and were sequenced by Novogene Co., Ltd (Cambridge, UK) on a NovaSeq 6000 platform with a single-end 50 bp chemistry.

The raw sequencing data were deposited in the NCBI SRA database (BioProject: PRJNA788379) under the identifiers SRR17227515–SRR17227578.

## **Bioinformatic analysis of high-throughput sequencing data**

### **RNA-seq analysis**

The quality of the raw reads was checked with FastQC v0.11.8 [60]. Trimming of adapter sequences and filtering of low-quality reads were carried out with cutadapt v2.8 [61]. The clean reads were aligned to the AtRTD2 reference transcriptome [62] and the normalized expression values (transcript per million, TPM) were calculated with kallisto v0.44.0 [63] with the following parameters: -b 10 --bias --rf-stranded. The TPM values calculated for the alternative transcript isoforms were aggregated to represent gene-level expression and were

normalized between samples with sleuth v0.29.0 [63]. To test whether there is a significant difference between the means of the Bur (or Col) virus-infected and all the other samples, a Wald test was applied. The calculated  $P$ -values were corrected for multiple testing using the Benjamini-Hochberg method [64]. Genes with  $qval$  (corrected  $P$ -value, at 1% false discovery rate)  $< 0.01$ ,  $b$  (estimated fold-change)  $> \log(2)$ , and  $mean\_obs$  (mean raw expression level)  $> 1$  were considered as differentially expressed genes (DEGs).

The DEG expressions were visualized in a heatmap, for which  $Z$ -scores were calculated in the following way: for every gene, the mean and standard deviation of the expression across samples were calculated, then from every individual value the mean was subtracted and the resulting value was divided by the standard deviation. The  $Z$ -score tells how many standard deviations an expression value is from the mean expression of an individual gene. In this way, genes with different mean expression levels can be compared. The heatmaps were generated with the `geom_tile` function of the `ggplot2` R package [65].

Principal component analysis of the RNA-seq data was performed using the `prcomp` R package. Normalized gene expression values (transcript per million, TPM) were scaled ( $\log_2$ -transformed) and centered before the analysis. The plots were generated with the `autoplot` function of the `ggfortify` R package [66].

Comparison of the DEG sets from different conditions was analyzed with the `UpSet` R package [67].

The Gene Ontology (GO) term analysis of the DEGs was performed using the PlantRegMap server [68] with the default settings. The enrichment factor was calculated by dividing the observed number of genes in a particular category by the expected number of genes in that category. The plot was generated with the `geom_point` function of the `ggplot2` R package [65].

## Small RNA-seq analysis

The raw sequences were processed and filtered using cutadapt v2.8 [61] with the following parameters: `-a TGGGAATTCTCGGGTGCCAAGG -m 20 -M 25 -q 20 --max-n=0 --discard-untrimmed`. The trimmed sequences were further filtered to remove tRNA- and rRNA-derived sequences using the short read aligner bowtie2 v2.3.5.1 with the following parameters: `-k 1 -D 20 -R 3 -N 1 -L 10 -i S,1,0.50` and a reference sequence set containing tRNA, rRNA, snRNA, and snoRNA sequences from *Viridiplantae*. The unaligned sequences were collected and aligned to the *Arabidopsis thaliana* TAIR10 reference genome with ShortStack v3.8.5 [69] with the following parameters: `--align_only --bowtie_m all --ranmax none --keep_qual`. The raw count table of the genome-mapped sRNA sequences was produced using the fasta and view functions of the samtools suite v1.9 [70] and custom shell scripts. The sRNS-producing loci were predicted using ShortStack v3.8.5 with the following parameters: `--dicermin 20 --dicermax 25`. The normalization and differential expression analysis of the sequences or the sRNA loci were performed with the DESeq2 R package [71]. Only those sequences/loci were accepted that passed through the following filter: *padj* (corrected *P*-value, at 1% false discovery rate)  $< 0.05$ ,  $\log_2\text{FoldChange} > \log(1.5)$ , and *baseMean* (mean raw expression level)  $> 1$ . The DESeq2 normalization factors were used for the calculation of the sequence length distribution plot.

The sRNA target analysis was performed using the psRNATarget server [72] with the default settings except that the target accessibility calculation was allowed. We provided the 114 Bur virus upregulated sRNAs and the 88 Bur virus downregulated transcripts that were extracted from the AtRTD2 transcripts [62] using the notseq function of the EMBOSS package v6.6 [73]. Only the predicted cleavage hits were considered for further analysis.

## RT-qPCR analysis

The primers to measure *ARF4* (AT5G60450) expression (ARF4\_qPCR\_F: 5'- GCT CCT CTT GAC TAC AAA CAA CAG -3', ARF4\_qPCR\_R: 5'- GGC GAA ACT TCC ACT CTA CTC C -3') were designed with PerlPrimer v1.2.4 [74] in a way that at least one primer must span an intron to prevent amplification from genomic DNA. The *FIL* (AT2G45190) was amplified using the primers described before (FIL\_qPCR\_F: 5'- TGG TAC AGC AAC CAC ATC GGA CAG -3', FIL\_qPCR\_R: 5'- GCC AAA CCA TCC TTG CGG TTA ATG -3') [75]. The *PDF2* gene (AT1G13320) was used as an internal reference, primers to measure *PDF2* were published earlier (PDF2\_qPCR\_F: 5'- TCA TTC CGA TAG TCG ACC AAG -3', PDF2\_qPCR\_R: 5'- TTG ATT TGC GAA ATA CCG AAC -3') [76].

For the first-strand cDNA synthesis, 1 µg of total RNA was treated with DNase I, then reverse transcribed with RevertAid H Minus Reverse Transcriptase (Thermo Fisher Scientific, Waltham, MA, USA) and a random primer according to the manufacturer's instructions. The cDNAs were diluted ten times, and 1 µL was used in a 10 µL qPCR. The qPCRs were performed using SYBR™ Select Master Mix (Thermo Fisher Scientific, Waltham, MA, USA), following the manufacturer's instructions. The reactions were run in a LightCycler® 96 Real-Time PCR machine (Roche, Basel, Switzerland). The reaction profile was the following: preincubation at 95 °C for 1 min; amplification at 95 °C for 15 s and 60 °C for 30 s, repeated 45 times; melting at 95 °C for 10 s, then 65 °C for 1 min and continuous heating to 97 °C; cooling at 37 °C for 30 s. Four biological and two technical replicates were measured per sample. The values for every biological replicate were calculated as the mean of the two technical replicates. The relative expression values were calculated by the LightCycler96® v1.1 software (Roche, Basel, Switzerland). Unpaired, two-tailed Student's *t*-tests (at 0.95 confidence level) were performed using the *t.test* R package to assess if there

is a significant difference between the means of the mock- and virus-infected samples in the Col and Bur plants either in the TVCV or TuMV series.

## **QTL mapping**

To obtain a coarse genetic map, a publicly available Minimal Recombinant Inbred Line set (INRA, Versailles, France) derived from a cross of *Arabidopsis thaliana* Bur and Col ecotypes was amended with 23 more lines selected randomly from the remaining lines [77]. Altogether, 43 lines and the two parental lines were used in the experiment. Five plants per line were infected either with TVCV-ApH or mock as described above and the leaf deformation (upward curling) was scored from 0 to 5 (0 = infected Col, 5 = infected Bur). The average score of the five plants was recorded and used for assessing correlation with the genotypes of the RILs that were obtained from INRA (<http://publiclines.versailles.inrae.fr/page/20>). The analysis was carried out with Windows QTL Cartographer v2.5\_011 [78] using the single marker association test.

## **Results**

### **TVCV and TuMV infection causes leaf deformation only on Bur ecotype**

We observed earlier that upon infection with turnip vein-clearing virus (TVCV, a tobamovirus), the Bur ecotype of *Arabidopsis thaliana* displayed much more severe symptoms than Col [45]. The most obvious symptom was the leaf deformation which included polarity defect (upward curling leaves) and wavy leaf margin that could not be observed on Col plants under the conditions tested (Fig 1). We were curious whether this difference in symptom development is specific to this virus-host interaction or can be

observed with other viruses as well. Therefore, we infected the above-mentioned two *Arabidopsis* ecotypes with turnip mosaic virus (TuMV, a potyvirus). We observed the same difference in symptom development between the two ecotypes, namely, the Bur plants showed upward curling leaves while the Col plants did not have leaf deformity (Fig 1).

**Fig 1. Viral symptoms of TVCV- and TuMV-infected *Arabidopsis thaliana* Col and Bur ecotypes.**

Four-week-old plants grown under long-day conditions at 21°C were infected either with virus or mock solutions. After two weeks, plants were photographed and samples for RNA purifications were collected from mock- and virus-infected plants. For the RNA purification, samples were taken from symptomatic leaves of Bur or leaves of similar developmental stages from the Col and mock samples. To obtain one replicate, seven leaves from different plants were pooled (one leaf per plant). Four replicates were prepared in such a way for every mock and virus-infected sample.

Both viruses replicated and reached similar viral RNA levels in each ecotype, suggesting that viral replication, short-distance, and long-distance viral movement are not involved in the observed symptom differences. The differential leaf development defect in the two ecotypes upon both virus infections suggests that there are host factors whose ecotype-specific misregulation leads to an altered leaf developmental program only in the Bur plants.

**TVCV and TuMV infection elicit different transcriptome responses in Col and Bur ecotypes**

Upon viral infection, the transcriptome of the infected host changes significantly [79–83]. The virus modulates the host environment in favor 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 virus similarly in many ways. However, since there were significant differences in the observable symptoms, we anticipated that there would be gene expression changes that reflect these differences. In order to reveal these differential gene expression profiles, we collected leaves from TVCV- and TuMV-infected *Arabidopsis Col* and *Bur* plants along with mock-infected ones in four biological replicates for RNA-seq experiments and purified total RNA from them. Next, the polyA-tail mRNA fraction was sequenced either on an Illumina HiSeq X (TVCV) or a NovaSeq 6000 (TuMV) platform. The clean reads were mapped to the AtRTD2 *Arabidopsis thaliana* reference transcriptome [62] with kallisto [63] and the gene expression values were normalized between samples with sleuth [63] (Table S1). We performed a principal component analysis to reveal which experimental variables are responsible for the most variances between samples (Fig 2).

**Fig 2. Principal Component Analysis of the RNA-seq data.**

PCA was performed using the prcomp R package. Normalized gene expressions (Transcript Per Million, TPM) were scaled and centered before the analysis. (A) The first component (PC1) separates the samples by ecotype and explains 36.85% of the total variances between the samples while the second component (PC2) separates the samples by virus series and explains further 23.35% of the variances. (B) The third component (PC3) separates samples that are either mock or virus-infected, which is responsible for a further 8.06% of the variances. The four biological replicates are marked with the same color for easier identification.

According to the analysis, the most important factor that separates samples and is responsible for 38.23% of the total variances (PC1) is the ecotype, while the second most important factor (PC2) separates the TVCV and TuMV series (Fig 2A). This latter is either

because of the different technology that was used for sequencing the two series in two different laboratories or could reflect a true biological difference (or both). The third most important factor (PC3) that separates the mock- and virus-infected samples is responsible only for 8.06% of the total variances observed between samples (Fig 2B). This figure also shows that the mock- and virus-infected samples in the Bur ecotype are more separated than in the Col ecotype (Fig 2B). We interpret these results that there is a substantial difference between the two ecotypes in their responses to viral infection.

## **Finding the genes responsible for the common leaf polarity**

### **defect in the TVCV- and TuMV-infected Bur plants**

To find out which genes are responsible for the upward curling leaves in the TVCV- or TuMV-infected Bur ecotype, we performed a differential gene expression analysis using sleuth [63], separately for the two virus series. We expected that genes responsible for the leaf deformation that was observed only in the virus-infected Bur ecotype will either be induced or reduced only in the Bur virus samples both in the TVCV and TuMV series compared to all the other samples in the series, namely, Bur mock, Col mock and Col virus samples. Alternatively, it is possible that genes are preferentially induced or reduced in the Col virus samples that prevent symptom development. We considered genes differentially expressed only if the  $Q$ -value from the statistical test was lower than 0.01 (1% false discovery rate), the absolute value of the effect ( $b$  parameter, or estimated fold-change, log-transformed) was higher than  $\log(2)$ , and the mean raw expression ( $mean\_obs$ ) was higher than 1. In this way, we got a list of the differentially expressed genes for the Bur virus or Col virus samples for the TVCV and TuMV series (Fig 3).



**Fig 3. Expression pattern of the differentially expressed genes between Bur and Col plants infected either with TVCV or TuMV.**

To get a list of differentially expressed genes (DEGs) that significantly change only in the Bur (or Col) virus-infected samples, a Wald-test was applied (i.e., Bur virus samples against all the others). Only genes with  $qval < 0.01$ ,  $b$  (estimated fold-change)  $> \log(2)$  and  $mean\_obs$  (mean raw expression level)  $> 1$  were accepted as DEGs. The number of DEGs is noted in the title. Normalized expression values (transcript per million, TPM) of individual, differentially expressed genes in (A) TVCV and (B) TuMV-infected samples were centered by calculating Z-scores. Z-score tells how many Standard Deviations an expression value is from the mean expression of an individual gene. In this way, genes with different mean expression levels can be compared. The expressions of DEGs in samples are represented in a heatmap. The expressions of the four biological replicates are shown for every sample.

There were 278 genes that were up- and 548 that were downregulated in the Bur virus samples, while 386 and 162 genes were up- and downregulated in the Col virus samples in the TVCV series (Fig 3A, Table S2). These numbers were 184, 317, 456, and 115 in the TuMV series, respectively (Fig 3B, Table S2). To filter for the genes that were regulated similarly in the two virus series, we performed a set analysis with the UpSet R package (Fig 4).

**Fig 4. Commonly regulated genes in TVCV and TuMV-infected plants.**

Sets of DEGs in eight different conditions were compared and visualized using the UpSet R package. The vast majority of the DEGs were specific to one condition (represented by single dots), while the number of common elements between the different conditions (connected dots) was much smaller. The most abundant common elements were observed between the sets of downregulated genes in Bur virus samples both in TVCV and TuMV-infected plants

(blue) and the upregulated genes in Bur virus samples both in TVCV and TuMV-infected plants (orange). All the other common elements were much smaller in size, sets with zero elements are not shown.

According to this analysis, the majority of the differentially expressed genes were unique to a certain set and probably not responsible for the commonly observed leaf polarity misregulation. The most abundant intersection with 88 genes was the one that contained downregulated genes in Bur virus samples both in the TVCV and TuMV series (Table S3b), while the second most abundant with 36 genes was the one containing the upregulated genes in Bur virus samples in both virus series (Table S3a). All the other intersections were much smaller and not relevant for our investigation. We performed a GO term enrichment analysis with the above-mentioned two gene sets using the PlantRegMap server [68]. Among the upregulated genes, categories were enriched with moderate or low significance, related to plant cell wall functions (GO:0071554, GO:0071555, GO:0071669, etc), cellular developmental process (GO:0048869), response to hormone (GO:0009725), cell differentiation (GO:0030154) while among the downregulated genes, several terms related to photosynthesis and light regulation (GO:0009768, GO:0009765, GO:0019684, etc), hormone metabolic process (GO:0042445), regulation of hormone level (GO:0010817), regionalization (GO:0003002), pattern specification process (GO:0007389), etc., were significantly enriched (Fig 5, Table S3c–d).

**Fig 5. GO term enrichment analysis of common DEGs of TVCV and TuMV-infected Bur plants.**

The common DEGs of the TVCV and TuMV-infected Bur plants (orange and blue colored sets in Fig 4) were subjected to GO-term enrichment analysis using the PlantRegMap server with the default settings. Only terms in the Biological Process category are shown, the results

of the full analysis can be found in Table S3<sup>c-d</sup>. Enrichment was calculated by dividing the observed number of genes in a particular category by the expected number of genes in that category.

Interestingly, the latter two categories (GO:0003002: regionalization and GO:0007389: pattern specification process) in the downregulated genes contained *FIL* (*FILAMENTOUS FLOWER*, AT2G45190) that encodes a member of the YABBY family of transcriptional regulators involved in abaxial cell type specification in leaves [84]. Actually, *FIL* is the most significantly differentially down-regulated (in the Bur virus samples) gene in the common set of the TVCV and TuMV series (Table S3<sup>b</sup>). *FIL* directly regulates *ARF4* (*AUXIN RESPONSE FACTOR 4*, AT5G60450) [85] that was found in the response to hormone and the cellular developmental process GO categories in the enriched sets of upregulated genes. Furthermore, the important leaf abaxial identity determinant *ARF4* is known to be post-transcriptionally regulated by tasiARF-guided AGO1, as described in the Introduction part. The TAS3-derived tasiARF (trans-acting ARF-targeting small interfering RNA) is a special kind of small interfering RNA (siRNA) that is produced from the TAS3 non-coding RNA (ncRNA) precursor and is a well-known, conserved regulator of leaf polarity (through the regulation of *ARF4*) in land plants [50–56]. During the production of tasiARF, the TAS3 transcript is first cleaved by a miR390 loaded AGO7 [22,57], then, a double-stranded intermediate product is produced by the RNA-dependent RNA polymerase RDR6/SGS3 complex. This dsRNA is then diced in a phased manner, mostly by DCL4, producing a set of 21 nt long siRNAs. From this siRNA population, only a few sequences are conserved evolutionary. The conserved tasiARF targets the transcripts of *ARF2*, *ARF3*, and *ARF4* transcription factor genes, the protein product of which negatively regulates auxin-dependent gene expression [58–60]. The tasiARF is cell to cell mobile and forms a gradient across the leaf blade being more abundant on the adaxial side [61,62]. This results in the confinement

of ARF4 in the abaxial side of the leaf. If this gradient is disturbed by some external or internal factor, for example, virus infection or a mutation, it could manifest in a deformed leaf.

## **Virus-responsive sRNA profiles of the ecotypes are different**

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. We performed a Principal Component Analysis of the normalized sRNA expression values to check the quality of our dataset (Fig 6).

### **Fig 6. Principal Component Analysis of the sRNA-seq data.**

PCA was performed using the prcomp R package. DESeq2-normalized sRNA expressions were scaled and centered before the analysis. (A) The first component (PC1) separates the samples by virus series and explains 29.1% of the total variances between the samples while the second component (PC2) separates the samples by ecotype and explains further 19.88% of the variances. (B) The third component (PC3) separates samples that are either mock or virus-infected, which is responsible for a further 10.46% of the variances. The four biological replicates are marked with the same color for easier identification.

According to the PCA, the first principal component separates samples by the virus series which could either mean that the two viruses elicit quite different responses or could reflect the different sample preparation and sequencing platforms that were used for the generation of data (or both). The second principal component separates the samples by ecotype, while the third one by treatment. The substantial differences in the responses of the two ecotypes to the two viruses suggested that there would be a limited number of commonly regulated sRNAs that could be responsible for the observed common symptoms.

Next, we analyzed the length distribution of the sRNAs in the samples. 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 switches: the 21–22-nt size classes become overwhelmingly dominant [13,86,87,44]. This is mainly due to the 21- and 22-nt viral siRNAs that are produced by DCL4 and DCL2 enzyme activities, respectively, and amplified by RDR1 or RDR6 in *Arabidopsis* [5,88–90]. Analysis of our sRNA libraries showed the same shift in size classes. However, when we filtered out the virus-derived vsRNAs by mapping the sRNA sequences to the *Arabidopsis* genome, the remaining host-derived sRNAs also showed a tremendous 21-nt size class enrichment in the virus-infected samples (Fig 7).

**Fig 7. Sequence-length distribution of the host-derived sRNAs in Bur and Col plants infected either with TVCV or TuMV.**

The raw sRNA sequences were first adapter-trimmed, quality- and length-filtered, then rRNA and tRNA sequences were removed. The remaining sRNA sequences were mapped to the *Arabidopsis thaliana* TAIR10 reference genome with ShortStack. The raw abundances of the genome-mapped sequences were normalized with DESeq2. The normalization factors were used to scale the raw abundances of the size classes measured by FastQC. Error bars represent the standard error of the mean of four biological replicates.

This can either be attributed to the mostly 21-nt-long miRNAs or host-derived, also vastly 21-nt-long virus-activated siRNAs (vasiRNAs), which are produced in an RDR1 and DCL4-dependent manner from plant precursor transcripts [42]. The vasiRNAs regulate host genes that can modulate the antiviral response of the plant. We hypothesized that some vasiRNAs may be produced in an ecotype-specific manner and can regulate the genes that we have found in the comparative transcriptome analysis. To find sRNAs whose expression

either increased or decreased only in the Bur virus samples, we performed a differential expression analysis of the genome-mapped, clean sRNAs which had a mean read count  $> 1$  (102195 unique, 20–25-nt-long sequences) using DESeq2. We applied a filtering rule that was less strict than in the case of RNA-seq analysis ( $\log_2\text{FoldChange} > \log_2(1.5)$ ,  $\text{padj} < 0.05$ ,  $\text{baseMean} > 0$ ) because according to our experience, sRNA expressional changes are usually smaller and the data are noisier. In this way, we got sets of differentially expressed sRNAs in the Bur (or Col) virus samples in the TVCV and TuMV series (Fig 8).

**Fig 8. Expression pattern of the differentially expressed sRNA sequences between Bur and Col plants infected either with TVCV or TuMV.**

To get a list of differentially expressed sRNAs (DEs) that significantly change only in the Bur (or Col) virus-infected samples, a Wald-test was applied (i.e., Bur virus samples against all the others). Only sequences with  $\text{padj} < 0.05$ ,  $\text{Log}_2\text{FoldChange} > \log(1.5)$ , and  $\text{baseMean}$  (mean raw expression level)  $> 0$  were accepted as DEs. The number of DEs is noted in the title. Normalized expression values of the individual, differentially expressed sequences in (A) TVCV and (B) TuMV-infected samples were centered by calculating Z-scores. Z-score tells how many Standard Deviations an expression value is from the mean expression of an individual sequence. In this way, sequences with different mean expression levels can be compared. The expressions of DEs in samples are represented in a heatmap. The expressions of the four biological replicates are shown for every sample.

The analysis revealed that sRNAs tended to be upregulated upon virus infection rather than downregulated: there were 2266 sequences that were selectively upregulated and 38 downregulated in the Bur virus samples in the TVCV series, and 521 and 48 in the TuMV series. These numbers were similar regarding the selectively up- and downregulated sRNAs in the Col virus samples in both virus series (Fig 8). The majority of upregulated sRNAs in

both virus series were 21–22-nt-long while the downregulated sRNAs were mainly 23–24-nt-long (Table S4).

To find commonly regulated sRNAs between the TVCV and TuMV series, we analyzed the differentially expressed sRNAs with the UpSet R package like in the case of RNA-seq data (Fig 9).

**Fig 9. Commonly regulated sRNA sequences in TVCV and TuMV-infected plants.**

Sets of differentially expressed sRNA sequences (DES) in eight different conditions were compared and visualized using the UpSet R package. The vast majority of the DESs were specific to one condition (represented by single dots), while the number of common elements between the different conditions (connected dots) was much smaller. Blue color marks the sRNA sequences that are downregulated in Bur virus samples both in TVCV and TuMV-infected plants while orange color marks the sRNA sequences that are upregulated in Bur virus samples both in TVCV and TuMV-infected plants. Sets with zero elements are not shown.

We focused on the sets that have common sequences between the TVCV and TuMV Bur virus samples, either up- or downregulated to find potential candidates that can regulate the DEGs identified in the RNA-seq analysis. We found 133 upregulated and one downregulated sRNA sequence that was common between the TVCV and TuMV Bur virus samples. Annotation of these sequences revealed that most of the 133 upregulated sequences are 21-nt long and associated with intergenic regions, transposable elements, *TASI*, and *TAS2* genes, some protein-coding genes and pseudogenes, and some miRNA genes (*MIR158B*, *MIR162B*, *MIR390A*, *MIR396A*, *MIR835A*, *MIR841A*, *MIR846A*), although only one mature miRNA sequence was found (ath-miR846-3p), all the other MIR-associated sequences are iso-miRs or siRNAs. It is worth mentioning that 41% of the sequences contain one mismatch

(including the *TAS1C*- and *TAS2*-derived, and miR390, miR396, miR835a, miR841a sequences) which suggests that they are Bur-specific sequences. We checked if the 133 Bur virus upregulated sequences can be found in the Col virus upregulated sequences allowing 2 mismatches and we found that 19 sequences were also present in the Col set, so we removed them. We were curious if the remaining 114 Bur virus upregulated sRNAs can target the 88 Bur virus downregulated genes identified in the RNA-seq analysis, therefore, we performed a target prediction using the psRNATarget server [72]. According to the prediction, 86 out of the 114 sRNAs can potentially regulate 69 out of the 88 genes (Table S5). When we performed a GO term enrichment analysis of the potentially vasiRNA-regulated genes, the regionalization and pattern specification categories remained significantly enriched (Fig S1). That includes a We found that a transposon-derived 21-nt-long siRNA (5'- AUU GCA GAG AUG GAU GUA CAA -3') that could potentially target *FIL* *in trans*. The *FIL*-targeting sequence is identical in the two ecotypes but induced only in the Bur virus sample. One other gene in the GO categories regionalization (GO:0003002) and pattern specification process (GO:0007389) is also potentially regulated by a vasiRNA: *DOT3* (*DEFECTIVELY ORGANIZED TRIBUTARIES 3*, AT5G10250) is targeted by an intergenic siRNA that contains one mismatch compared to the Col reference sequence, but the Col version is not induced in the Col virus sample. The most significantly enriched categories are related to photosynthesis and chloroplast functions. This is consistent with the previous observations that vasiRNAs target genes coding for different members of the photosynthetic apparatus to inhibit the replication of viruses by tuning down the energy production of the plant cell [43]. Apparently, there is an ecotype difference in this response that could contribute to the symptom severity.

The miR396 has been described to target the GRF family of transcription factors that are required for coordination of cell division and differentiation during leaf development in *Arabidopsis* [91]. The GRFs are not among the 88 downregulated genes we found. We



checked their expression pattern in our data but they do not show any consistent pattern. There might be other genes related to leaf development directly or indirectly that could be regulated post-transcriptionally by vasiRNAs.

The one commonly downregulated sequence is the tasiARF that regulates *ARF4* which was found in the 36 Bur virus upregulated genes (Table S4).

## **tasiARF expression anti-correlates with the expression of its target *ARF4***

The finding that tasiARF is the only commonly downregulated sRNA between the TVCV and TuMV Bur virus samples prompted us to look up the expression profile of the tasiARF-related pathway components, namely, the phase-initiating miR390 that targets *TAS3* (the precursor of tasiARF), *AGO7*, the protein product of which binds miR390, *RDR6*, *SGS3*, *DRB4*, and *DCL4*, the protein products of which are involved in the synthesis and dicing of the dsRNA precursor, the tasiARF, and its target *ARF4*, the identified vasiRNA that potentially targets *FIL*, and *FIL* itself that together with *ARF4* are involved in the leaf polarity determination (Fig 10).

### **Fig 10. Expression pattern of the *TAS3* regulatory module components.**

The normalized expression values of the sRNAs and genes involved in the *TAS3* regulatory module are shown in the TVCV and TuMV series. Error bars represent the standard error of the mean of four biological replicates. The actual data points are also shown.

Analyzing Fig 10, we can observe the following: miR390 is virus-induced in both ecotypes, especially by TVCV infection. The *AGO7*, *RDR6*, and *DRB4* mRNA levels did not change significantly, while the *SGS3* was induced only in the Bur virus samples in both virus

series. *DCLA* was induced in the Bur virus sample but the induced level was the same as the levels in the Col mock and virus samples. The tasiARF levels significantly drop in the Bur virus samples both in the TVCV and TuMV series. The expression profile of *ARF4* is anti-correlated with the tasiARF and the *FIL* profile in both virus series. The newly identified vasiRNA-FIL is dramatically induced in the Bur virus samples both the TVCV and TuMV series and anti-correlates with the *FIL* expression. The *ARF4* and *FIL* expression levels were validated by RT-qPCR (Fig 11).

**Fig 11. Validation of the *ARF4* and *FIL* expression patterns with RT-qPCR.**

(A) RT-qPCRs were performed on the samples used for the high-throughput sequencing. Four biological and two technical replicates were measured per sample. The values for every biological replicate were calculated as the mean of the two technical replicates and shown as dots in the figure. The error bars represent the standard error of the mean of four biological replicates. The *PDF2* gene (AT1G13320) was used as an internal reference [76]. (B) Unpaired, two-tailed Student's *t*-tests at 95% confidence intervals were performed to assess if there is a significant difference between the means of the mock and virus-infected samples in the Col and Bur plants either in the TVCV or TuMV series. The resulting *P*-values for every pair tested are indicated in the table. Values less than 0.05 indicate a significant difference. (C) Expression pattern of the unspliced pre-mRNA transcript of *ARF4* (genomic or *gARF4*) in the RNA-seq data. The error bars represent the standard error of the mean of four biological replicates. The individual data points are also shown.

To decide whether the observed elevated level of *ARF4* is due to a transcriptional (i.e., induction by a transcription factor like *FIL*) or post-transcriptional regulation (i.e., by tasiARF), we checked the expression level of the unspliced pre-mRNA transcript of *ARF4* by amending the reference transcriptome with the genomic sequence of the *ARF4* gene

(*gARF4*) and performed an RNA-seq analysis as described before. According to this analysis, the *gARF4* transcript levels were very low in all samples and no significant change was observed that could explain the elevated level of *ARF4* in the Bur virus sample (Fig 11C). This suggests that the observed *ARF4* induction is a result of decreased post-transcriptional regulation by tasiARF.

Taken together, the induction of miR390 and *SGS3* would suggest a higher level of *TAS3* processing and eventually a higher level of tasiARF, but the case is the opposite: the tasiARF production is compromised during virus infection, especially in the Bur ecotype. This observation suggests that the virus blocks one or more key steps of phased siRNA production. This can either be the miR390-mediated cleavage of the *TAS3* ncRNA or the RDR6/*SGS3*-mediated dsRNA precursor synthesis, or the phased cleavage of the dsRNA precursor by *DCL4*. The miR390-mediated cleavage is specific to the *TAS3* pathway, while the other components are shared by other tasiRNA-producing processes.

## **Viral infection induces the production of tasiRNAs except for the tasiARF**

In *Arabidopsis thaliana*, there are eight *TAS* loci of four families and many other phased siRNA-producing loci. Most of the sRNAs emerging from these loci are not AGO-competent and have no known biological role. However, some of them have evolved to target host genes. Only the *TAS3*-derived tasiARF is conserved among the land plants [24–26]. We wondered if the virus also blocks the production of other tasiRNAs or if this phenomenon is specific to the *TAS3* pathway, therefore, we analyzed their expression profiles. Surprisingly, unlike the tasiARF, all the investigated functional tasiRNAs were virus-induced (Fig 12), except for the phosphate starvation-induced *TAS4*-siR81 [39,92] which is scarcely expressed under our circumstances.

**Fig 12. Expression pattern of functional tasiRNAs other than the *TAS3*-derived tasiARF.**

The normalized expression values of the functional siRNAs are shown in the TVCV and TuMV series. We considered only the mature, functional sequence; other sequence variants or non-functional siRNAs from the same TAS locus were not included in the analysis. Error bars represent the standard error of the mean of four biological replicates. The actual data points are also shown.

These results suggest that the *TAS3* regulatory module is blocked at the *TAS3*-specific miR390 cleavage since all the other steps are carried out by shared components including the Bur virus-induced *SGS3* that might be responsible for the enhanced production of tasiRNAs from other loci. This blockage could happen either through ecotype-specific sequestering of miR390 or the AGO7 protein (or some unknown component of the tasiARF biogenesis pathway) by the viral silencing suppressor. Alternatively, the AGO1-loading of tasiARF could be blocked which might lead to the destabilization of this siRNA specimen. The exact mechanism causing depletion of the tasiARF remains to be elucidated.

**QTL mapping suggests that the observed leaf deformation is a result of multigenic interaction**

To support our findings, we performed a limited quantitative trait locus (QTL) mapping experiment using 43 selected recombinant inbred lines (RILs) between Bur and Col ecotypes that were established by INRA [77]. For the experiments, five plants per line were infected either with TVCV-ApH or mock and the leaf deformation was scored by inspection. The genotype data for the markers were obtained from INRA and along with the phenotype data,

were used as input for the mapping software Windows QTL Cartographer [78]. A Single Marker Association test identified several markers on different chromosomes associated with the leaf deformation (Doc S1). The strongest association was observed at the bottom of chromosome 5, where the *ARF4* (AT5G60450) is located (among others). However, the presence of other possible QTLs suggests that the leaf deformations are not caused by the erroneous expression of one gene (*ARF4*), rather, it might be the result of the interaction of multiple genes.

## **Ecotype-specific genetic differences in the tasiARF biogenesis**

### **pathway components**

We hypothesized that the VSRs could interfere with tasiARF biogenesis either by sequestering miR390 or by blocking some functions of AGO7, or some unknown component of the biogenesis pathway in an ecotype-specific manner. One possibility was that the structure of the miR390 sRNA duplex was different in the two ecotypes because of a possibly different sequence of the star strand, resulting in a differential binding by the VSRs. The sequence and structure of the miR390 sRNA duplex were shown to be important for selective binding by AGO7 [29,93] and possibly affect its binding to VSRs as well. We investigated this possibility by predicting the miR390 loci (*MIR390A* on chromosome 2, and *MIR390B* on chromosome 5) in the Col and Bur ecotypes with ShortStack using the appropriate sRNA and genome sequences of the individual ecotypes separately. The analysis revealed no difference in the sequence of miR390 precursors or miRNA duplex structures.

Alternatively, VSRs could bind and block one of the components of the tasiARF biogenesis pathway in an ecotype-specific manner. To investigate this possibility, we compared the amino acid sequences of the proteins involved in the tasiARF biogenesis. We downloaded the protein sequences of AGO1, AGO2, AGO7, ARF4, and FIL from 19

different *Arabidopsis thaliana* ecotypes (including Bur and Col) from the homepage of the 1001 genome project (<http://mtweb.cs.ucl.ac.uk/mus/www/19genomes/index.html>), and aligned the sequence variants with Clustal omega [94]. According to the alignment, there were five amino acid changes in the Bur version of AGO7 protein compared to the Col version (all in the N-terminal part), while there were two and one amino acid changes in ARF4 and AGO2 proteins, respectively. There was no difference in the sequence of AGO1 and FIL proteins in the Col and Bur ecotypes (Fig S2). The Bur version of AGO7 was more different than the versions in the rest of the ecotypes. The two amino acid changes in ARF4 were not conserved (E/K and T/I in Col and Bur, respectively), and again, the Bur variant was rare among the investigated ecotypes. The FIL and AGO1 sequences were remarkably conserved among the investigated 19 ecotypes suggesting a strong selection pressure on them. There was some degree of variability in the AGO2 sequence but the Bur was not the most divergent ecotype in this respect.

Taken together, the best candidate for the ecotype-specific blockage of tasiARF production is AGO7 (or an unknown component of the TAS3 module) but the exact mechanism of blockage remains to be elucidated.

## Discussion

Viral symptom development is regulated by both viral and host factors. On the virus side, viral silencing suppressors were found to be the major symptom determinants. However, only a scarce amount of much less data is available for such symptom determinants on the plant side. During an infection, the virus modulates the molecular environment of the host cell in a way that favors the replication and spread of the virus. Since the viral genome has a limited coding capacity, the virus utilizes host factors for its replication, cell-to-cell movement, virion assembly, and release. Genetic variations in these host factors could affect

the severity of viral symptoms. Such variations occur naturally and determine the host range in which the virus can replicate efficiently. Host factors can also be manipulated artificially which may result in a restricted viral replication or enhanced immune response [95,96].

*Arabidopsis thaliana* is one of the most studied model organisms of plant biology. There are hundreds of ecotypes of this species that differ significantly in their geno- and phenotypes and therefore, are susceptible to viral infection differently. Indeed, in a **genome-wide association study screen** under field conditions, 317 *Arabidopsis* accessions were phenotyped for TuMV viral RNA accumulation and found significant differences among the susceptibility of the different ecotypes [97]. We also screened *Arabidopsis* ecotypes that display more severe symptoms upon virus infection (unpublished results) and we found that the Bur ecotype displays more severe leaf developmental defects than the Col ecotype when infected either with TVCV or TuMV. Since leaf development is regulated by sRNAs at many points [15,16,98,99], we expected that the mRNA and sRNA profiles of the Col and Bur ecotypes would be significantly different when infected with viruses. Namely, we expected that genes and sRNAs that were responsible for the observed severe symptoms only in the Bur ecotype would change significantly (in the Bur virus sample only compared to the Col mock, Col virus, or Bur mock samples). To find such genes and sRNAs, we performed transcriptome and sRNAome analyses. Our analysis revealed that the levels of *FIL* (*YABBY1*, *AFO*) and *ARF4*, two transcription factors that are major abaxial polarity determinants in *Arabidopsis* leaf development [99], and the small RNA specimens **called** tasiARF **and vasiRNA-FIL** that negatively regulates *ARF4* [30] **and *FIL*, respectively**, change significantly in the Bur ecotype when infected either with TVCV or TuMV, while such a change cannot be observed in the Col background. This led us to assume that the corruption of the *TAS3*-mediated regulatory module might be responsible for the observed leaf deformation in the virus-infected Bur plants. According to our analysis, miR390 is virus-induced in both ecotypes, especially upon TVCV infection but the tasiARF production is blocked only in the

Bur virus sample. We observed that the production of other tasiRNAs is not blocked, rather, they are induced upon virus infection, especially in the Bur virus samples. This suggested that both viruses likely specifically block the *TAS3* pathway at the phase initiating AGO7-mediated cleavage. The Bur-specific depletion of tasiARF might lead to an ectopic expression of *ARF4* in the adaxial part of the leaf which could result in a misregulated auxin-dependent cell expansion in this domain, manifested as an upward curling leaf and wavy leaf margin.

*ARF4* is called a repressive, class B ARF [100–102] because it lacks the glutamine-rich transactivator domain that other activating ARFs have. Repressive ARFs might act by forming a dysfunctional heterodimer with other ARFs, and can form repressive chromatin in the promoter of the regulated gene [103]. In a previous study, *ARF8* was shown to be the major symptom determinant in transgenic plants expressing three different viral silencing suppressors [104]. Unlike *ARF4*, *ARF8* is an activating ARF and regulated post-transcriptionally by miR167 rather than tasiARF. In the *Arabidopsis* plants (Col background) overexpressing the TuMV silencing suppressor HCPro, *ARF4* was found to be upregulated that coincided with upward curling leaves, just like in our experiment. However, in the *arf4* mutant, the symptoms do not disappear, unlike in the *arf8* mutant. This might be because of the different background (the *arf4* mutant is in Col) that suggests a multigenic interaction that allows the symptoms to develop only in the Bur background.

In a different work that utilized a dexamethasone-inducible *FIL:FIL-GR* construct, researchers showed that *FIL* positively regulates *ARF4* and *KAN1* but not *ARF3* [85]. In our viral-infected samples, we observed the downregulation of *FIL* and upregulation of *ARF4* which seemingly contradicts the above-mentioned finding. However, the drop of the tasiARF level in our Bur virus sample in both the TVCV and TuMV series anti-correlates well with the *ARF4* level which might underline the importance of the post-transcriptional regulation of *ARF4*. It is also important to note that in the above-mentioned study, *FIL* was selectively



induced in the abaxial part of the leaf where the *ARF4* is normally confined by the tasiARF gradient, while in our study, the *ARF4* expression was probably increased in the adaxial part due to the absence of inhibitory tasiARF. Also, the newly identified virus-induced vasiRNA-FIL might interfere with the delicate regulation of leaf polarity in an ecotype-specific manner.

Understanding the host-pathogen interaction network to diminish viral symptom development is one of the central goals of virological research. In this study, we have shown that two viruses having very different replication strategies, (TVCV and TuMV which are genuine representatives of *Tobamovirus* and *Potyvirus* families, respectively) cause very similar upward curling leaf and wavy leaf margin symptoms on the Bur but not on Col ecotype plants. The molecular basis for the ecotype difference therefore likely lies within the presence or absence of specific host factors that react in a very similar way to the infection of by the two pathogens (in the Bur ecotype). By studying the transcriptomes and sRNAomes of the two sets of infections we hypothesize that the basis of the symptom development in Bur plants may be the enhanced impairment of the *TAS3* pathway. Both viruses possess a VSR that binds sRNAs, constituting a common target of action acting player: (i.e. for example, miR390 duplex binding by the two VSRs would impair *TAS3*-derived tasiARF biogenesis. An ecotype-specific structural difference in the miR390 duplex could result in selective binding by the VSRs. Still in such a case, a host cellular status/characteristic should enable the ecotype specific symptom development through VSR action. However, we could not find evidence for a different duplex structure of miR390a or miR390b in the two ecotypes. The RISC-loading efficiency of miR390 is very low and the overexpression of miR390 does not improve it, because the bottleneck is the availability of AGO7 [11]. It is possible that the miR390 loading into AGO7 is differentially affected by the presence of the virus either because the VSR directly manipulates the loading process or indirectly by relocating the AGO7 within the cell to a compartment that lacks miR390. Since some VSRs can bind ARGONAUTE proteins to block their antiviral action, we compared the amino acid

sequences of the AGO proteins that could be involved in the production of tasiARF. There was no difference in the AGO1 sequence but AGO2 and especially AGO7 were more variable between the investigated ecotypes. ~~A more plausible~~ We think that the best explanation ~~would be is~~ that the VSR binds and blocks ~~some functions of~~ the AGO7 protein selectively in the Bur ecotype but that would require a parallel evolution of the VSRs of different viruses to target the same AGO7 protein. Such parallel evolution of viral factors to target the same host factor is not unprecedented [105]. A study investigating the evolution of plant VSRs argues that plant VSRs are subject to episodic positive selection due to frequent jumps between different hosts rather than a one-to-one co-evolution with a single host [106]. This is consistent with our findings since both viruses we worked with have a broad host range (the turnip vein-clearing virus isolate we used, TVCV-ApH was isolated from *Alliaria petiolata*, a.k.a. garlic mustard [45]). It is unclear why the virus ~~favors~~ would favor blocking AGO7, although it was described that AGO7 has some antiviral function [107,108]. Another interesting possibility is that since the virus replication is often associated with an endomembrane system and the AGO7-mediated tasiRNA production is also endomembrane-associated [109], therefore, the viral replication could interfere with the AGO7-dependent tasiARF production. However, this interference should not be a passive process since it should happen only in the Bur ecotype.

Our QTL mapping results suggest that a multigenic interaction is behind the observed leaf deformations. ~~, although the disrupted abaxial-adaxial gradient of tasiARF could be a major contributor.~~ Many of the genes that were selectively up- and downregulated in the Bur virus samples could be related to leaf development: many of the upregulated genes regulate cell wall structure while many downregulated genes are involved in auxin biosynthesis (i.e. *TAR2*) or signaling (i.e. *IAA29*), and could be directly involved in ARF4-mediated auxin signaling. It is worth mentioning that there are numerous transcription factors in both sets (see Table S3) that could form an intricate regulatory network with or without ARF4 the net

effect of which would lead to the observed phenotype. It is difficult to tell which players are the most upstream in the signaling chain that leads to the observed ecotype-specific leaf deformation, although we believe that the disrupted abaxial-adaxial gradient of tasiARF could be a major contributor. To reveal the exact nature of the selective blockage of tasiARF production in the virus-infected Bur ecotype, further research is needed in the future.

## Supporting information

**Fig S1. GO term enrichment analysis of the downregulated, potential vasiRNA target genes.**

**Fig S2. Amino acid alignment of ARF4, FIL, AGO1, AGO2, and AGO7 in 19 different ecotypes.**

**Table S1. Sequencing statistics of the RNA-seq and sRNA-seq data.**

**Table S2. Normalized expression values of all *Arabidopsis* genes; Wald test results and normalized expression values of differentially expressed genes.**

**Table S3. Commonly regulated genes between the TVCV and TuMV series and their GO term enrichment analysis.** Content of the two sets (Bur virus up- and downregulated genes in both virus series), and the full result of their GO term enrichment analysis.

**Table S4. Wald test results and normalized expression values of differentially expressed sRNAs; commonly regulated sRNAs between the TVCV and TuMV series.**

**Table S5. Target prediction of the Bur virus upregulated sRNAs among the Bur virus downregulated genes.** Result of the target prediction performed using the psRNATarget server. Only the predicted Cleavage hits are shown. The sRNAs and the target genes are annotated.

**Data S1. Normalized expression values of all sRNAs in the samples.** Annotated expression table of sRNA sequences having a mean expression > 1. The table contains

102195 unique, 20–25-nt-long sequences ordered by mean expression value (highly expressed first).

**Doc S1. Result of the QTL mapping.** The result of a single marker analysis performed by WinQTL Cartographer.

## Acknowledgments

We would like to thank Tibor Csorba for his valuable suggestions to improve the manuscript.

## References

1. Hull R. Chapter 4 - Symptoms and Host Range. In: Hull R, editor. *Plant Virology* (Fifth Edition). Boston: Academic Press; 2014. pp. 145–198. Available: <http://www.sciencedirect.com/science/article/pii/B9780123848710000042>
2. Carrington JC, Kasschau KD, Mahajan SK, Schaad MC. Cell-to-Cell and Long-Distance Transport of Viruses in Plants. *Plant Cell Online*. 1996;8: 1669–1681. doi:10.1105/tpc.8.10.1669
3. Ding S-W, Voinnet O. Antiviral immunity directed by small RNAs. *Cell*. 2007;130: 413–426. doi:10.1016/j.cell.2007.07.039
4. 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. doi:10.1128/JVI.79.12.7812-7818.2005
5. 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. doi:10.1126/science.1128214
6. Voinnet O. Induction and suppression of RNA silencing: insights from viral

- infections. *Nat Rev Genet.* 2005;6: 206–220. doi:10.1038/nrg1555
7. Burgyán J, Havelda Z. Viral suppressors of RNA silencing. *Trends Plant Sci.* 2011;16: 265–272. doi:10.1016/j.tplants.2011.02.010
  8. Csorba T, Kontra L, Burgyán J. Viral silencing suppressors: Tools forged to fine-tune host-pathogen coexistence. *Virology.* 2015;479–480: 85–103. doi:10.1016/j.virol.2015.02.028
  9. Mi S, Cai T, Hu Y, Chen Y, Hodges E, Ni F, et al. Sorting of small RNAs into *Arabidopsis* Argonaute complexes is directed by the 5' terminal nucleotide. *Cell.* 2008;133: 116–127. doi:10.1016/j.cell.2008.02.034
  10. Takeda A, Iwasaki S, Watanabe T, Utsumi M, Watanabe Y. The Mechanism Selecting the Guide Strand from Small RNA Duplexes is Different Among Argonaute Proteins. *Plant Cell Physiol.* 2008;49: 493–500. doi:10.1093/pcp/pcn043
  11. Dalmadi Á, Gyula P, Bálint J, Szittyá G, Havelda Z. AGO-unbound cytosolic pool of mature miRNAs in plant cells reveals a novel regulatory step at AGO1 loading. *Nucleic Acids Res.* 2019;47: 9803–9817. doi:10.1093/nar/gkz690
  12. Satish D, Mukherjee SK, Gupta D. The landscape of microRNAs in plant viral infections. *Plant Gene.* 2021; 100293. doi:10.1016/j.plgene.2021.100293
  13. Zhang B, Li W, Zhang J, Wang L, Wu J. Roles of Small RNAs in Virus-Plant Interactions. *Viruses.* 2019;11: 827. doi:10.3390/v11090827
  14. Bazzini AA, Hopp HE, Beachy RN, Asurmendi S. Infection and coaccumulation of tobacco mosaic virus proteins alter microRNA levels, correlating with symptom and plant development. *Proc Natl Acad Sci.* 2007;104: 12157–12162. doi:10.1073/pnas.0705114104
  15. Yang T, Wang Y, Teotia S, Zhang Z, Tang G. The Making of Leaves: How Small RNA Networks Modulate Leaf Development. *Front Plant Sci.* 2018;9. doi:10.3389/fpls.2018.00824

16. Pulido A, Laufs P. Co-ordination of developmental processes by small RNAs during leaf development. *J Exp Bot.* 2010;61: 1277–1291. doi:10.1093/jxb/erp397
17. Marín-González E, Suárez-López P. “And yet it moves”: Cell-to-cell and long-distance signaling by plant microRNAs. *Plant Sci.* 2012;196: 18–30. doi:10.1016/j.plantsci.2012.07.009
18. Chitwood DH, Nogueira FTS, Howell MD, Montgomery TA, Carrington JC, Timmermans MCP. Pattern formation via small RNA mobility. *Genes Dev.* 2009;23: 549–554. doi:10.1101/gad.1770009
19. Xia R, Xu J, Meyers BC. The Emergence, Evolution, and Diversification of the miR390-TAS3-ARF Pathway in Land Plants. *Plant Cell.* 2017;29: 1232–1247. doi:10.1105/tpc.17.00185
20. Fahlgren N, Montgomery TA, Howell MD, Allen E, Dvorak SK, Alexander AL, et al. Regulation of AUXIN RESPONSE FACTOR3 by TAS3 ta-siRNA Affects Developmental Timing and Patterning in Arabidopsis. *Curr Biol.* 2006;16: 939–944. doi:10.1016/j.cub.2006.03.065
21. Garcia D, Collier SA, Byrne ME, Martienssen RA. Specification of Leaf Polarity in Arabidopsis via the trans-Acting siRNA Pathway. *Curr Biol.* 2006;16: 933–938. doi:10.1016/j.cub.2006.03.064
22. Marin E, Jouannet V, Herz A, Lokerse AS, Weijers D, Vaucheret H, et al. miR390, Arabidopsis TAS3 tasiRNAs, and Their AUXIN RESPONSE FACTOR Targets Define an Autoregulatory Network Quantitatively Regulating Lateral Root Growth. *Plant Cell.* 2010;22: 1104–1117. doi:10.1105/tpc.109.072553
23. Hunter C, Willmann MR, Wu G, Yoshikawa M, Gutiérrez-Nava M de la L, Poethig SR. Trans-acting siRNA-mediated repression of ETTIN and ARF4 regulates heteroblasty in Arabidopsis. *Development.* 2006;133: 2973–2981. doi:10.1242/dev.02491

24. Axtell MJ, Jan C, Rajagopalan R, Bartel DP. A Two-Hit Trigger for siRNA Biogenesis in Plants. *Cell*. 2006;127: 565–577. doi:10.1016/j.cell.2006.09.032
25. Nagasaki H, Itoh J, Hayashi K, Hibara K, Satoh-Nagasawa N, Nosaka M, et al. The small interfering RNA production pathway is required for shoot meristem initiation in rice. *Proc Natl Acad Sci*. 2007;104: 14867–14871. doi:10.1073/pnas.0704339104
26. Douglas RN, Wiley D, Sarkar A, Springer N, Timmermans MCP, Scanlon MJ. ragged seedling2 Encodes an ARGONAUTE7-Like Protein Required for Mediolateral Expansion, but Not Dorsiventrality, of Maize Leaves. *Plant Cell*. 2010;22: 1441–1451. doi:10.1105/tpc.109.071613
27. Adenot X, Elmayan T, Laressergues D, Boutet S, Bouché N, Gascioli V, et al. DRB4-Dependent TAS3 trans-Acting siRNAs Control Leaf Morphology through AGO7. *Curr Biol*. 2006;16: 927–932. doi:10.1016/j.cub.2006.03.035
28. Pekker I, Alvarez JP, Eshed Y. Auxin Response Factors Mediate Arabidopsis Organ Asymmetry via Modulation of KANADI Activity. *Plant Cell*. 2005;17: 2899–2910. doi:10.1105/tpc.105.034876
29. Montgomery TA, Howell MD, Cuperus JT, Li D, Hansen JE, Alexander AL, et al. Specificity of ARGONAUTE7-miR390 interaction and dual functionality in TAS3 trans-acting siRNA formation. *Cell*. 2008;133: 128–141. doi:10.1016/j.cell.2008.02.033
30. Allen E, Xie Z, Gustafson AM, Carrington JC. microRNA-Directed Phasing during Trans-Acting siRNA Biogenesis in Plants. *Cell*. 2005;121: 207–221. doi:10.1016/j.cell.2005.04.004
31. Schwab R, Maizel A, Ruiz-Ferrer V, Garcia D, Bayer M, Crespi M, et al. Endogenous TasiRNAs Mediate Non-Cell Autonomous Effects on Gene Regulation in *Arabidopsis thaliana*. *PLOS ONE*. 2009;4: e5980. doi:10.1371/journal.pone.0005980

32. Meyers BC, Kaushik S, Nandety RS. Evolving disease resistance genes. *Curr Opin Plant Biol.* 2005;8: 129–134. doi:10.1016/j.pbi.2005.01.002
33. Leonetti P, Stuttmann J, Pantaleo V. Regulation of plant antiviral defense genes via host RNA-silencing mechanisms. *Virology*. 2021;18: 194. doi:10.1186/s12985-021-01664-3
34. Shivaprasad PV, Chen H-M, Patel K, Bond DM, Santos BACM, Baulcombe DC. A microRNA superfamily regulates nucleotide binding site-leucine-rich repeats and other mRNAs. *Plant Cell.* 2012;24: 859–874. doi:10.1105/tpc.111.095380
35. Zhang Y, Xia R, Kuang H, Meyers BC. The Diversification of Plant NBS-LRR Defense Genes Directs the Evolution of MicroRNAs That Target Them. *Mol Biol Evol.* 2016;33: 2692–2705. doi:10.1093/molbev/msw154
36. Sós- Hegedűs A, Domonkos Á, Tóth T, Gyula P, Kaló P, Szittyá G. Suppression of NB-LRR genes by miRNAs promotes nitrogen-fixing nodule development in *Medicago truncatula*. *Plant Cell Environ.* 2020;43: 1117–1129. doi:10.1111/pce.13698
37. Zhai J, Jeong D-H, De Paoli E, Park S, Rosen BD, Li Y, et al. 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. doi:10.1101/gad.177527.111
38. Felippes FF, Weigel D. Triggering the formation of tasiRNAs in *Arabidopsis thaliana*: the role of microRNA miR173. *EMBO Rep.* 2009;10: 264–270. doi:10.1038/embor.2008.247
39. Rajagopalan R, Vaucheret H, Trejo J, Bartel DP. A diverse and evolutionarily fluid set of microRNAs in *Arabidopsis thaliana*. *Genes Dev.* 2006;20: 3407–3425. doi:10.1101/gad.1476406
40. Li S, Liu J, Liu Z, Li X, Wu F, He Y. HEAT-INDUCED TAS1 TARGET1 Mediates Thermotolerance via HEAT STRESS TRANSCRIPTION FACTOR A1a-Directed



- Pathways in Arabidopsis. *Plant Cell*. 2014;26: 1764–1780.  
doi:10.1105/tpc.114.124883
41. Si-Ammour A, Windels D, Arn-Boulidoires E, Kutter C, Ailhas J, Meins F, et al. miR393 and Secondary siRNAs Regulate Expression of the TIR1/AFB2 Auxin Receptor Clade and Auxin-Related Development of Arabidopsis Leaves. *Plant Physiol*. 2011;157: 683–691. doi:10.1104/pp.111.180083
42. Cao M, Du P, Wang X, Yu Y-Q, Qiu Y-H, Li W, et al. Virus infection triggers widespread silencing of host genes by a distinct class of endogenous siRNAs in Arabidopsis. *Proc Natl Acad Sci*. 2014;111: 14613–14618.  
doi:10.1073/pnas.1407131111
43. Leonetti P, Ghasemzadeh A, Consiglio A, Gursinsky T, Behrens S, Pantaleo V. Endogenous activated small interfering RNAs in virus- infected Brassicaceae crops show a common host gene- silencing pattern affecting photosynthesis and stress response. *New Phytol*. 2021;229: 1650–1664. doi:10.1111/nph.16932
44. Pitzalis N, Amari K, Graindorge S, Pflieger D, Donaire L, Wassenegger M, et al. Turnip mosaic virus in oilseed rape activates networks of sRNA-mediated interactions between viral and host genomes. *Commun Biol*. 2020;3: 1–16.  
doi:10.1038/s42003-020-01425-y
45. Tóth T, Gyula P, Salamon P, Kis S, Sós-Hegedűs A, Szittyá 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.  
doi:10.1371/journal.pone.0224398
46. Lartey RT. Occurrence of a vein-clearing tobamovirus in turnip. *Plant Dis*. 1993;77: 21. doi:10.1094/PD-77-0021
47. Melcher U. Turnip vein-clearing virus, from pathogen to host expression profile. *Mol Plant Pathol*. 2003;4: 133–140. doi:10.1046/j.1364-3703.2003.00159.x

48. Csorba T, Bovi A, Dalmay T, Burgyán J. The p122 subunit of Tobacco mosaic virus replicase is a potent silencing suppressor and compromises both small interfering RNA- and microRNA-mediated pathways. *J Virol.* 2007;81: 11768–11780.  
doi:10.1128/JVI.01230-07
49. Ding XS, Liu J, Cheng N-H, Folimonov A, Hou Y-M, Bao Y, et al. The Tobacco mosaic virus 126-kDa protein associated with virus replication and movement suppresses RNA silencing. *Mol Plant-Microbe Interact MPMI.* 2004;17: 583–592.  
doi:10.1094/MPMI.2004.17.6.583
50. Vogler H, Akbergenov R, Shivaprasad PV, Dang V, Fasler M, Kwon M-O, et al. Modification of small RNAs associated with suppression of RNA silencing by tobamovirus replicase protein. *J Virol.* 2007;81: 10379–10388.  
doi:10.1128/JVI.00727-07
51. Shattuck VI. *The Biology, Epidemiology, and Control of Turnip Mosaic Virus.* Horticultural Reviews. John Wiley & Sons, Ltd; 1992. pp. 199–238.  
doi:10.1002/9780470650523.ch4
52. Riechmann JL, Laín S, García JA 1992. Highlights and prospects of potyvirus molecular biology. *J Gen Virol.* 1992;73: 1–16. doi:10.1099/0022-1317-73-1-1
53. Urcuqui-Inchima S, Haenni A-L, Bernardi F. Potyvirus proteins: a wealth of functions. *Virus Res.* 2001;74: 157–175. doi:10.1016/S0168-1702(01)00220-9
54. Chung BY-W, Miller WA, Atkins JF, Firth AE. An overlapping essential gene in the Potyviridae. *Proc Natl Acad Sci.* 2008;105: 5897–5902.  
doi:10.1073/pnas.0800468105
55. Wei T, Zhang C, Hong J, Xiong R, Kasschau KD, Zhou X, et al. Formation of Complexes at Plasmodesmata for Potyvirus Intercellular Movement Is Mediated by the Viral Protein P3N-PIPO. *PLOS Pathog.* 2010;6: e1000962.  
doi:10.1371/journal.ppat.1000962

56. Anandalakshmi R, Pruss GJ, Ge X, Marathe R, Mallory AC, Smith TH, et al. A viral suppressor of gene silencing in plants. *Proc Natl Acad Sci.* 1998;95: 13079–13084. doi:10.1073/pnas.95.22.13079
57. Garcia-Ruiz H, Carbonell A, Hoyer JS, Fahlgren N, Gilbert KB, Takeda A, et al. Roles and Programming of Arabidopsis ARGONAUTE Proteins during Turnip Mosaic Virus Infection. *PLoS Pathog.* 2015;11: e1004755. doi:10.1371/journal.ppat.1004755
58. Lakatos L, Csorba T, Pantaleo V, Chapman EJ, Carrington JC, Liu Y-P, et al. Small RNA binding is a common strategy to suppress RNA silencing by several viral suppressors. *EMBO J.* 2006;25: 2768–2780. doi:10.1038/sj.emboj.7601164
59. Szittyá G, Salamon P, Burgyán J. The complete nucleotide sequence and synthesis of infectious RNA of genomic and defective interfering RNAs of TBSV-P. *Virus Res.* 2000;69: 131–136. doi:10.1016/S0168-1702(00)00178-7
60. Andrews S. FastQC: a quality control tool for high throughput sequence data. In: Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data [Internet]. 2010. Available: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
61. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal.* 2011;17: 10–12. doi:10.14806/ej.17.1.200
62. Zhang R, Calixto CPG, Marquez Y, Venhuizen P, Tzioutziou NA, Guo W, et al. A high quality Arabidopsis transcriptome for accurate transcript-level analysis of alternative splicing. *Nucleic Acids Res.* 2017;45: 5061–5073. doi:10.1093/nar/gkx267
63. Pimentel H, Bray NL, Puente S, Melsted P, Pachter L. Differential analysis of RNA-seq incorporating quantification uncertainty. *Nat Methods.* 2017;14: 687–690. doi:10.1038/nmeth.4324

64. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Stat Soc Ser B Methodol.* 1995;57: 289–300.
65. Wickham H. *Ggplot2: elegant graphics for data analysis.* New York: Springer; 2009.
66. Tang Y, Horikoshi M, Li W. *ggfortify: unified interface to visualize statistical results of popular R packages.* *R J.* 2016;8: 474–485.
67. Conway JR, Lex A, Gehlenborg N. *UpSetR: an R package for the visualization of intersecting sets and their properties.* *Bioinformatics.* 2017;33: 2938–2940.  
doi:10.1093/bioinformatics/btx364
68. Tian F, Yang D-C, Meng Y-Q, Jin J, Gao G. *PlantRegMap: charting functional regulatory maps in plants.* *Nucleic Acids Res.* 2020;48: D1104–D1113.  
doi:10.1093/nar/gkz1020
69. Axtell MJ. *ShortStack: Comprehensive annotation and quantification of small RNA genes.* *RNA.* 2013 [cited 29 Mar 2016]. doi:10.1261/rna.035279.112
70. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. *The Sequence Alignment/Map format and SAMtools.* *Bioinformatics.* 2009;25: 2078–2079.  
doi:10.1093/bioinformatics/btp352
71. Love MI, Huber W, Anders S. *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.* *Genome Biol.* 2014;15: 550. doi:10.1186/s13059-014-0550-8
72. Dai X, Zhao PX. *psRNATarget: a plant small RNA target analysis server.* *Nucleic Acids Res.* 2011;39: W155–W159. doi:10.1093/nar/gkr319
73. Rice P, Longden I, Bleasby A. *EMBOSS: The European Molecular Biology Open Software Suite.* *Trends Genet.* 2000;16: 276–277. doi:10.1016/S0168-9525(00)02024-2
74. Marshall OJ. *PerlPrimer: cross-platform, graphical primer design for standard,*

- bisulphite and real-time PCR. *Bioinformatics*. 2004;20: 2471–2472.  
doi:10.1093/bioinformatics/bth254
75. Czechowski T, Bari RP, Stitt M, Scheible W-R, Udvardi MK. Real-time RT-PCR profiling of over 1400 *Arabidopsis* transcription factors: unprecedented sensitivity reveals novel root- and shoot-specific genes. *Plant J*. 2004;38: 366–379.  
doi:10.1111/j.1365-313X.2004.02051.x
76. Lilly ST, Drummond RSM, Pearson MN, MacDiarmid RM. Identification and Validation of Reference Genes for Normalization of Transcripts from Virus-Infected *Arabidopsis thaliana*. *Mol Plant Microbe Interact*. 2011;24: 294–304.  
doi:10.1094/MPMI-10-10-0236
77. Simon M, Loudet O, Durand S, Bérard A, Brunel D, Sennesal F-X, et al. Quantitative Trait Loci Mapping in Five New Large Recombinant Inbred Line Populations of *Arabidopsis thaliana* Genotyped With Consensus Single-Nucleotide Polymorphism Markers. *Genetics*. 2008;178: 2253–2264.  
doi:10.1534/genetics.107.083899
78. Wang S, Basten CJ, Zeng Z-B. Windows QTL Cartographer 2.5. Department of Statistics, North Carolina State University, Raleigh, NC.; 2012. Available: <http://statgen.ncsu.edu/qtlcart/WQTLCart.htm>
79. Whitham SA, Quan S, Chang H-S, Cooper B, Estes B, Zhu T, et al. Diverse RNA viruses elicit the expression of common sets of genes in susceptible *Arabidopsis thaliana* plants. *Plant J*. 2003;33: 271–283. doi:10.1046/j.1365-313X.2003.01625.x
80. Whitham SA, Yang C, Goodin MM. Global Impact: Elucidating Plant Responses to Viral Infection. *Mol Plant Microbe Interact*. 2006;19: 1207–1215.  
doi:10.1094/MPMI-19-1207
81. Bengyella L, Waikhom SD, Allie F, Rey C. Virus tolerance and recovery from viral induced-symptoms in plants are associated with transcriptome reprogramming. *Plant*

- Mol Biol. 2015;89: 243–252. doi:10.1007/s11103-015-0362-6
82. Postnikova OA, Nemchinov LG. Comparative analysis of microarray data in *Arabidopsis* transcriptome during compatible interactions with plant viruses. *Virology*. 2012;9: 101. doi:10.1186/1743-422X-9-101
83. Medzihradszky A, Gyula P, Sós- Hegedűs A, Szittyá 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: 1748–1758. doi:10.1111/mpp.12875
84. Siegfried KR, Eshed Y, Baum SF, Otsuga D, Drews GN, Bowman JL. Members of the YABBY gene family specify abaxial cell fate in *Arabidopsis*. *Development*. 1999;126: 4117–4128.
85. Bonaccorso O, Lee JE, Puah L, Scutt CP, Golz JF. FILAMENTOUS FLOWER controls lateral organ development by acting as both an activator and a repressor. *BMC Plant Biology*. 2012;12: 176. doi:10.1186/1471-2229-12-176
86. Hu Q, Hollunder J, Niehl A, Kørner CJ, Gereige D, Windels D, et al. Specific Impact of Tobamovirus Infection on the *Arabidopsis* Small RNA Profile. *PLoS ONE*. 2011;6: e19549. doi:10.1371/journal.pone.0019549
87. Tatineni S, Riethoven J-JM, Graybosch RA, French R, Mitra A. Dynamics of Small RNA Profiles of Virus and Host Origin in Wheat Cultivars Synergistically Infected by Wheat Streak Mosaic Virus and Triticum Mosaic Virus: Virus Infection Caused a Drastic Shift in the Endogenous Small RNA Profile. *PLOS ONE*. 2014;9: e111577. doi:10.1371/journal.pone.0111577
88. Blevins T, Rajeswaran R, Shivaprasad PV, Beknazariants D, Si-Ammour A, Park H-S, et al. Four plant Dicers mediate viral small RNA biogenesis and DNA virus induced silencing. *Nucleic Acids Research*. 2006;34: 6233–6246. doi:10.1093/nar/gkl886
89. Bouché N, Laressergues D, Gascioli V, Vaucheret H. An antagonistic function for

- Arabidopsis DCL2 in development and a new function for DCL4 in generating viral siRNAs. *EMBO J.* 2006;25: 3347–3356. doi:10.1038/sj.emboj.7601217
90. Wang X-B, Wu Q, Ito T, Cillo F, Li W-X, Chen X, et al. RNAi-mediated viral immunity requires amplification of virus-derived siRNAs in *Arabidopsis thaliana*. *Proc Natl Acad Sci.* 2010;107: 484–489. doi:10.1073/pnas.0904086107
91. Wang L, Gu X, Xu D, Wang W, Wang H, Zeng M, et al. miR396-targeted AtGRF transcription factors are required for coordination of cell division and differentiation during leaf development in *Arabidopsis*. *J Exp Bot.* 2011;62: 761–773. doi:10.1093/jxb/erq307
92. Howell MD, Fahlgren N, Chapman EJ, Cumbie JS, Sullivan CM, Givan SA, et al. Genome-Wide Analysis of the RNA-DEPENDENT RNA POLYMERASE6/DICER-LIKE4 Pathway in *Arabidopsis* Reveals Dependency on miRNA- and tasiRNA-Directed Targeting. *Plant Cell.* 2007;19: 926–942. doi:10.1105/tpc.107.050062
93. Endo Y, Iwakawa H, Tomari Y. *Arabidopsis ARGONAUTE7* selects miR390 through multiple checkpoints during RISC assembly. *EMBO Rep.* 2013;14: 652–658. doi:10.1038/embor.2013.73
94. Madeira F, Pearce M, Tivey ARN, Basutkar P, Lee J, Edbali O, et al. Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Res.* 2022; gkac240. doi:10.1093/nar/gkac240
95. Zaidi SS-A, Tashkandi M, Mansoor S, Mahfouz MM. Engineering Plant Immunity: Using CRISPR/Cas9 to Generate Virus Resistance. *Front Plant Sci.* 2016;7: 1673. doi:10.3389/fpls.2016.01673
96. Borrelli VMG, Brambilla V, Rogowsky P, Marocco A, Lanubile A. The Enhancement of Plant Disease Resistance Using CRISPR/Cas9 Technology. *Front Plant Sci.* 2018;9: 1245. doi:10.3389/fpls.2018.01245
97. Rubio B, Cosson P, Caballero M, Revers F, Bergelson J, Roux F, et al. Genome-wide

- association study reveals new loci involved in *Arabidopsis thaliana* and Turnip mosaic virus (TuMV) interactions in the field. *New Phytol.* 2019;221: 2026–2038. doi:10.1111/nph.15507
98. D’Ario M, Griffiths-Jones S, Kim M. Small RNAs: Big Impact on Plant Development. *Trends Plant Sci.* 2017;22: 1056–1068. doi:10.1016/j.tplants.2017.09.009
  99. Yamaguchi T, Nukazuka A, Tsukaya H. Leaf adaxial–abaxial polarity specification and lamina outgrowth: evolution and development. *Plant Cell Physiol.* 2012;53: 1180–1194. doi:10.1093/pcp/pcs074
  100. Ulmasov T, Hagen G, Guilfoyle TJ. Activation and repression of transcription by auxin-response factors. *Proc Natl Acad Sci.* 1999;96: 5844–5849. doi:10.1073/pnas.96.10.5844
  101. Ulmasov T, Hagen G, Guilfoyle TJ. Dimerization and DNA binding of auxin response factors. *Plant J.* 1999;19: 309–319. doi:https://doi.org/10.1046/j.1365-313X.1999.00538.x
  102. Weijers D, Wagner D. Transcriptional Responses to the Auxin Hormone. *Annu Rev Plant Biol.* 2016;67: 539–574. doi:10.1146/annurev-arplant-043015-112122
  103. Chung Y, Zhu Y, Wu M-F, Simonini S, Kuhn A, Armenta-Medina A, et al. Auxin Response Factors promote organogenesis by chromatin-mediated repression of the pluripotency gene SHOOTMERISTEMLESS. *Nat Commun.* 2019;10: 886. doi:10.1038/s41467-019-08861-3
  104. Jay F, Wang Y, Yu A, Taconnat L, Pelletier S, Colot V, et al. Misregulation of AUXIN RESPONSE FACTOR 8 Underlies the Developmental Abnormalities Caused by Three Distinct Viral Silencing Suppressors in *Arabidopsis*. *PLOS Pathog.* 2011;7: e1002035. doi:10.1371/journal.ppat.1002035
  105. Zhang H, Li L, He Y, Qin Q, Chen C, Wei Z, et al. Distinct modes of manipulation



- of rice auxin response factor OsARF17 by different plant RNA viruses for infection.  
Proc Natl Acad Sci. 2020;117: 9112–9121. doi:10.1073/pnas.1918254117
106. Murray GGR, Kosakovsky Pond SL, Obbard DJ. Suppressors of RNAi from plant viruses are subject to episodic positive selection. Proc R Soc B Biol Sci. 2013;280: 20130965. doi:10.1098/rspb.2013.0965
107. Qu F, Ye X, Morris TJ. Arabidopsis DRB4, AGO1, AGO7, and RDR6 participate in a DCL4-initiated antiviral RNA silencing pathway negatively regulated by DCL1. Proc Natl Acad Sci. 2008;105: 14732–14737. doi:10.1073/pnas.0805760105
108. Carbonell A, Carrington JC. Antiviral roles of plant ARGONAUTES. Curr Opin Plant Biol. 2015;27: 111–117. doi:10.1016/j.pbi.2015.06.013
109. Jouannet V, Moreno AB, Elmayer T, Vaucheret H, Crespi MD, Maizel A. Cytoplasmic Arabidopsis AGO7 accumulates in membrane-associated siRNA bodies and is required for ta-siRNA biogenesis. EMBO J. 2012;31: 1704–1713. doi:10.1038/emboj.2012.20