



HAL
open science

Small RNA-Omics for Virome Reconstruction and Antiviral Defense Characterization in Mixed Infections of Cultivated Solanum Plants

Silvia Turco, Victor Golyaev, Jonathan Seguin, Céline Gilli, Laurent Farinelli, Thomas Boller, Olivier Schumpp, Mikhail Poogin

► **To cite this version:**

Silvia Turco, Victor Golyaev, Jonathan Seguin, Céline Gilli, Laurent Farinelli, et al.. Small RNA-Omics for Virome Reconstruction and Antiviral Defense Characterization in Mixed Infections of Cultivated Solanum Plants. *Molecular Plant-Microbe Interactions*, 2018, 31 (7), pp.707-723. 10.1094/MPMI-12-17-0301-R. hal-02628890

HAL Id: hal-02628890

<https://hal.inrae.fr/hal-02628890>

Submitted on 27 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

Small RNA-Omics for Virome Reconstruction and Antiviral Defense Characterization in Mixed Infections of Cultivated *Solanum* Plants

Silvia Turco,¹ Victor Golyaev,¹ Jonathan Seguin,¹ Céline Gilli,² Laurent Farinelli,³ Thomas Boller,¹ Olivier Schumpp,^{4,†} and Mikhail M. Pooggin^{1,5,†}

¹Department of Environmental Sciences, Botany, Zurich-Basel Plant Science Center, University of Basel, Hebelstrasse 1, 4056 Basel, Switzerland; ²Agroscope, 1964 Conthey, Switzerland; ³Fasteris AG, 1228 Plan-les-Ouates, Switzerland; ⁴Agroscope, 1260 Nyon, Switzerland; and ⁵INRA, UMR BGPI, 34398 Montpellier, France

Accepted 7 February 2018.

In plants, RNA silencing–based antiviral defense generates viral small RNAs (sRNAs) faithfully representing the viral genomes. We employed sRNA sequencing and bioinformatics (sRNA-omics) to characterize antiviral defense and to reconstruct the full genomic sequences and their variants in the evolving viral quasispecies in cultivated solanaceous plants carrying mixed infections. In naturally infected *Solanum tuberosum* (potato), one case study revealed a virome comprising *Potato virus Y* (genus *Potyvirus*) and *Potato virus X* (genus *Potexvirus*), which was reconstructed by de novo–assembling separate genome-size sRNA contigs. Another case study revealed a virome comprising NTN and O strains of *Potato virus Y*, whose sRNAs assembled in chimeric contigs, which could be disentangled on the basis of reference genome sequences. Both viromes were stable in vegetative potato progeny. In a cross-protection trial of *Solanum lycopersicum* (tomato), the supposedly protective mild strain CH2 of *Pepino mosaic virus* (genus *Potexvirus*) was tested for protection against strain LP of the same virus. Reciprocal mechanical inoculations eventually resulted in co-infection of all individual plants with CH2 and LP strains, reconstructed as separate sRNA contigs. LP invasions into CH2-preinfected plants and vice versa were accompanied by alterations of consensus genome sequences in viral quasispecies, indicating a potential risk of cross-protection measures. Additionally, the study also revealed, by reconstruction from sRNAs, the presence of the mechanically nontransmissible *Southern tomato virus* (genus *Amalgavirus*) in some plants. Our in-depth analysis of sRNA sizes, 5'-nucleotide frequencies and hotspot maps revealed similarities in sRNA-generating mechanisms in potato and tomato, differential silencing responses to virome components and potential for sRNA-directed cross-targeting between viral strains which could not, however, prevent the formation of stable viromes.

RNA silencing is a sequence-specific mechanism regulating gene expression and mediating defense against invasive nucleic acids such as transposons and viruses in most eukaryotes. In all RNA-silencing pathways, the key effector molecules are small RNAs (sRNAs) that bind Argonaute (AGO) family proteins and guide the resulting RNA-induced silencing complexes toward complementary target sequences to direct gene silencing posttranscriptionally through mRNA cleavage and translational inhibition or transcriptionally through chromatin modification (Ghildiyal and Zamore 2009).

In plants, sRNAs are classified into micro RNAs (miRNAs) and small interfering RNAs (siRNAs) generated by Dicer-like (DCL) enzymes from hairpin-like structures of *MIR* gene transcripts or perfect double-stranded RNA (dsRNA) precursors, respectively (Borges and Martienssen 2015). Both miRNAs and siRNAs are sorted by AGOs based on the identity of the 5' nucleotide, size, and other factors (Fang and Qi 2016). RNA silencing can be reinforced by RNA-dependent RNA polymerase (RDR) activities generating dsRNA precursors of secondary and heterochromatic siRNAs (Borges and Martienssen 2015).

Plant-silencing pathways are mediated by multigenic families of DCLs, AGOs, and RDRs with specialized and partially redundant functions. The genetic requirements for biogenesis and action of endogenous and exogenous sRNAs have been thoroughly investigated in *Arabidopsis* and rice, while those in solanaceous plants, such as *Solanum lycopersicum* (tomato) and *Solanum tuberosum* (potato), are still poorly understood. Recent studies have begun to uncover the mechanisms of endogenous sRNA biogenesis and action in *S. lycopersicum* (Kravchik et al. 2014a and b; Yifhar et al. 2012). Like its orthologs in *Arabidopsis* and rice, tomato DCL1 generates 21- to 22-nt miRNAs playing crucial roles in plant development and physiology (Kravchik et al. 2014b). Like *Arabidopsis* DCL4, tomato DCL4 generates 21-nt transacting siRNAs that target auxin response–factor genes regulating leaf development (Yifhar et al. 2012). Tomato DCL3 generates not only 24-nt heterochromatic siRNAs (Kravchik et al. 2014a), like *Arabidopsis* DCL3 (Xie et al. 2004), but also 24-nt miRNAs and phased siRNAs (Kravchik et al. 2014a), like rice DCL3a and DCL3b, respectively (Song et al. 2012; Wu et al. 2010). Interestingly, four tomato paralogs of DCL2 (Bai et al. 2012) appear to serve as substitutes for other DCLs (Kravchik et al. 2014a and b) and possibly generate 22-nt siRNAs, like *Arabidopsis* DCL2 (Bouché et al. 2006; Xie et al. 2004).

By analogy with *Arabidopsis*, multiple DCLs, RDRs, and AGOs can potentially contribute to antiviral defense in

†Corresponding authors: Mikhail M. Pooggin;
E-mail: Mikhail.Pooggin@inra.fr and
Olivier Schumpp; E-mail: olivier.schumpp@agroscope.admin.ch

Funding: This work was funded by FP7 People: Marie-Curie Actions grant 608422 to Mikhail M. Pooggin and a Swiss Government Excellence Scholarship to Victor Golyaev.

*The e-Xtra logo stands for “electronic extra” and indicates that three supplementary figures, one supplementary table, and six supplementary datasets are published online.

Solanum species. In *Arabidopsis*, four DCLs mediate the biogenesis of 21-nt (DCL4 and DCL1), 22-nt (DCL2), and 24-nt (DCL3) virus-derived siRNAs (Akbergenov et al. 2006; Aregger et al. 2012; Blevins et al. 2006, 2011; Bouché et al. 2006; Fusaro et al. 2006; Malpica-López et al. 2018; Xie et al. 2004). RNA viruses replicating in the cytoplasm are targeted mostly by DCL4 and DCL2, while DNA viruses visiting the nucleus are targeted by all the four DCLs (Pooggin 2013, 2016). Although DCLs can directly process dsRNAs produced during RNA virus replication and DNA virus bidirectional transcription, RDR6 and RDR1 can, in some cases, generate dsRNA precursors of secondary viral siRNAs (Aregger et al. 2012; Wang et al. 2010, 2011). Multiple *Arabidopsis* AGOs function in antiviral defense through their association with both viral siRNAs and host miRNAs (Carbonell and Carrington 2015). RNA virus-derived 21- to 22-nt siRNAs are sorted, by at least three AGOs, following the 5'-nucleotide rule, with AGO1 binding mostly 5'U-siRNAs, AGO2 5'A-siRNAs, and AGO5 5'C-siRNAs (Takeda et al. 2008; Wang et al. 2011). So far, it is unknown which AGO (if any) has specificity for 5'G-siRNAs.

In all virus-infected plants examined so far, viral siRNA populations are extremely diverse, being derived from the entire virus genome in both sense and antisense polarities. The sRNA sequencing and bioinformatics (sRNA-omics) approach based on deep sequencing and bioinformatic analysis of sRNAs allows for universal diagnostics of DNA and RNA viruses as well as de novo reconstruction of their complete genomes (Kreuze et al. 2009; Seguin et al. 2014). Moreover, sRNA-omics enables reconstruction of the consensus genome sequence, identification of single nucleotide polymorphisms (SNPs) in evolving viral quasispecies (Kutnjak et al. 2015; Seguin et al. 2014 and 2016) and characterization of endogenous and antiviral RNA silencing pathways in cultivated plants (Fuentes et al. 2016; Rajeswaran et al. 2014a and 2014b).

Mixed infections are often observed in cultivated and wild plants (Syller and Grupa 2016). Two unrelated viruses in many cases interact in a synergistic manner, mutually promoting systemic infection and disease development, like, for example, the potyvirus *Potato virus X* (PVX) and the potyvirus *Potato virus Y* (PVY). In contrast, closely related strains of the same virus or isolates of the same viral strain usually compete with each other. Such antagonistic interactions are used in agriculture for protection of cultivated plants from severe viral strains (isolates) by deliberate preinoculation with mild strains (isolates) of the same virus (strain) (Syller and Grupa 2016; Ziebell and Carr 2010). Cross-protection and superinfection exclusion phenomena based on antagonistic interactions between viral strains or isolates or related viruses sharing homologous sequences can be explained by activation of the plant RNA silencing machinery generating viral siRNAs against one virus, which can also target another incoming virus in a sequence-specific manner (Ratcliff et al. 1999). In this case, the higher sequence identity between the two viruses, the better cross-protection is expected. On the other hand, suppression of RNA silencing by viral suppressor proteins (Csorba et al. 2015) can potentially lead to synergistic interactions between two related viruses or viral strains (Vanitharani et al. 2004), thus interfering with RNA silencing-based cross-protection and increasing incidence of mixed infections. It should be noted, however, that not all virus interactions in mixed infections can be explained by RNA silencing, implying other alternative mechanisms (Syller and Grupa 2016).

In this study, we applied sRNA-omics for reconstruction of viral genome sequences and characterization of defense responses to mixed infections in potato and tomato plants. We found that differential sRNA-based silencing responses and cross-targeting between virome components do not appear to destabilize a virome during vegetative propagation, while

mechanical co-inoculations of new plants are accompanied by appearance and fixation of new genetic variants in virome quasispecies. Our case studies of cultivated potato and tomato plants highlight the ability of closely related viruses to coexist in single plants and alter their genome sequences, which represents a potential risk to generate pathogenic recombinants when using 'mild strains' as agents for cross-protection.

RESULTS AND DISCUSSION

Stability of the PVY-PVX virome quasispecies despite differential silencing responses to PVY and PVX in vegetatively propagated potato plants.

With an ultimate goal to investigate RNA silencing-based interactions in mixed viral infections of cultivated solanaceous plants, we first performed a survey of potato tubers purchased in different local shops in Basel, Switzerland in 2014. We hypothesized that vegetatively propagated potato plants would frequently be co-infected with different viruses, while closely related strains of the same virus would be excluded from one plant due to cross-targeting by virus-derived siRNAs (or other mechanisms underlying superinfection exclusion phenomena).

Upper leaves from three potato plants, grown up from three eyes of one tuber (BPO-150, BPO-151, BPO-152), and one more plant from another tuber (BPO-147), which exhibited virus-like disease symptoms (yellow mosaic and deformations), as well as upper leaves from two healthy-looking plants grown from two separate tubers (BPO-148, BPO-149), were harvested in pools of several leaves per plant and, after total RNA preparation and sRNA quality validation, were subjected to sRNA-omics. Illumina sequencing of 10 cDNA libraries from 18- to 30-nt RNA fractions (all in one lane of HiSeq2500) yielded 29 to 38 million reads per library. The size range of 20 to 25 nt, constituting the majority of reads in each library (84 to 94%), was selected for further bioinformatic analysis, since it is known to be populated by viral siRNAs as well as plant miRNAs and siRNAs.

Our algorithm of de novo virome reconstruction consisted of the following steps: i) Oases assembly of redundant or non-redundant 20- to 25-nt reads, using k-mer values 15, 17, 19, and 21; ii) Seqman scaffolding and merging of all Oases contigs longer than 49 nts; iii) BWA mapping redundant 20- to 25-nt reads to Seqman contigs and most closely related reference sequence from National Center for Biotechnology Information (NCBI) GenBank, followed by MISIS-2 consensus sequence verification and reference-based gap filling, respectively.

In each of the three symptomatic plants (BPO-150 to BPO-152) stemming from one tuber, complete genomes of PVX and PVY were reconstructed. From one sRNA library (BPO-152), a single Seqman contig of 6,436 nt represented a complete PVX genomic RNA, while, from other libraries, two (BPO-150) or four (BPO-151) separate Seqman contigs covered a near-complete genome and the gap-filling step was executed to reconstruct the complete genome. Likewise, a complete PVY genomic RNA of 9,721 nt including an 18-nt poly(A) tail was reconstructed, either as single Seqman contigs (BPO-150 and BPO-152) with a few single nucleotide indels corrected manually or as two separate Seqman contigs (BPO-151) fused by gap-filling. In all three plants co-infected with PVX and PVY, their consensus genome sequences supported by the majority of redundant sRNA reads were found to be identical. Moreover, the positions and percentages of single nucleotide substitutions deviating from the consensus sequence (SNP profiles) were near identical between the plants (Supplementary Fig. S1). This indicates that PVX-PVY virome quasispecies are stable in the three separate plants grown from three different eyes of the same infected potato tuber. The consensus sequence of our PVX isolate (deposited to NCBI GenBank as MF405302) is 98% identical (122 SNPs) to that of PVX-OG

from Japan, deposited to GenBank in 2004 (AB196000), while the consensus sequence of our PVY isolate (deposited to GenBank as MF405303) is 99% identical (34 SNPs) to an IUNG-4 isolate of PVY-NTN strain from Poland, deposited in 2011 (JF927752). This illustrates a broad geographic distribution of PVY and PVX strains composing the virome we identified here from samples collected in Switzerland in 2014.

To characterize the RNA silencing response to the PVX-PVY virome, 20- to 25-nt sRNA reads from infected and healthy plants were mapped with zero mismatches to the reference genome sequences of *S. tuberosum* (nuclear, chloroplastic, and mitochondrial) as well as PVX and PVY (in combination and individually), and then, the mapped endogenous, and viral sRNAs were sorted by size (20, 21, 22, 23, 24, 25 nt), polarity (forward, reverse), and 5'-terminal nucleotide identity (5'A, 5'C, 5'G, 5'U) and were counted (Supplementary Dataset S1). In all libraries, the virome (PVY+PVX)-derived sRNAs constituted a large fraction (26 to 33%) of total (plant+viral) 20- to 25-nt sRNAs (Fig. 1A, BPO-150 to BPO-152). In each library, PVY-derived sRNAs were four to five times more abundant than PVX-derived sRNAs (Fig. 1B), indicating a difference in the plant silencing response to individual components of the PVY-PVX virome. The stronger response to the PVY component can potentially be explained by the mode of action of a PVY-encoded silencing suppressor HC-Pro, whose homolog from another potyvirus can sequester and inactivate viral siRNAs (Garcia-Ruiz et al. 2015). HC-Pro-mediated inactivation of viral siRNAs would in turn lead to higher accumulation of PVY RNA and its dsRNA derivatives

processed into viral siRNAs (discussed below). Besides the difference in sRNA accumulation levels, sRNA size class profiles of PVY and PVX differed substantially. While both viruses spawn predominantly 21-nt sRNAs and less abundant 22-nt sRNAs, the 22-nt class is relatively more abundant in PVX sRNAs (Fig. 1D). This suggests that potato DCL4 and DCL2 likely generating 21-nt and 22-nt siRNAs, respectively, may target PVY and PVX differentially. Analysis of 5'-nt identities of viral sRNAs revealed that both PVY and PVX spawn predominantly 5'U, 5'A, and 5'C sRNAs for each of the two major size classes and much less abundant 5'G sRNAs. This suggests that viral 21- to 22-nt sRNAs may predominantly be associated with potato orthologs of AGO1 (5'U), AGO2 (5'A), and AGO5 (5'C). It should be noted that the nucleotide content of the PVX (%A, C, G, U = 31, 24, 23, 23, respectively) and the PVY (%A, C, G, U = 31, 18, 23, 27) genomes cannot account for the observed paucity in virus-derived 5'G-sRNAs (9 to 13% of total for each virus). Furthermore, PVY HC-Pro binding to DCL-produced viral siRNA duplexes would probably be unspecific, since viral siRNAs associated with the potyviral HC-Pro in *Arabidopsis* did not exhibit any 5'-nt preference (Garcia-Ruiz et al. 2015). Thus, the observed enrichment of 5'U, 5'A, and 5'C in viral sRNAs in our case may reflect their association with the respective AGOs.

In the endogenous potato sRNAs, the 24-nt class dominated, in addition to 21-nt and less abundant 22-nt classes (Fig. 1C). The 24-nt sRNAs, being underrepresented in viral, mitochondrial, or chloroplastic sRNA populations, are derived predominantly from the potato nuclear genome and have a

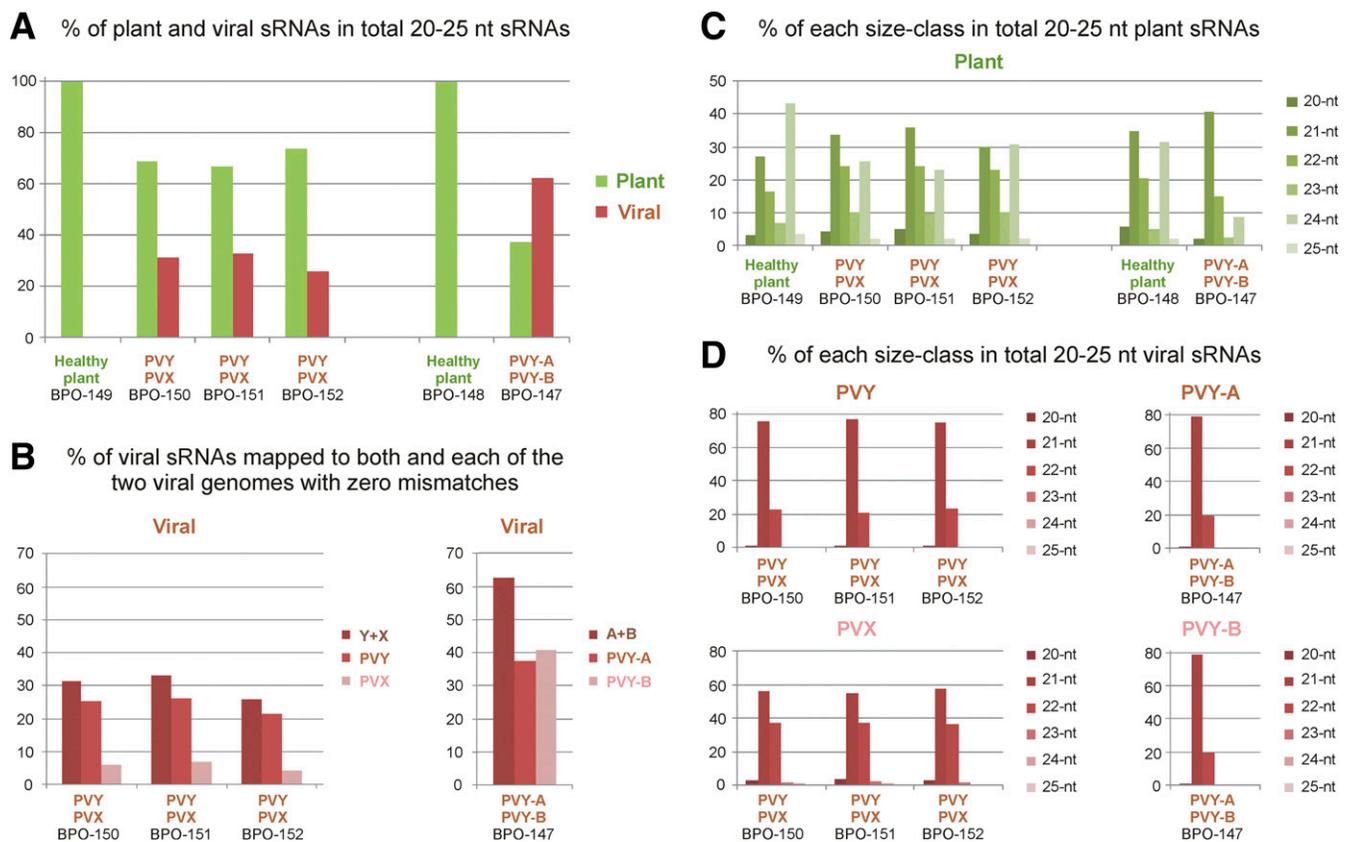


Fig. 1. Illumina sequencing counts of endogenous and viral small RNAs (sRNAs) in *Solanum tuberosum*. The 20- to 25-nt sRNA libraries from healthy (BPO-148 and BPO-149) and *Potato virus Y* and *Potato virus X* (PVY+PVX; BPO-150 to BPO-152) or PVY-A^N+PVY-B^O (BPO-147) virome-infected plants were mapped to the combined virome or each of the individual viral reference sequences and the plant genome reference sequence with zero mismatches and were counted. **A**, Percentage of virome- and plant-derived sRNAs in the pool of total 20- to 25-nt reads. **B**, Percentage of each of the individual virus-derived sRNAs in the combined virome sRNAs. **C** and **D**, Percentage of each size class in the 20- to 25-nt pool of plant-derived (C) or each virus-derived (D) sRNA read. Note that subtitles above graphs indicate the identity of mapped sRNAs, i.e., endogenous plant sRNAs (plant), total virus-derived sRNAs (viral), specific virus-derived sRNAs (PVX or PVY), and viral strain-derived sRNAs (PVY-A or PVY-B).

strong bias to 5'A (71 to 72%), thus resembling *Arabidopsis* heterochromatic siRNAs (Havecker et al. 2010). Notably, 24-nt potato sRNAs are relatively less abundant in virus-infected plants, in which 21-nt and 22-nt sRNAs become relatively more abundant (Fig. 1C). Analysis of 5'-nt frequencies revealed that the nuclear genome-derived 21-nt sRNAs with 5'U, likely representing miRNAs generated by DCL1 and associated with AGO1, are highly abundant in healthy plants (76 to 80% of total 21-nt reads) and much less abundant in PVX+PVY-infected plants (52 to 57%). Biological significance of the observed changes in host sRNA profiles remains to be investigated.

Analysis of single-nucleotide resolution sRNA maps (Supplementary Dataset S2) generated by MISIS-2 (Seguin et al. 2016) revealed that viral sRNA species of major size classes (21-nt and 22-nt) cover the entire genomes of PVY and PVX in both forward and reverse orientations without gaps. The hotspots of sense and antisense sRNA production are almost equally distributed along the PVY genome, whereas those are more abundant in a 5' portion of the PVX genome containing an RdRP open reading frame (ORF) compared with a 3' portion of the genome containing triple gene block (TGB1, TGB2, TGB3) and coat protein (CP) ORFs (Fig. 2). Unlike PVY that translates a single large poly-protein from genomic RNA, PVX translates TGB1-3 and CP from two separate subgenomic RNAs (sgRNAs). Almost equal abundance of sense and antisense reads along PVY and PVX genome sequences for each major sRNA size (Fig. 2) suggests that 21-nt and 22-nt viral sRNAs are processed by respective DCLs (likely DCL4 and DCL2) in the form of duplexes from perfect dsRNA precursors covering the entire virus genome.

Comparison of single-nucleotide resolution maps of viral sRNAs in BPO-150, BPO-151, and BPO-152 (Fig. 2) indicates that the biogenesis mechanisms for PVY- and PVX-derived 21-nt and 22-nt siRNAs are both qualitatively and quantitatively similar in all three PVY-PVX virome-infected plants. Taken together, the potent antiviral RNA silencing machinery constantly but differentially targeting PVY and PVX does not appear to destabilize the virome complex or modify the consensus sequences in the virome quasispecies during vegetative propagation of potato plants under greenhouse conditions.

Reconstruction of two distinct PVY strains co-infecting a single potato plant.

In a plant grown from the second infected tuber (BPO-147), de novo virome reconstruction revealed the presence of two strains of PVY, whose sRNAs were assembled in two Seqman contigs, the 9,710-nt contig_6, covering a complete PVY including a 24-nt poly(A) tail, and the 1,495-nt contig_9, representing the 5' terminal portion of PVY. Contig_6 contained numerous degenerate nucleotides throughout its sequence, except for the 5' terminal portion of 1,535 nt, suggesting its chimeric nature (Fig. 3A). Indeed, mapping of Oases contigs back to contig_6 and contig_9, followed by map inspection using the Interactive Genomic Viewer (IGV), revealed two contig categories with high sequence similarity, which were merged by Seqman, creating the chimeric region of contig_6 with degenerate nucleotides as well as two contig categories with much lower sequence similarities (74%), which built the 5'-terminal part of contig_6 lacking degenerate nucleotides and the separate contig_9 (Fig. 3A). BLAST analysis revealed that the 5' part of contig_9 is 99% identical to the corresponding region of many isolates of PVY strain NTN, whereas contig_6 is 99% identical to the 5' part of many isolates of PVY strain O. To confirm the presence of two strains in our virome (operationally named PVY-A^N and PVY-B^O), we used a reverse transcription-polymerase chain reaction (RT-PCR) sequencing approach in which several overlapping amplicons of approximately 700 to 900 bp were obtained, in each case, using two

forward primers corresponding to a region of sRNA sequence dissimilarity in the pair with one reverse primer corresponding to a region of sRNA sequence identity. This approach enabled us to disentangle 4,002-nt 3' terminal parts of PVY-A^N and PVY-B^O, which shared 87% identity and matched best to the corresponding parts of an isolate IUNG-4 of PVY-NTN (JF927752, 99% identical) and isolate CO2140 of PVY-O (HQ912914, 99% identical). We then disentangled the remaining middle portions of PVY-A^N and PVY-B^O (sharing 95% identity) by mapping redundant sRNA reads to the reference sequences of IUNG-4 and CO2140, visualizing SNPs by MISIS-2 and, if needed, correcting the reference sequence by assigning two major nucleotide variants to respective strains (Fig. 3A).

It should be noted that reconstruction of PVY-A^N and PVY-B^O (deposited to GenBank as MF422609 and MF422610, respectively) would have been impossible without the reference sequences, given high nucleotide identity throughout the genome sequences, except for 5' terminal parts. The latter parts share only 74% identity, allowing the short-read assembler Velvet/Oases to generate separate nonchimeric sRNA contigs. In contrast, the higher nucleotide identities (87 to 95%) in the middle and the 3' terminal portions of PVY-A^N and PVY-B^O prevented separation of Oases contigs by Seqman, thus creating the single chimeric contig (Fig. 3A). Given the coexistence of PVY-A^N and PVY-B^O in one plant virome, we cannot exclude potential recombination events leading to swaps of short or long sequences between the two strains, further complicating reliable separation of the two consensus sequences. Our attempts to physically separate PVY-A^N and PVY-B^O in vegetative progeny of the infected plant failed because both strains were detected by RNA blot hybridization with strain-specific probes in seedlings grown from each of the four progeny tubers tested. Thus, the virome composed of two PVY strains appears to be as stable as the PVY-PVX virome described above.

Comparison of potato RNA silencing responses to distinct viromes.

To evaluate RNA silencing response to two PVY strains present in one plant, we mapped sRNAs with zero mismatches to the reconstructed reference sequences of PVY-A^N and PVY-B^O, either in combination, allowing equal (50:50%) distribution of identical sRNAs between the two strains, or separately, allowing identical sRNAs from both strains to be mapped on each strain. We also mapped sRNAs to the 5' terminal (head) and the 3' terminal (tail) parts of PVY-A^N and PVY-B^O. The sum of viral reads from both strains together constituted 62.5% of total (plant+viral) 20- to 25-nt reads (Fig. 1A). This shows that the PVY-A^N-PVY-B^O virome triggers a two times stronger silencing response than the PVX-PVY virome, although a PVY component of the latter attracted approximately 80% of the overall response (Fig. 1B). The identical sRNAs from both PVY-A^N and PVY-B^O constituted approximately 25% of all viral sRNAs, which elevated the counts of viral sRNAs mapped separately to PVY-A^N (37.6% of total) or PVY-B^O (40.8% of total) (Fig. 1B). This shows a high cross-targeting potential of viral siRNAs derived from PVY-A^N against PVY-B^O and vice versa. In contrast, cross-targeting potential is negligible for the PVX-PVY virome, since PVX and PVY do not share any significant proportion of identical sRNA reads.

In-depth analysis and comparison of viral sRNA size profiles, 5'-nt frequencies, and hotspot maps revealed remarkable similarities between PVY-A^N and PVY-B^O as well as a PVY component of the PVX-PVY virome (Figs. 2 and 3B). This indicates that the potato RNA silencing machinery targets the potyvirus PVY in a consistent and reproducible manner in different individual plants. Moreover, the biogenesis of viral 21-nt and 22-nt siRNAs is not substantially influenced by a second virome component (be it another strain of PVY or an unrelated potyvirus). In a previous study, deep sequencing of

MISIS-generated maps of viral sRNAs aligned to PVY and PVX consensus genome sequences with zero mismatches

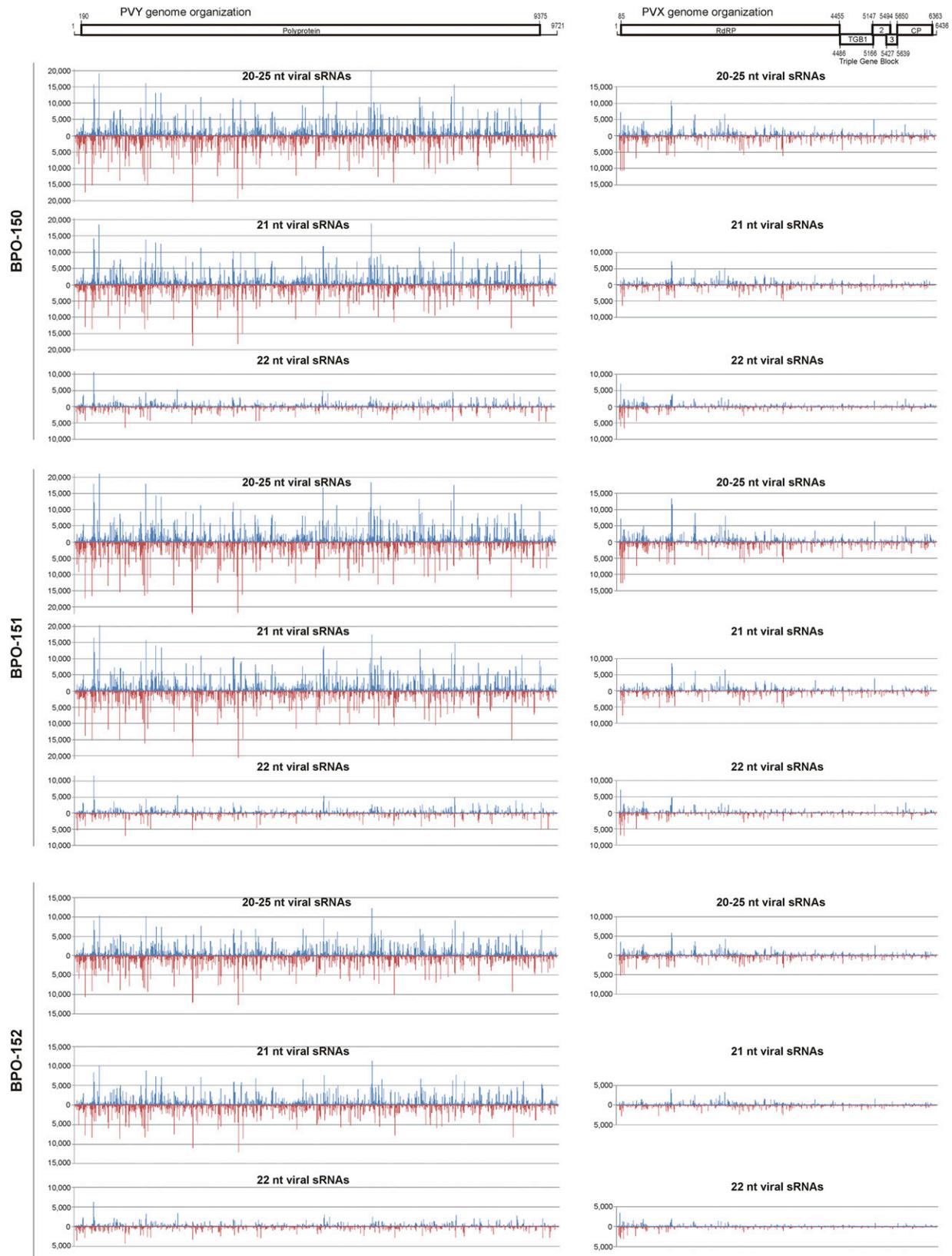


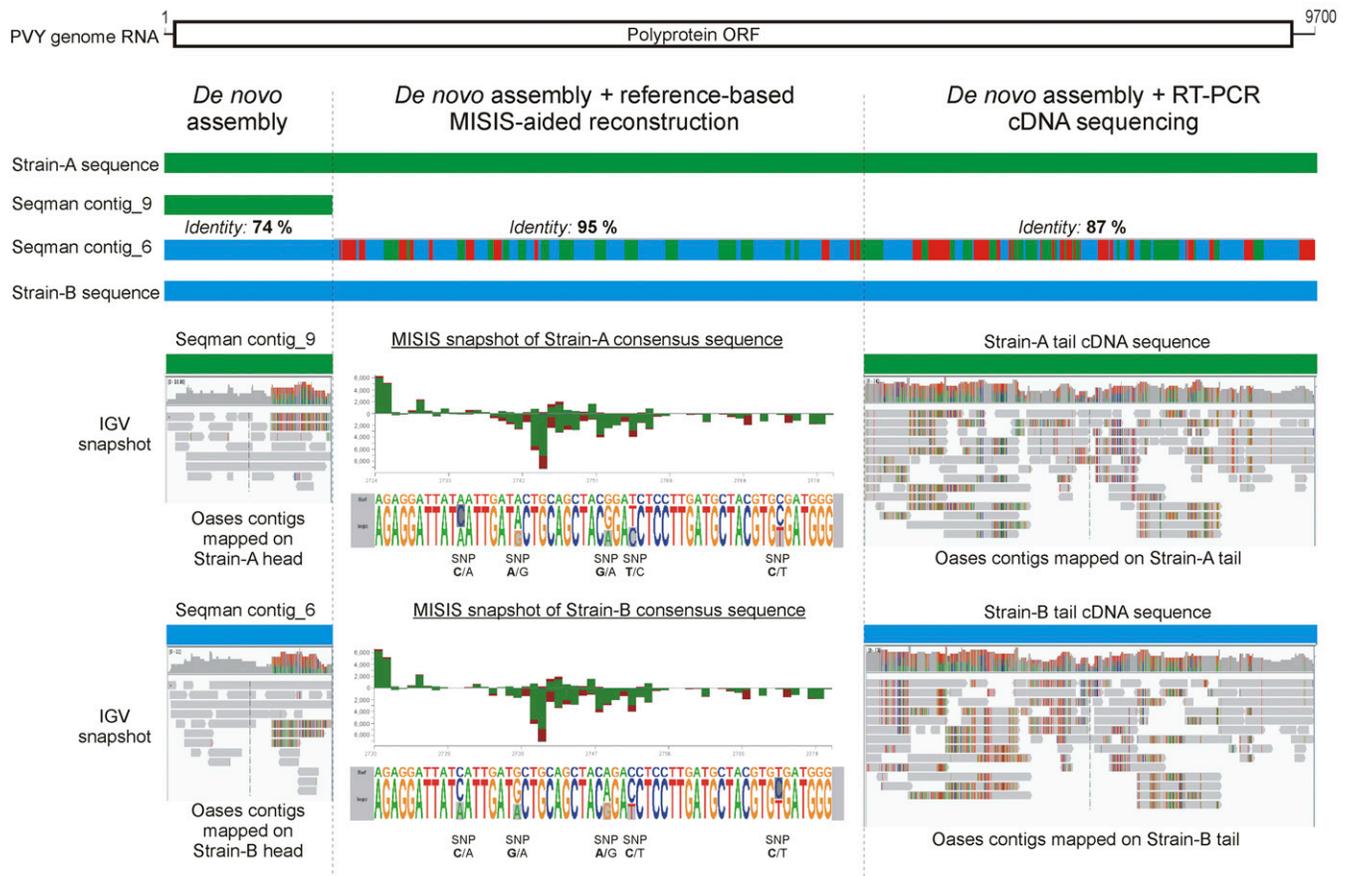
Fig. 2. Single-nucleotide resolution maps of viral small RNAs (sRNAs) from leaf tissues of three plants grown from a potato tuber infected with a *Potato virus Y* and *Potato virus X* (PVY-PVX) virome. For each plant sample (BPO-150, BPO-151, and BPO-152), the histograms plot the numbers of total 20- to 25-nt, 21-nt, or 22-nt viral sRNA reads at each nucleotide position of 9,721-nt PVY and 6,436-nt PVX genome sequences (mapped with zero mismatches). The bars above the axis represent sense reads starting at each position and those below represent antisense reads ending at the respective position. Scaled PVY and PVX genome diagrams are shown above the histograms, with the open reading frames boxed and their encoded proteins and nucleotide positions indicated.

sRNAs from potato plants singly infected with PVY-O, PVY-NTN, or PVY-N revealed both similarities and differences in viral sRNA biogenesis and endogenous sRNA profile alteration (Naveed et al. 2014). Interestingly, the silencing responses to PVY-O and PVY-NTN were similar and generally resembled those observed here for our PVY-O and PVY-NTN isolates

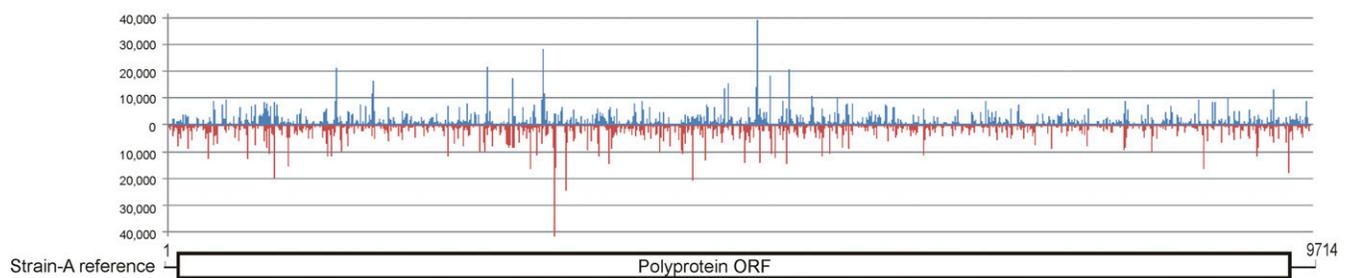
present in the PVY-A^N-PVY-B^O virome and PVY-NTN isolate in the PVX-PVY virome, thus supporting our main conclusions.

Our findings that the plant-silencing machinery almost equally targets each component of the PVY-A^N-PVY-B^O virome generating massive quantities of viral 21- and 22-nt siRNAs from each strain and that the cross-targeting potential of viral

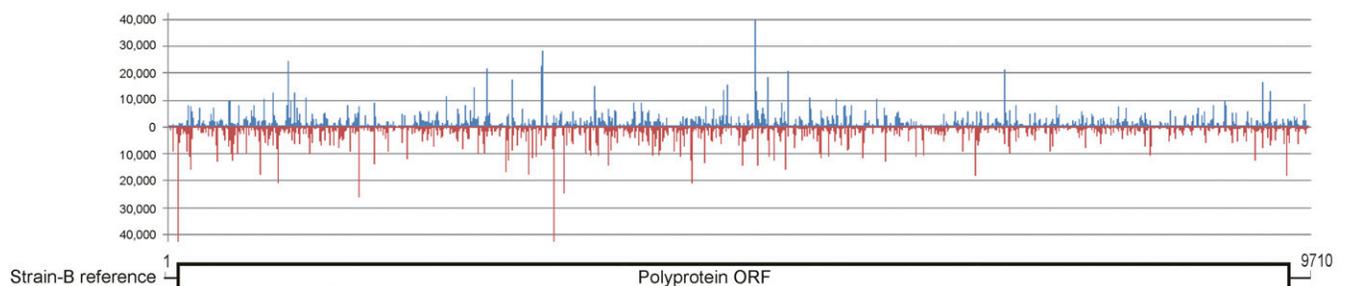
A Reconstruction of PVY twin strains



B MISIS map of viral 20-25 nt sRNAs aligned to Strain-A consensus sequence with zero mismatches



MISIS map of viral 20-25 nt sRNAs aligned to Strain-B consensus sequence with zero mismatches



siRNAs does not seem to affect the virome integrity or stability in vegetative progeny, suggest that PVY-A^N and PVY-B^O interact with each other neutrally (or synergistically) rather than antagonistically. In a previous study, co-inoculation of potato and tobacco plants with different isolates of three PVY strains have revealed both antagonistic and neutral interactions, depending on the strain and isolate combination and a host plant (Syller and Grupa 2014). The interactions between PVY-NTN and PVY-O were usually antagonistic, resulting in suppression of PVY-O accumulation by PVX-NTN, as compared with single infections (Syller and Grupa 2014). Nonetheless, it cannot be excluded that dual infections with certain genetic variants of these strains can, under natural conditions, form a stable virome, like the PVY-A^N-PVY-B^O virome described here.

Interactions between and molecular evolution of mild CH2 and LP strains of *Pepino mosaic virus* (PepMV) in *Solanum lycopersicum*.

To investigate if the mechanism underlying cross-protection and superinfection exclusion phenomena is based on RNA silencing, we performed a cross-protection trial of tomato plants at Agroscope in 2014. In this trial, a mild isolate of PepMV strain CH2 (De Nayer et al. 2011; Hanssen et al. 2010), marketed under tradename PMV-01 (Sciencia Terrae), was used for plant protection against an isolate of PepMV strain LP, originating from a tomato cultivar Merlice plant collected in Champagny, Switzerland in October 2013. We hypothesized that viral siRNAs derived from the protective commercial strain CH2 in CH2-preinfected plants would cross-target identical or near-identical sequences of the challenging strain LP and thereby prevent co-infections.

On January 23, 2014, young plants of tomato cultivar Merlice grown in a production greenhouse (Conthey, Switzerland) were spray-inoculated with the protective strain CH2. In another greenhouse (Nyon, Switzerland) at about the same time, young plants of tomato cultivar Mt. Favet were inoculated with the strain LP. After about two months (on March 07, 2014) four Mt. Favet plants infected with LP were transferred from Nyon to Conthey, where they were cocultivated with CH2-preinfected Merlice plants in order to evaluate any protective effect of CH2 against LP. The cocultivation procedure involved deliberate mechanical cross-contamination between all the greenhouse plants, using leaf-cutting and fruit-harvesting tools, thus allowing repetitive inoculations of CH2-infected plants by sap from LP-infected plants and vice versa. Upper leaves of two selected Merlice plants preinfected with CH2 and challenge-inoculated with LP were sampled at two timepoints, 1.5 months (April 17, 2014) and 6.5 months (September 30, 2014) postintroduction of LP. Likewise, upper leaves of two selected Mt. Favet plants preinfected with LP and challenge-inoculated with CH2 were sampled at two timepoints (July 18 and September 30, 2014). Total RNA was extracted from a total of eight leaf samples of the infected plants as well as from upper leaf samples of two control nontreated Mt. Favet plants (harvested in February 2014) and was analyzed by sRNA-omics followed by Northern blot hybridization.

Illumina sequencing of all cDNA libraries in one lane of HiSeq2500 yielded 20.2 to 31.5 million reads per library. The 20- to 25-nt reads were taken for bioinformatic analysis as described above for the potato libraries. De novo virome reconstruction revealed the presence of CH2 and absence of LP in both plants preinfected with CH2 at the first timepoint, indicating CH2-mediated cross-protection during 1.5 months of repetitive challenge-inoculation with LP. However, at the second timepoint, both plants were found to be co-infected with CH2 and LP, indicating a breach in cross-protection at some point during the next 5 months of repetitive challenge-inoculation with LP. Likewise, virome reconstruction revealed the presence of both LP and CH2 in both plants preinfected with LP after 4.5 months of repetitive challenge-inoculation with CH2 as well as at the later timepoint (6.5 months). Northern blot hybridization analysis confirmed the sRNA-omics findings and further revealed that LP invasion had reduced the accumulation levels of gRNA and, more pronouncedly, sgRNAs of CH2 (Fig. 4, CH2-specific probe, lanes 1 and 3 versus 2 and 4), suggesting antagonistic interactions between CH2 and LP in mixed infections. In previous studies, competition between PepMV strains in mixed infections did not appear to affect viral loads, compared with control single infections (Hanssen et al. 2010; Gómez et al. 2009), although sgRNAs were not analyzed.

De novo reconstructed consensus sequences of CH2 from the two single infections (BPO160 and BPO-162; deposited to GenBank as MF422611 and MF422613, respectively) differed at three nucleotide positions of the 6,412-nt genome. MISIS-aided analysis revealed that, in plant 2, the nucleotide substitutions A2041G, T3358C, and T5224G are supported by 80 to 85% reads at the first timepoint (BPO-162) and by almost 100% reads at the second timepoint (BPO-163) (Supplementary Fig. S2). These three substitutions were absent in plant 1 at the first (BPO-160) or second (BPO-161) timepoints as well as in two plants preinfected with LP at the first (BPO-156 and BPO-158) or second (BPO-157 and BPO-159) timepoints. In one of the latter plants, two additional nucleotide substitutions (G2875A and T3265C) were found, both supported by the majority of reads at the two timepoints (BPO-158 and BPO-159) (deposited to GenBank as MF422615). These findings led us to conclude that, in one of the two Merlice plants inoculated with CH2, the replicating virus has acquired the three mutations that started to dominate in viral quasispecies at the early timepoint and eventually outcompeted an original consensus genome sequence, as evident at the later timepoint. Only the original genetic variant of CH2 (MF422611) invaded both LP-infected Mt. Favet plants, and then, either stably persisted in one plant or acquired two mutations in another one. The on-going evolution of CH2 quasispecies is further illustrated by comparison of our BPO-160 variant (MF422611) derived from PMV-01 with the initial sequence of this isolate deposited in Genbank in 2011 (JN835466). CH2 has accumulated 16 SNPs throughout the genome sequence as well as a 1-nt deletion and a 1-nt insertion separated by 9 nt in the RdRP ORF, which together altered seven amino acids in RdRP and one amino acid in each of the other proteins (TGB1, TGB2, TGB3, and CP).



Fig. 3. Reconstruction of *Potato virus Y* (PVY) strains co-infecting a potato plant. **A**, Schematic illustration of the reconstruction of the genome sequences of two PVY strains through three steps: i) de novo assembly of viral small RNAs (sRNAs) generating two Seqman contigs, the short contig_9 representing a 5' portion of PVY-A^N and the long genome-length chimeric contig_6 containing sequences from both strains (in respective shades) and degenerate nucleotide sequences (in a different shade); ii) reverse transcription-polymerase chain reaction sequencing-based separation of PVY-A^N and PVY-B^O genome sequences in their 3'-portions; and iii) reference-based separation of the middle portions of PVY-A^N and PVY-B^O genomes aided by MISIS-2. Interactive Genomic Viewer (IGV) snapshots of the Oases contigs mapped to PVY-A^N and PVY-B^O genome sequences illustrate two categories of the contigs for each of the two strains (solid gray or fully mixed shades) as well as the chimeric contigs containing parts from another strain (partially mixed shades) (Thorvaldsdóttir et al. [2013] provides details of IGV mapview). MISIS snapshots illustrate single nucleotide polymorphism positions in the DNA logo with two major nucleotides representing either PVY-A^N or PVY-B^O (Seguin et al. [2016] provides details of MISIS mapview). PVY genome organization is shown above with the polyprotein open reading frame (ORF) boxed. **B**, Histograms plotting the numbers of total 20- to 25-nt viral sRNA reads at each nucleotide position of 9,714-nt PVY-A^N and 9,710-nt PVY-B^O (mapped with zero mismatches). The bars above the axis represent sense reads starting at each position and those below represent antisense reads ending at the respective position. PVY genome diagrams are shown above the histograms, with the polyprotein ORF boxed and nucleotide positions indicated.

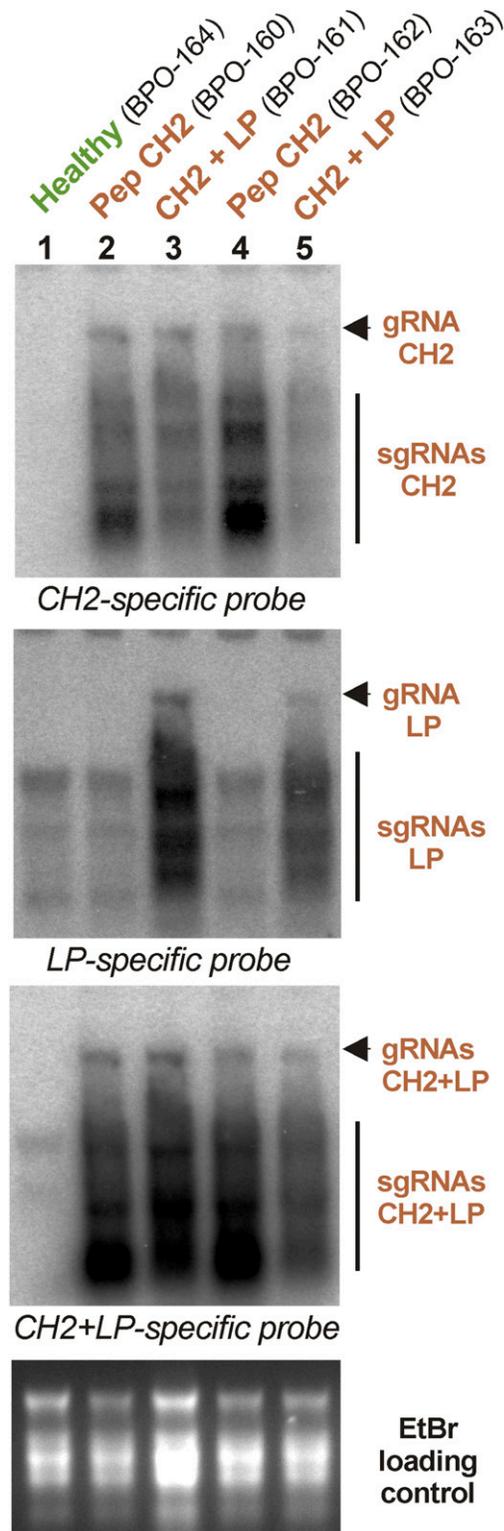


Fig. 4. Blot hybridization analysis of viral RNA accumulation in *Solanum lycopersicum* plants infected with one (CH2) or two strains (CH2+LP) of *Pepino mosaic virus*. Aliquots (5 µg) of the total RNA samples from healthy (BPO-164) and virus-infected plants (BPO-160 to BPO-163) pre-analyzed by sRNA sequencing were separated on formaldehyde-containing agarose gel, were stained with ethidium bromide (EtBr), and were blotted to a membrane. The membrane was successively hybridized with DNA oligonucleotide probes specific for the reconstructed reference sequences of CH2, LP, or both (CH2+LP) strains and was scanned. The scanned images after each hybridization are shown with the probe indicated below and the positions of viral genomic RNA (gRNA) and subgenomic RNAs (sgRNAs) indicated on the right. EtBr staining of the gel is shown as loading control.

PMV-01 is known to ensure only partial protection against non-CH2 strains of the virus (De Nayer et al. 2011). We confirm and extend this finding by showing that PMV-01/CH2 stops LP dissemination for at least 1.5 months, but repetitive challenge inoculation does break CH2-mediated cross-protection against LP. Furthermore, we demonstrate that the PMV-01 quasiespecies evolves on passages to new host plants through appearance and fixation of mutations. This on-going evolution may eventually result in either a loss of protectiveness, an alteration of symptom severity, or both. Indeed, mild (FJ457096) and aggressive (FJ457097) isolates of PepMV CH2 differ only by 38 SNPs altering nine amino acids (Hanssen et al. 2009). Moreover, two studies have shown that a single point mutation is sufficient to increase PepMV aggressiveness. In one study, a point mutation altering one amino acid in TGB3 at position 67 (K [AAA] to E [GAA]) could convert a mild PepMV pathotype into a necrotic one (Hasiów-Jaroszewska et al. 2011), while, in another study, single amino acid substitutions in CP at position 155 (E [GAA] to K [AAA]) or position 166 (D [GAT] to G [GGT]) resulted in severe yellowing symptoms (Hasiów-Jaroszewska et al. 2013). Our analysis of the evolving CH2 quasiespecies revealed the nonconsensus nucleotides A and G at the two respective positions of CP, albeit in a tiny proportion of reads (six of 30,337 and five of 29,537 total reads from the six CH2-infected plants), while no polymorphism was found at the corresponding nucleotide position of TGB3 in any CH2-infected plant (0 of 2,725 total reads). The quasiespecies nature of PepMV has previously been established by analyzing PCR-amplified TGB1 sequences from several naturally infected tomato plants (Hasiów-Jaroszewska et al. 2010). Our genome-wide analysis of consensus sequence variation reveals appearance of SNPs and fixation of new consensus nucleotides in PepMV quasiespecies within individual plants during the 8-month trial.

De novo reconstruction of the LP consensus sequences was somewhat complicated by the presence of CH2 in all the LP-infected plant samples and required MISIS-aided corrections at SNP positions within several stretches of high similarity between LP and CH2 genome sequences (sharing 78% overall identity). With this reservation, the LP consensus sequence reconstructed from two CH2-preinfected plants, in which the CH2 consensus sequence is known from the earlier timepoint, differ only by one nucleotide, T4191 in BPO-161 (deposited to GenBank as MF422612) versus C4191 in BPO-163 (deposited to GenBank as MF422614). In both plants preinfected with LP, the LP consensus sequence appeared, with the above reservations, to be identical to that from BPO-163, suggesting that this LP variant invaded both Merlice plants preinfected with CH2 and then acquired the point mutation in one of them (BPO-161). BLAST analysis revealed that the 6,408-nt LP consensus sequence from BPO-163 (MF422614) is 99.8% identical (14 SNPs) to a sequence of PepMV isolate LP-2001 (AJ606361) obtained in Peru from an asymptomatic Peruvian tomato (*Solanum peruvianum*, also known as *Lycopersicon peruvianum*) plant (López et al. 2005) (hence the name LP for our isolate). Interestingly, a nearly identical isolate, SM.74 (AM109896), differing by 18 SNPs from the BPO-163 sequence, has been obtained in Peru from a pepino (*Solanum muricatum*) plant (Pagán et al. 2006). Thus, PepMV LP mutated upon arrival in Europe from Peru and eventually infected the tomato plant that was collected in Switzerland for our cross-protection trial.

To verify if our LP isolate undergoes quasiespecies evolution in single infections, we repeated (in 2016) inoculation of tomato Mt. Favet plants with newly prepared sap from the LP-infected leaf tissues used for the cross-protection trial and sequenced sRNAs from uppermost leaves of LP-infected (HYT-25) and mock-inoculated (HYT-26) plants. De novo reconstruction of the

consensus genome sequence of LP from HYT25 (deposited to GenBank as MF422616) revealed 13 SNPs, compared with that from BPO-163 (A138C, C1548T, A1609G, T1997C, A2272G, C2394A, T3138C, G3225A, T3486C, C4418T, G5861A, C5987T, T6254C), each supported by the majority of redundant and nonredundant sRNA reads. Interestingly, seven of these 13 substituted nucleotides (C1548, T3138, G3225, T3486, G5861, C5987, T6254) are shared between BPO-163 and the Peruvian isolates LP-2001 and SM.74, indicating that the corresponding substitutions are new genetic variants fixed in viral quasispecies of the HYT-25 plant. On the other hand, nucleotides C138, G1609, C1997, G2272, A2394, and T4418 are shared between HYT-25 and both Peruvian isolates, indicating that those six nucleotides were substituted and fixed through independent evolution before preinoculation of the tomato plants (followed by stable transmission of this genetic variant into the CH2-infected plant BPO-163). Thus, similar to CH2, the LP quasispecies undergoes continuous microevolution of the consensus genome sequence on passages in tomato plants cultivated under greenhouse conditions.

Tomato RNA silencing response to CH2 and LP strains of PepMV in single and mixed infections.

We analyzed the tomato silencing response to CH2 and LP by mapping 20- to 25-nt reads to the tomato genome and the reconstructed consensus sequences of CH2 and LP (individually and in combination) with zero mismatches and by counting the endogenous and viral sRNAs based on size, polarity, and 5' nucleotide identity (Supplementary Dataset S3). The proportion of all viral (CH2+LP) sRNAs in a total (tomato+viral) 20- to 25-nt sRNA population varied from 1.6 to 4.7% (Fig. 5A), thus being comparable to the proportions reported for tomato plants co-infected with EU and US1 strains of PepMV (Li et al. 2012) and also to the proportion of PVX-derived sRNAs in potato plants (Fig. 1B). Notably, viral sRNA accumulation was approximately two to 2.5 times higher in the plants infected with CH2 alone (3.8 to 4.7%), compared with the plants co-infected with CH2 and LP (1.6 to 1.8%). In all co-infected plants, the levels of CH2-derived sRNAs somewhat exceeded those of LP-derived sRNAs, although the difference was less pronounced in the plants preinfected with LP (Fig. 5B). Comparison of size-class profiles revealed a striking difference between single (CH2) and dual (CH2+LP) infections; while the 21-nt class is dominant in all cases for both CH2 and LP, the 22-nt class of CH2-derived sRNAs is much more prominent in single infections (Fig. 5D). This suggests that LP may selectively interfere with production of 22-nt siRNAs from CH2, which contributes to an overall reduction of viral sRNA accumulation in the plants preinfected with CH2 upon invasion of LP (Fig. 5A) and correlates with reduced levels of CH2 gRNAs and sgRNAs (Fig. 4). Analysis of sRNAs from the tomato plant infected with LP alone (HYT-25; Supplementary Fig. S3) confirmed the results for dual infections in that LP-derived sRNAs were dominated by the 21-nt class and constituted a smaller proportion in total 20- to 25-nt sRNAs (0.6%) than CH2-derived sRNAs. Analysis of 5' nt frequencies revealed a substantial bias to 5'U and 5'A in both 21-nt and 22-nt viral sRNAs derived from both CH2 and LP in dual infections, while this bias was less pronounced in single infections due to higher frequencies of 5'C. Thus, PepMV-derived 21- and 22-nt sRNAs appear to be sorted by tomato AGO1, AGO2, and AGO5, similar to PVX-derived sRNAs in potato plants.

Analysis of size profiles and 5'-nt frequencies of endogenous tomato further confirmed a differential impact of single (CH2) and dual (CH2+LP) infections of the plant RNA-silencing machinery, although the size profile alterations are not always consistent between biological replicates (Fig. 4B). Notably,

21-nt 5'U sRNAs, which dominate in healthy plants (79%) and likely represent DCL1-dependent AGO1-associated miRNAs (Kravchik et al. 2014b), are less dominant in all infected plants (62 to 73%). Likewise, 24-nt 5'A sRNA, which dominate in healthy plants (73%) and likely represent DCL3-dependent, AGO4-associated heterochromatic and phased sRNAs (Kravchik et al. 2014a), become less dominant (61 to 63%) in infected plants. Biological significance of the impact of PepMV infections on endogenous sRNAs remains to be investigated.

Analysis of single-nucleotide resolution maps of CH2- and LP-derived sRNAs revealed similarities to those of PVX-derived sRNAs. Both 21-nt and 22-nt viral sRNA species cover the entire viral genome in sense and antisense orientation (Supplementary Datasets S4 and S5) and the sRNA hotspots generally map to the 5' portion of the RdRP ORF (Figs. 6 and 7). A notable difference between CH2 and LP is that a sgRNA region (TGB-CP ORFs) of the CH2 genome spawns relatively more abundant sRNAs than that of the LP (or PVX) genome. Furthermore, sRNA hotspot patterns appear to be unique for CH2 and LP and remarkably reproducible in all the plants at both timepoints (Figs. 6 and 7), indicating sequence-specificity of the viral siRNA-generating machinery targeting LP and CH2. Similar observations have been reported for two distinct DNA begomoviruses in tomato plants (Fuentes et al. 2016).

To assess a cross-targeting potential of CH2-derived sRNAs against LP, we mapped 20- to 25-nt sRNA reads from CH2-infected plants at the first timepoint, when no LP was detectable (BPO-160 and BPO-162), to the LP genome. A significant proportion of CH2-derived sRNAs (1.2 and 1.4%, respectively) mapped to the LP genome with zero mismatches (note that the proportion is elevated to 20.5 and 22.5%, respectively, if up to two mismatches are allowed). The majority of these identical sRNAs concentrate in five hotspots of 21- and 22-nt sRNAs of both polarities across the LP genome (Fig. 6, BPO-160 and BPO-162), which contain continuous stretches of more than 20 nt with 100% identity between CH2 and LP sequences. These 21-nt and 22-nt siRNAs derived from CH2 can potentially target LP gRNA, gRNA(-), and sgRNAs for cleavage and degradation. Such cross-targeting might have contributed to CH2-mediated cross-protection against LP for at least 1.5 months. However, a relatively small quantity of the cross-targeting sRNAs and a relatively short length of the cross-targeted sites in the LP genome may account for the failure of CH2-infected plants to resist repetitive challenge-inoculation with LP for a longer term. Consistent with our findings, mixed infections of distinct PepMV strains have been observed in a previous cross-protection trial (Hanssen et al. 2010) as well as under natural conditions (Gómez et al. 2009), indicating that PepMV strains can form a stable virome. On the other hand, a previous trial has suggested that the protecting effect of a mild variant of CH2 (PVM-01) against an aggressive variant of CH2 could be based on RNA silencing because of 99.4% sequence identity (Hanssen et al. 2010). However, our above described evidence for the PVY-A^N-PVY-B^O virome indicates that even a very high cross-targeting potential for viral sRNAs does not prevent the coexistence of two viral strains in a stable virome.

It is worth emphasizing that we obtained only indirect evidence for the hypothesis that CH2-derived siRNAs can potentially target the invading LP strain, which includes three lines: i) cross-protection of CH2-preinfected plants during at least 1.5 months of repetitive challenge-inoculation with LP, ii) identification of CH2-derived siRNAs that map to the LP genome in sense and antisense orientation with zero mismatches (Fig. 6), and iii) potential association of CH2 siRNAs with AGO1, AGO2, and AGO5, based on the observed predominance of 5'U, 5'A, and 5'C in 21- and 22-nt viral siRNAs in CH2-infected plants. Direct evidence for the ability of viral siRNAs to program AGO1, AGO2,

and AGO5 proteins for siRNA-directed cleavage of complementary viral RNA sequences and to inhibit replication of the targeted viral RNA has been obtained in an in vitro study (Schuck et al. 2013), further supporting our hypothesis.

Identification and reconstruction of *Southern tomato virus* (STV) by sRNA-omics.

In addition to CH2 and LP strains of PepMV, de novo virome reconstruction revealed the presence of STV in the two Merlice

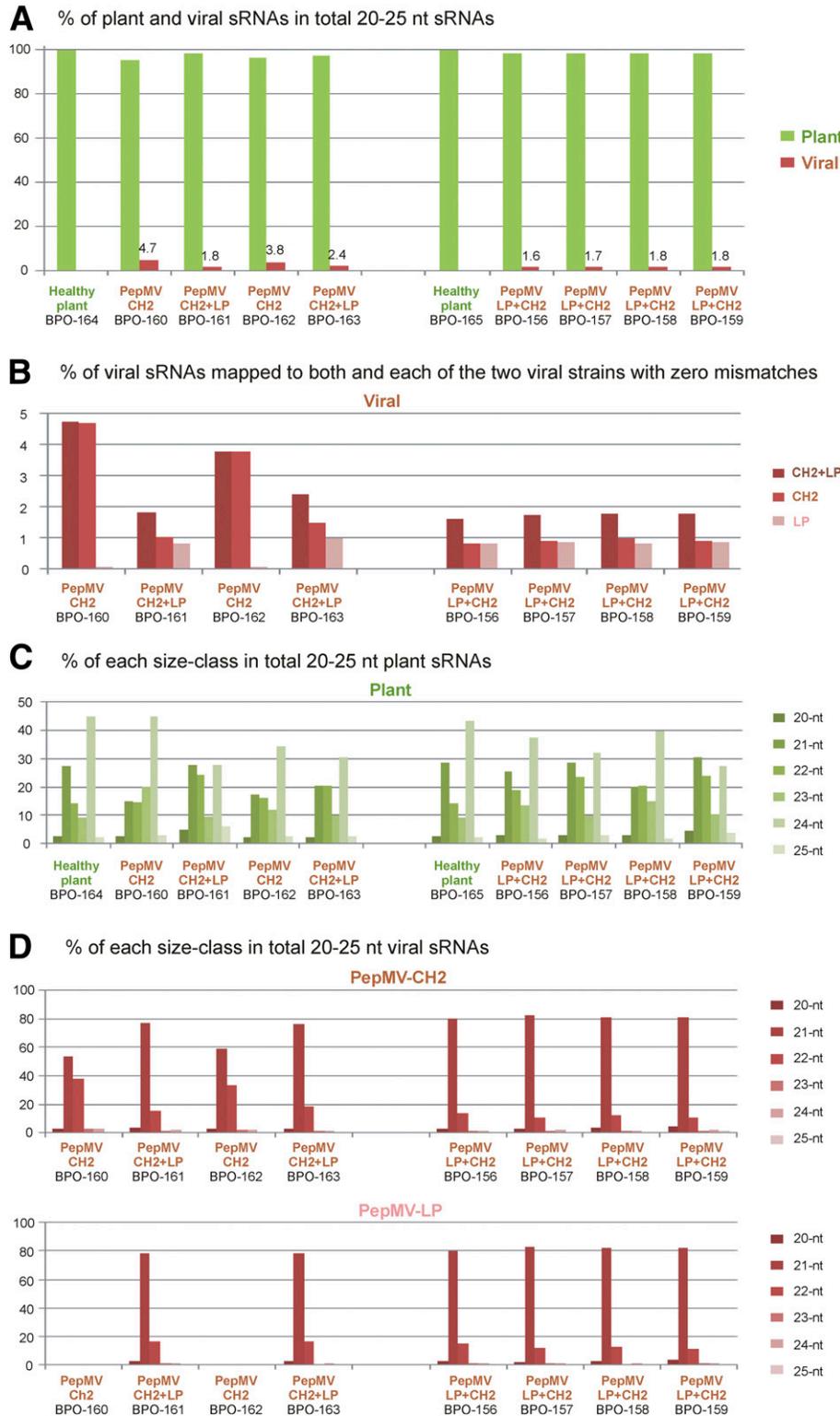


Fig. 5. Illumina sequencing counts of endogenous and viral small RNAs (sRNAs) in *Solanum lycopersicum* plants infected with one (CH2) or two (CH2+LP) strains of *Pepino mosaic virus* (PepMV). The 20- to 25-nt sRNA libraries from two healthy (BPO-164 and BPO-165), two CH2-preinfected (BPO-160 and BPO-161 or BPO-162 and BPO-163) and two LP-preinfected (BPO-156 and BPO-157 or BPO-158 and BPO-159) plants were mapped to the combined (CH2+LP) virome or individual viral genome sequences and the plant genome reference sequence with zero mismatches and were counted. **A**, The percentage (%) of virome- and plant-derived sRNAs in the pool of total 20- to 25-nt reads. **B**, The percentage of each individual virus-derived sRNA in the combined virome sRNAs. **C** and **D**, The percentage of each size class in the 20- to 25-nt pool of plant-derived (**C**) or each virus-derived (**D**) sRNA read.

tomato plants preinfected with CH2 at both timepoints. Interestingly, STV was not detected in any of the Mt. Favet plants (BPO-160 to BPO-163 versus BPO-156 to BPO-159), suggesting that it had persisted in the Merlice plants and was not transmitted to the Mt. Favet plants during their cocultivation. STV belongs to

the family *Amalgaviridae*, has a dsRNA genome of approximately 3.4 kb with two overlapping ORFs, and is transmitted vertically, via seeds, but not mechanically (Sabanadzovic et al. 2009). The latter observation is consistent with our finding that STV was not cotransmitted with the sap together with PepMV.

MISIS-generated maps of viral sRNAs aligned to PepMV CH2 and LP consensus genome sequences with zero mismatches

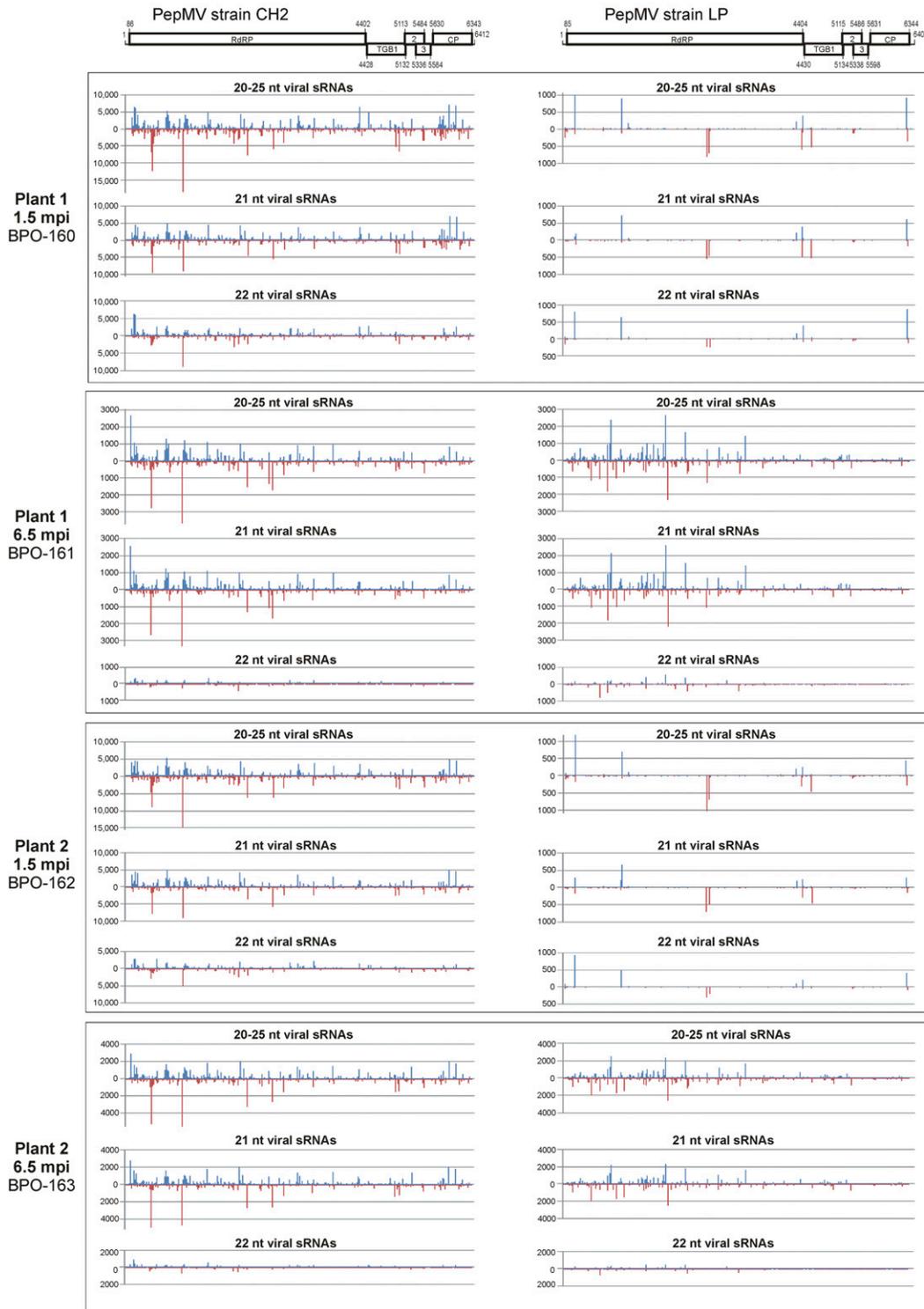


Fig. 6. Single-nucleotide resolution maps of viral small RNAs (sRNAs) from two tomato plants preinfected with *Pepino mosaic virus* (PepMV) CH2 and challenge-inoculated with PepMV LP. For each plant sample at the two timepoints, 1.5 months postinoculation (mpi) (BPO-160 and BPO-162) and 6.5 mpi (BPO-161 and BPO-163), the histograms plot the numbers of total 20- to 25-nt, 21-nt, or 22-nt viral sRNA reads at each nucleotide position of 6412 nt CH2 and 6408 nt LP genome sequences (mapped with zero mismatches). The bars above the axis represent sense reads starting at each position and those below represent antisense reads ending at the respective position. A scaled CH2 and LP genome diagrams are shown above the histograms, with the open reading frames boxed and their encoded proteins and nucleotide positions indicated. Note that the plants are not yet infected with LP at 1.5 mpi.

MISIS-generated maps of viral sRNAs aligned to PepMV CH2 and LP consensus genome sequences with zero mismatches

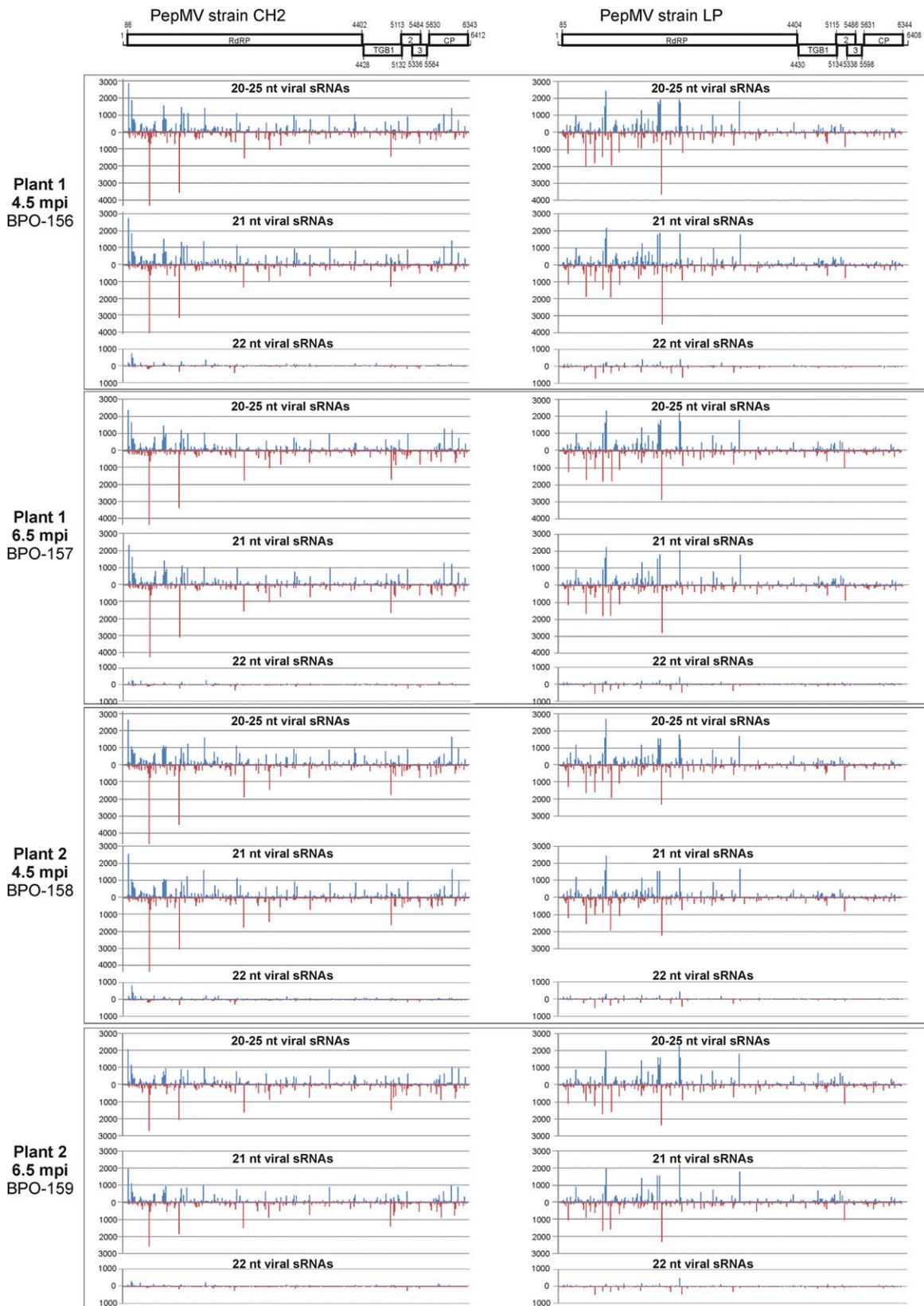


Fig. 7. Single-nucleotide resolution maps of viral small RNAs (sRNAs) from two tomato plants preinfected with *Pepino mosaic virus* (PepMV) LP and challenge-inoculated with PepMV CH2. For each plant sample at the two timepoints, 4.5 months postinoculation (mpi) (BPO-156 and BPO-158) and 6.5 mpi (BPO-157 and BPO-159), the histograms plot the numbers of total 20- to 25-nt, 21-nt, or 22-nt viral sRNA reads at each nucleotide position of the 6,412-nt CH2 and 6,408-nt LP genome sequences (mapped with zero mismatches). The bars above the axis represent sense reads starting at each position and those below represent antisense reads ending at the respective position. Scaled CH2 and LP genome diagrams are shown above the histograms, with the open reading frames boxed and their encoded proteins and nucleotide positions indicated.

The consensus sequences of our STV isolates from two plants differ at one nucleotide position of the 3,438-nt genome G1490 in plant 1 (BPO-160/BPO-161; deposited to GenBank as MF422617) and A1490 in plant 2 (BPO-162/BPO-163; deposited to GenBank as MF422618). Our isolates are 99.3% identical (24 and 23 SNPs, respectively) to the STV isolate CN-12 (KT438549) identified in

tomato plants in China in 2012 (Padmanabhan et al. 2015). Interestingly, the Chinese isolate CN-12 has exhibited severe disease symptoms, while our Swiss genetic variants of STV appeared to be symptomless.

Accumulation of STV-derived sRNAs was very low in both plants at the first timepoint in the presence of CH2, while it was

MISIS-generated maps of viral sRNAs aligned to STV-161 and STV-163 consensus genome sequences with zero mismatches

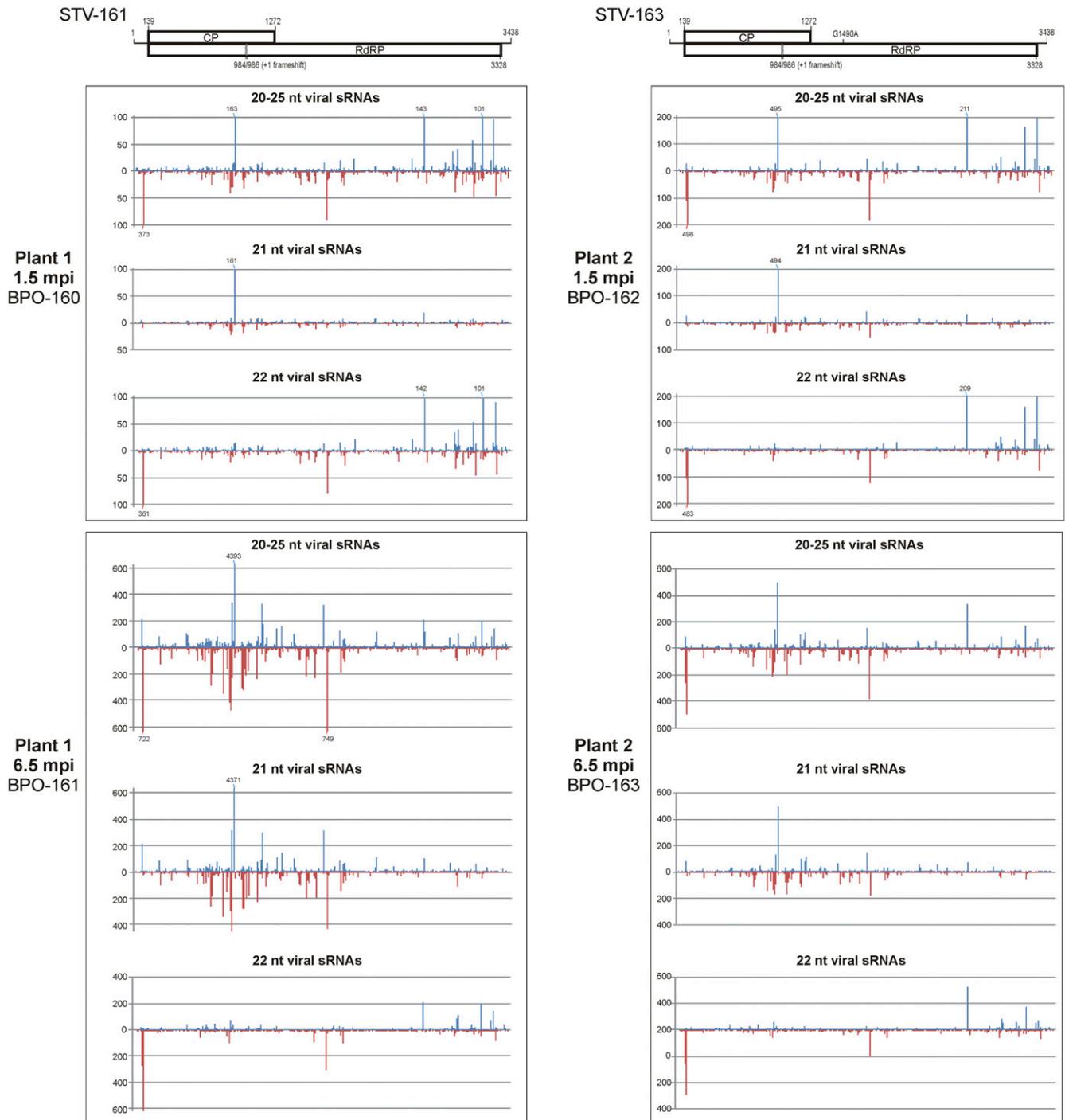


Fig. 8. Single-nucleotide resolution maps of contaminant *Southern tomato virus* (STV)-derived small RNAs (sRNAs) from two tomato plants preinfected with *Pepino mosaic virus* (PepMV) CH2 and challenge-inoculated with PepMV LP. For each plant sample at the two timepoints, 1.5 months postinoculation (mpi) (BPO-160 and BPO-162) and 6.5 mpi (BPO-161 and BPO-163), the histograms plot the numbers of total 20- to 25-nt, 21-nt, or 22-nt viral sRNA reads at each nucleotide position of two isolates (STV-161 and STV-163, differing by one single nucleotide polymorphism) of the 3,438-nt STV genome sequence (mapped with zero mismatches). The bars above the axis represent sense reads starting at each position and those below represent antisense reads ending at the respective position. Scaled STV-161 and STV-163 genome diagrams are shown above the histograms, with the open reading frames boxed and their encoded proteins and nucleotide positions indicated. Note that the plants are not yet co-infected with LP at 1.5 mpi.

strongly elevated at the second timepoint in the presence of CH2 and LP. This suggests that LP co-infection might have activated replication of STV, which would, in turn, boost the silencing response to STV. Moreover, the size profile of STV sRNAs was altered dramatically in that the 21-nt class, being the second most abundant in the presence of CH2 alone, became the first most abundant, outnumbering the 22-nt class, in the presence of both CH2 and LP (Fig. 8; Supplementary Dataset S6). Taking into consideration the negative effect of LP on production of 22-nt sRNAs from CH2 (as described above), LP (but not CH2) appears to directly interfere with DCL2 activity generating 22-nt viral sRNAs, which would indirectly increase the proportion of 21-nt siRNAs generated by DCL4 from STV and CH2.

Concluding remarks.

The sRNA-omics approach has been successfully used for detection and de novo reconstruction of consensus genome sequences of plant RNA and DNA viruses (Kreuze et al. 2009; Seguin et al. 2014). Notably, viral sRNA population from a potato plant experimentally infected with PVY could faithfully reconstitute both the consensus sequence and the SNP profile of viral genomic RNA molecules packaged into virions (Kutnjak et al. 2015). Here, we demonstrate that the consensus genome sequences of both PVY and PVX stably persisting in three potato plants grown from a co-infected tuber are identical. Moreover, the positions and percentages of minor nucleotide variants that deviate from the consensus genome sequences supported by more than half of the sRNA reads are remarkably similar between the plants. Taken together, these findings imply that, following a natural co-infection event and eventual coexistence (with possible coevolution) in vegetatively propagated plant tissues, the two viruses have formed a stable virome quasispecies. This is despite the fact that the two components of the virome are differentially targeted by the plant RNA silencing-based defense system, resulting in four- to fivefold more abundant siRNAs derived from PVY, as compared with PVX. By contrast, our findings for PepMV indicate an on-going evolution of the viral quasispecies, with appearance and fixation of SNPs in each infected plant whose leaf tissues were sampled at two different timepoints postinoculation. This may indicate virus adaptation to a new host plant cultivar or new environmental conditions upon mechanical inoculation. In the case of co-infections with LP and CH2 strains of PepMV, the observed microevolution of viral consensus sequences may additionally be driven by the viral sRNAs derived from one strain that can potentially target identical or near-identical genomic sequences of another strain upon its invasion in a singly infected plant. This hypothesis is worthy of more systematic investigation.

The natural co-infection of vegetatively cultivated potato with two distinct strains of the same virus that can form a stable virome is not unprecedented. For example, two strains of PVX have been identified in a field potato sample from Peru (Kutnjak et al. 2014). Interestingly, these strains could be discriminated by sRNA sequencing and near-completely reconstructed from viral sRNAs, using a reference-based approach. Our results for the PVY-A^N-PVY-B^O virome (as well as PepMV LP+CH2 viromes) demonstrate that identical and near-identical sequences shared by the virome components may prevent de novo reconstruction and reliable separation of closely related genetic variants of the same virus. The regions of sequence identity can potentially promote homologous recombination creating recombinant genomes in the evolving virome quasispecies, resulting in appearance of new recombinant strains or new chimeric viruses. Uncovering such events and reconstituting the complexity of mixed virome quasispecies with closely related genetic variants would require high-throughput sequencing of single long nucleic acid molecules.

Our comparative analysis of RNA silencing-based antiviral responses in potato and tomato revealed remarkable similarities in the biogenesis of viral siRNAs that appear to be processed by DCL4 and DCL2 from dsRNA precursors representing the entire virus genome and presumably sorted by at least three AGOs (AGO1, AGO2, and AGO5), which would be consistent with the findings in *Arabidopsis*. Interestingly, the potexviruses PVX and PepMV spawn much less-abundant siRNAs, compared with PVY. The difference in viral siRNA levels could potentially reflect differences in viral titres. Our results for PepMV strains revealed that the higher accumulation of CH2-derived siRNAs in single infections well correlates with the higher accumulation of CH2 gRNA and sgRNAs. Additionally or alternatively, viral siRNA accumulation may depend on the mode of action of a viral silencing-suppressor protein. Thus, the potyviral suppressor HC-Pro is known to bind and sequester viral siRNAs in *Arabidopsis* (Garcia-Ruiz et al. 2015), which correlates with their high abundance, while the potexviral suppressor TGB1 does not appear to possess any sRNA-binding capacity (Csorba et al. 2015). It remains to be investigated whether the mode of action or the relative strength of antisilencing activities of viral suppressor proteins could contribute to coexistence of distinct viruses and viral strains in a stable virome. In particular, interference with the repressive action of viral siRNAs by sequestration or other mechanisms (Csorba et al. 2015) would prevent potential cross-targeting in mixed infections, leading to synergistic interactions. The observed antagonistic interactions between CH2 and LP strains in mixed infections would be consistent with inability of TGB1 to prevent viral siRNA-mediated cross-targeting activity. Nonetheless, these two strains were able to co-invade plants.

Our cross-protection trial of cultivated tomato plants revealed the ability of the commercial protective mild isolate of PepMV strain CH2 (De Nayer et al. 2011; Hanssen et al. 2009, 2010), which has been approved for cross-protection measures in many European countries, under the tradename PMV-01, to coexist in planta with the LP strain of the same virus. The latter was introduced in Europe from Peru, raising concerns because of its potential to become pathogenic through one or both mutation and recombination with other strains. Our findings support these concerns in that the LP quasispecies undergoes fast evolution within an infected plant through appearance and fixation of point mutations. Likewise, the CH2 quasispecies was found to rapidly evolve, which may eventually create aggressive genetic variants with severe disease symptoms. Both the fast evolution of PepMV quasispecies and the formation of stable viromes between PepMV strains highlight a potential risk of using the protective mild strains of this and other viruses as agents for cross-protection. Among other requirements, genetic stability is considered to be important for effective and durable mild strain cross-protection (Ziebell and MacDiarmid 2017).

MATERIALS AND METHODS

Plant growth and treatment conditions.

Potato tubers were purchased in local shops of Basel-city in summer 2014 and planted in pots in a greenhouse at the University of Basel (16 h light at 220 $\mu\text{E}/\text{m}^2/\text{s}$ and 25°C, 8 h dark at 20°C and constant relative humidity of 65%). Four weeks post plantation, upper leaves of the plants exhibiting virus-like disease symptoms and of the control symptomless plants were harvested in pools and used for extraction of total RNA.

Cross-protection trial.

Hundred plantlets of tomato cv. Merlice were planted in a double-stem configuration in stonewool substrate in a production greenhouse of 150 m² with hydroponic cultivation at the Agro-scope outstation in Conthey, Switzerland. They were treated with

a protective isolate of PepMV strain CH2 (manufactured by Scientia Terrae B and traded as PMV-01), according to the manufacturer's recommendations. In parallel in a separate Agroscope greenhouse in Nyon, four plantlets of tomato cv. Mt. Favet were inoculated with PepMV strain LP that had been isolated from a tomato plant of cv. Merlice collected in Champagny, Switzerland in October 2013 and named PE387. These four plants were transferred and disseminated in the production greenhouse about 2 months postinoculation. Pruning and fruit harvest were performed weekly. Upper leaves of two selected Merlice plants preinfected with CH2 and challenge-inoculated with LP were sampled at two timepoints postintroduction of LP as specified above. Likewise, upper leaves of two selected Mt. Favet plants preinfected with LP and challenge-inoculated with CH2 were sampled at two timepoints indicated above.

Total RNA preparation.

Leaf tissue samples (1 g) were ground in liquid nitrogen, 5 ml of GHCL solution (6.5 M guanidine hydrochloride, 100 μ M Tris HCl, pH 8.0, 100 μ M sodium acetate, pH 5.5, 100 μ M β -mercaptoethanol) was added, and the resulting suspension was transferred into a sterile polypropylene tube and was mixed vigorously. After incubation at room temperature for 10 min, the suspension was centrifuged at 12,000 rpm for 10 min at 4°C and the supernatant was then transferred into a new tube. TRI Reagent (Sigma) (2.5 ml) and 1 ml of chloroform were added to the supernatant, the mixture was vortexed thoroughly and the centrifugation was repeated at the same conditions. The aqueous phase was transferred to a new tube and an equal volume of isopropanol was added, followed by vortexing. The mixture was incubated on ice for 30 min and was then centrifuged at 12,000 \times g, 4°C, for 20 min. The pellet was washed in 5 ml of 75% ethanol (prechilled on ice). The RNA was then pelleted at 12,000 \times g, 4°C, for 10 min, was vacuum-dried for 10 min, and was dissolved in diethyl pyrocarbonate (DEPC)-treated water at 65°C for 10 min.

Long and small RNA blot hybridization analyses.

For long RNA analysis, 5 to 10 μ g of vacuum-dried total RNA was resuspended in 5 μ l of DEPC-treated water, was mixed with 5 μ l of RNA gel loading dye (Thermo Fisher Scientific), was incubated at 70°C for 10 min, and was then loaded on formaldehyde-containing agarose gel (1.2% agarose, 3% formaldehyde, 1 \times MOPS buffer [0.02 M MOPS, pH 7.0, 1 mM EDTA, 5 mM NaOAc]). Following gel-electrophoresis for 2.5 h at 100 V, RNA was stained with ethidium bromide for loading control, and was then transferred by capillarity blotting to a Hybond N+ membrane (Amersham) for 24 h in transfer buffer (50 mM NaH₂PO₄, 5 mM EDTA, pH 6.5) and cross-linked to the membrane twice with 1,200 μ Joules \times 100 UV light in a Statalinker 1800 (Stratagene). For sRNA analysis, 10 μ g of potato total RNA and 5 μ g of tomato total RNA were vacuum-dried, were resuspended in 10 μ l of RNA gel loading dye (Thermo Fisher Scientific), and were incubated at 95°C for 3 min. The samples were then loaded on 15% polyacrylamide gel (19:1 acrylamide:bis-acrylamide and 8 M urea) and were run at 3 V for 4 h. RNA was then transferred to a Hybond N+ membrane (Amersham) by electroblotting in 1 \times Tris-borate-EDTA buffer at 10 V overnight at 4°C and was cross-linked to the membrane as described above.

Both types of membranes were taken for hybridization analysis as described by Akbergenov et al. (2006), using DNA oligonucleotide probes complementary to reconstructed viral genome sequences or endogenous plant RNAs (see Supplementary Table S1 for sequences of the DNA oligonucleotide probes and PCR primers and for the reference sequences used for bioinformatic analyses). The probes were end-labeled with P32 gamma ATP (Hartmann Analytic) by T4 polynucleotide kinase (Roche) and were purified through MicroSpin G-25

columns (Amersham). The hybridization was carried out overnight at 35°C in UltraHyb-oligo buffer (Ambion), followed by triple washing with 2 \times SSC (1 \times SSC is 0.15 M NaCl plus 0.015 M sodium citrate), 0.5% sodium dodecyl sulfate (SDS) for 30 min at 35°C. The membranes were exposed for 5 days to a phosphor screen and were scanned in a GE Typhoon 8600 imager (GE Healthcare Life Sciences). For repeated hybridizations, the membrane was stripped with 0.5 \times SSC, 0.5% SDS for 30 min at 80°C, and then, with 0.1 \times SSC, 0.5% SDS for 30 min at 80°C.

Illumina sequencing and bioinformatics analysis of sRNAs.

The 19- to 30-nt fractions of total RNA samples were sequenced at Fasteris, using a TruSeq small RNA library prep kit (Illumina) for cDNA library preparation and multiplexing the potato (BPO-147 to BPO-152) and tomato (BPO-156 to BPO-165) libraries in separate lanes of the Genome Analyzer HiSeq2500 or HiSeq 4000 (HYT-25 and BPO-26). After adapter trimming, sRNA reads were sorted in separated fastq files.

The redundant and nonredundant sRNA reads from 20 to 25 nt in length were assembled into contigs, using Velvet 1.2.10 (Zerbino and Birney 2008), followed by Oases 0.2.09 (Schulz et al. 2012). The Oases contigs assembled with different k-mers values (13, 15, 17, 19, 21) were combined in one single FASTA file and were mapped on the host reference genome (*S. tuberosum*: PGSC_DM_v3_superscaffolds, PGSC_DM_v3_2.1.9_superscaffolds_unanchored, *S. tuberosum*_Group_Phureja_chloroplast_DM1-3-516-R44, *S. tuberosum*_Group_Phureja_mitochondrion_DM1-3-516-R44 [Potato Genome Sequencing Consortium database] or *S. lycopersicum*: GCA_000188115.2.28 [Ensembl Genome database]) using BWA 0.7.12 (Burrows and Wheeler 1994). The unmapped contigs were further assembled using the Seqman module of Lasergene DNASTAR 12.0.0 Core Suite (DNASTar). The Seqman contigs were blasted on the NCBI nucleotide BLAST database (blastn), and the most closely related reference sequences of the viral strains were taken for further analysis. Redundant sRNA reads were mapped to the Seqman contigs and selected references and the resulting BAM files were analyzed by MISIS-2 (Seguin et al. 2016) to visualize sRNA coverage, identify SNPs and indels, and download the consensus sequences and sRNA single-nucleotide resolution maps as described by Seguin et al. (2016). The BAM files were then further analyzed using MISIS-2 and in-house Fasteris scripts to count viral and host plant sRNAs sorted by size (20, 21, 22, 23, 24, 25 nt, total 20- to 25-nt), polarity (forward, reverse, total), and 5'-nucleotide identity (5'A, 5'C, 5'G, 5'U) and to create the count tables. To enable reconstruction of PVY-A^N and PVY-B^O strains present in one virome, Oases contigs were mapped to Seqman contigs (or PVY-A^N and PVY-B^O reference sequences) using BWA and were visualized by IGV (Thorvaldsdóttir et al. 2013).

The bioinformatic analysis was performed at Fasteris or at the sciCORE scientific computing core facility of the University of Basel.

RT-PCR separation of PVY strains.

To validate the presence of two PVY strains in the potato sample BPO-147, we used an RT-PCR-based genome walking approach, with primers designed based on the sRNA contig sequences representing each strain. Total RNA (2 μ g) was incubated with 1 μ l of 10 mM dNTP mix and 2 pmol RT primer (PVY-RT_as) and was heated to 65°C for 5 min. After chilling in ice for 2 min, the RT was performed by adding to the mixture 4 μ l of 5 \times first-strand synthesis buffer (250 mM Tris-HCl [pH 8.3], 375 mM KCl, 15 mM MgCl₂, 0.1 M dithiothreitol [DTT]), 1 μ l of 0.1 M DTT, 1 μ l (40 U) of RNase inhibitor RNasin (Promega),

and 1 µl (200 U) of Superscript III RT (Invitrogen) and incubating at 55°C for 60 min. PCR amplification was performed in 25 µl of 1× PCR buffer, 0.5 µl of 10 mM dNTP mix, 2.5 U of *Taq* polymerase (Sigma), 20 pmol of PVY-RT primer, 20 pmol of strain-specific primer (147PVYO_1s and 147PVYN_1s, respectively), and 2 µl of the RT reaction mixture as template. The PCR amplification was carried out for 35 cycles (each consisting of 30 s at 95°C, 30 s at 60°C, and 60 s at 72°C) and the products were analyzed by 1% agarose gel electrophoresis. The PCR bands were extracted from the gel with the GenElute gel extraction kit (Sigma) and were sequenced at FASTERIS. Following this protocol, a genome walking was performed by designing antisense primers on the sequenced region, to amplify the corresponding cDNA together with a strain-specific primer.

ACKNOWLEDGMENTS

We thank N. Dubuis and J. Brodard for technical assistance in handling the tomato plants.

LITERATURE CITED

- Akbergenov, R., Si-Ammour, A., Blevins, T., Amin, I., Kutter, C., Vanderschuren, H., Zhang, P., Gruissem, W., Meins, F., Jr., Hohn, T., and Pooggin, M. M. 2006. Molecular characterization of geminivirus-derived small RNAs in different plant species. *Nucleic Acids Res.* 34: 462-471.
- Aregger, M., Borah, B. K., Seguin, J., Rajeswaran, R., Gubaeva, E. G., Zvereva, A. S., Windels, D., Vazquez, F., Blevins, T., Farinelli, L., and Pooggin, M. M. 2012. Primary and secondary siRNAs in geminivirus-induced gene silencing. *PLoS Pathog.* 8:e1002941.
- Bai, M., Yang, G. S., Chen, W. T., Mao, Z. C., Kang, H. X., Chen, G. H., Yang, Y. H., and Xie, B. Y. 2012. Genome-wide identification of Dicer-like, Argonaute and RNA-dependent RNA polymerase gene families and their expression analyses in response to viral infection and abiotic stresses in *Solanum lycopersicum*. *Gene* 501:52-62.
- Blevins, T., Rajeswaran, R., Aregger, M., Borah, B. K., Schepetilnikov, M., Baerlocher, L., Farinelli, L., Meins, F., Jr., Hohn, T., and Pooggin, M. M. 2011. Massive production of small RNAs from a non-coding region of *Cauliflower mosaic virus* in plant defense and viral counter-defense. *Nucleic Acids Res.* 39:5003-5014.
- Blevins, T., Rajeswaran, R., Shivaprasad, P. V., Beknazariants, D., Si-Ammour, A., Park, H. S., Vazquez, F., Robertson, D., Meins, F., Jr., Hohn, T., and Pooggin, M. M. 2006. Four plant Dicers mediate viral small RNA biogenesis and DNA virus induced silencing. *Nucleic Acids Res.* 34:6233-6246.
- Borges, F., and Martienssen, R. A. 2015. The expanding world of small RNAs in plants. *Nat. Rev. Mol. Cell Biol.* 16:727-741.
- Bouché, N., Lauressergues, D., Gascioli, V., and Vaucheret, H. 2006. An antagonistic function for *Arabidopsis* DCL2 in development and a new function for DCL4 in generating viral siRNAs. *EMBO J.* 25:3347-3356.
- Burrows, M., and Wheeler, D. J. 1994:Pages 1-24 in: A Block-Sorting Lossless Data Compression Algorithm. Digital Equipment Corporation Press, Palo Alto, CA, U.S.A..
- Carbonell, A., and Carrington, J. C. 2015. Antiviral roles of plant ARGONAUTES. *Curr. Opin. Plant Biol.* 27:111-117.
- Csorba, T., Kontra, L., and Burguán, J. 2015. Viral silencing suppressors: Tools forged to fine-tune host-pathogen coexistence. *Virology* 479-480: 85-103.
- De Nayer, F., Goen, K., Paeleman, A., Parra, N. O., Vanachter, A. C. R. C., Hanssen, I. M., Wittemans, L., and Vandewoestijne, E. 2011. Cross-protection as a control strategy for *Pepino mosaic virus* (PepMV) in greenhouse tomato. *Acta Hort.* 914:163-169.
- Fang, X., and Qi, Y. 2016. RNAi in Plants: An Argonaute-centered view. *Plant Cell* 28:272-285.
- Fuentes, A., Carlos, N., Ruiz, Y., Callard, D., Sánchez, Y., Ochagavía, M. E., Seguin, J., Malpica-López, N., Hohn, T., Lecca, M. R., Pérez, R., Doreste, V., Rehrauer, H., Farinelli, L., Pujol, M., and Pooggin, M. M. 2016. Field trial and molecular characterization of RNAi-transgenic tomato plants that exhibit resistance to Tomato yellow leaf curl geminivirus. *Mol. Plant-Microbe Interact.* 29:197-209.
- Fusaro, A. F., Matthew, L., Smith, N. A., Curtin, S. J., Dedic-Hagan, J., Ellacott, G. A., Watson, J. M., Wang, M. B., Brosnan, C., Carroll, B. J., and Waterhouse, P. M. 2006. RNA interference-inducing hairpin RNAs in plants act through the viral defence pathway. *EMBO Rep.* 7: 1168-1175.
- García-Ruiz, H., Carbonell, A., Hoyer, J. S., Fahlgren, N., Gilbert, K. B., Takeda, A., Giampetruzzi, A., García Ruiz, M. T., McGinn, M. G., Lowery, N., Martínez Baladejo, M. T., and Carrington, J. C. 2015. Roles and programming of *Arabidopsis* ARGONAUTE proteins during *Turnip mosaic virus* infection. *PLoS Pathog.* 11:e1004755.
- Ghildiyal, M., and Zamore, P. D. 2009. Small silencing RNAs: An expanding universe. *Nat. Rev. Genet.* 10:94-108.
- Gómez, P., Sempere, R. N., Elena, S. F., and Aranda, M. A. 2009. Mixed infections of *Pepino mosaic virus* strains modulate the evolutionary dynamics of this emergent virus. *J. Virol.* 83:12378-12387.
- Hanssen, I. M., Gutierrez-Aguirre, I., Paeleman, A., Goen, K., Wittemans, L., Lievens, B., Vanachter, A. C. R. C., Ravnika, M., and Thomma, B. P. H. J. 2010. Cross-protection or enhanced symptom display in greenhouse tomato co-infected with different *Pepino mosaic virus* isolates. *Plant Pathol.* 59:13-21.
- Hanssen, I. M., Paeleman, A., Vandewoestijne, E., Van Bergen, L., Bragard, C., Lievens, B., Vanachter, A. C. R. C., and Thomma, B. P. H. J. 2009. *Pepino mosaic virus* isolates and differential symptomatology in tomato. *Plant Pathol.* 58:450-460.
- Hasiów-Jaroszewska, B., Borodynko, N., Jackowiak, P., Figlerowicz, M., and Pospieszny, H. 2011. Single mutation converts mild pathotype of the *Pepino mosaic virus* into necrotic one. *Virus Res.* 159:57-61.
- Hasiów-Jaroszewska, B., Jackowiak, P., Borodynko, N., Figlerowicz, M., and Pospieszny, H. 2010. Quasispecies nature of *Pepino mosaic virus* and its evolutionary dynamics. *Virus Genes* 41:260-267.
- Hasiów-Jaroszewska, B., Paeleman, A., Ortega-Parra, N., Borodynko, N., Minicka, J., Czerwoniec, A., Thomma, B. P., and Hanssen, I. M. 2013. Ratio of mutated versus wild-type coat protein sequences in *Pepino mosaic virus* determines the nature and severity of yellowing symptoms on tomato plants. *Mol. Plant Pathol.* 14:923-933.
- Havecker, E. R., Wallbridge, L. M., Hardcastle, T. J., Bush, M. S., Kelly, K. A., Dunn, R. M., Schwach, F., Doonan, J. H., and Baulcombe, D. C. 2010. The *Arabidopsis* RNA-directed DNA methylation Argonautes functionally diverge based on their expression and interaction with target loci. *Plant Cell* 22:321-334.
- Kravchik, M., Damodharan, S., Stav, R., and Arazi, T. 2014a. Generation and characterization of a tomato DCL3-silencing mutant. *Plant Sci.* 221-222:81-89.
- Kravchik, M., Sunkar, R., Damodharan, S., Stav, R., Zohar, M., Isaacson, T., and Arazi, T. 2014b. Global and local perturbation of the tomato microRNA pathway by a trans-activated DICER-LIKE 1 mutant. *J. Exp. Bot.* 65:725-739.
- Kreuze, J. F., Perez, A., Untiveros, M., Quispe, D., Fuentes, S., Barker, I., and Simon, R. 2009. Complete viral genome sequence and discovery of novel viruses by deep sequencing of small RNAs: A generic method for diagnosis, discovery and sequencing of viruses. *Virology* 388:1-7.
- Kutnjak, D., Rupar, M., Gutierrez-Aguirre, I., Curk, T., Kreuze, J. F., and Ravnika, M. 2015. Deep sequencing of virus-derived small interfering RNAs and RNA from viral particles shows highly similar mutational landscapes of a plant virus population. *J. Virol.* 89:4760-4769.
- Kutnjak, D., Silvestre, R., Cuellar, W., Perez, W., Müller, G., Ravnika, M., and Kreuze, J. 2014. Complete genome sequences of new divergent potato virus X isolates and discrimination between strains in a mixed infection using small RNAs sequencing approach. *Virus Res.* 191:45-50.
- Li, R., Gao, S., Hernandez, A. G., Wechter, W. P., Fei, Z., and Ling, K. S. 2012. Deep sequencing of small RNAs in tomato for virus and viroid identification and strain differentiation. *PLoS One* 7:e37127.
- López, C., Soler, S., and Nuez, F. 2005. Comparison of the complete sequences of three different isolates of *Pepino mosaic virus*: Size variability of the TGBp3 protein between tomato and *L. peruvianum* isolates. *Arch. Virol.* 150:619-627.
- Malpica-López, N., Rajeswaran, R., Beknazariants, D., Seguin, J., Golyaev, V., Farinelli, L., and Pooggin, M. M. 2018. Revisiting the roles of tobamovirus replicase complex proteins in viral replication and silencing suppression. *Mol. Plant Microbe Interact.* 31:125-144.
- Naveed, K., Mitter, N., Harper, A., Dhingra, A., and Pappu, H. R. 2014. Comparative analysis of virus-specific small RNA profiles of three biologically distinct strains of Potato virus Y in infected potato (*Solanum tuberosum*) cv. Russet Burbank. *Virus Res.* 191:153-160.
- Padmanabhan, C., Zheng, Y., Li, R., Sun, S. E., Zhang, D., Liu, Y., Fei, Z., and Ling, K. S. 2015. Complete genome sequence of Southern tomato virus identified in China using next-generation sequencing. *Genome Announc.* 3:e01226-15.
- Pagán, I., Del Carmen Córdoba-Sellés, M., Martínez-Priego, L., Fraile, A., Malpica, J. M., Jordá, C., and García-Arenal, F. 2006. Genetic structure

- of the population of *Pepino mosaic virus* infecting tomato crops in Spain. *Phytopathology* 96:274-279.
- Pooggin, M. M. 2013. How can plant DNA viruses evade siRNA-directed DNA methylation and silencing? *Int. J. Mol. Sci.* 14:15233-15259.
- Pooggin, M. M. 2016. Role of small RNAs in virus host interaction. Pages 161-189 in: *Plant-Virus Interactions—Molecular Biology, Intra- and Intercellular Transport*. T. Klejnow, ed. Springer International Publishing, Cham, Switzerland.
- Rajeswaran, R., Golyaev, V., Seguin, J., Zvereva, A. S., Farinelli, L., and Pooggin, M. M. 2014a. Interactions of rice tungro bacilliform pararetrovirus and its protein P4 with plant RNA-silencing machinery. *Mol. Plant-Microbe Interact.* 27:1370-1378.
- Rajeswaran, R., Seguin, J., Chabannes, M., Duroy, P. O., Laboureau, N., Farinelli, L., Iskra-Caruana, M. L., and Pooggin, M. M. 2014b. Evasion of short interfering RNA-directed antiviral silencing in *Musa acuminata* persistently infected with six distinct banana streak pararetroviruses. *J. Virol.* 88:11516-11528.
- Ratcliff, F. G., MacFarlane, S. A., and Baulcombe, D. C. 1999. Gene silencing without DNA. rna-mediated cross-protection between viruses. *Plant Cell* 11:1207-1216.
- Sabanadzovic, S., Valverde, R. A., Brown, J. K., Martin, R. R., and Tzanetakis, I. E. 2009. Southern tomato virus: The link between the families Totiviridae and Partitiviridae. *Virus Res.* 140:130-137.
- Schuck, J., Gursinsky, T., Pantaleo, V., Burguán, J., and Behrens, S. E. 2013. AGO/RISC-mediated antiviral RNA silencing in a plant in vitro system. *Nucleic Acids Res.* 41:5090-5103.
- Song, X., Li, P., Zhai, J., Zhou, M., Ma, L., Liu, B., Jeong, D. H., Nakano, M., Cao, S., Liu, C., Chu, C., Wang, X. J., Green, P. J., Meyers, B. C., and Cao, X. 2012. Roles of DCL4 and DCL3b in rice phased small RNA biogenesis. *Plant J.* 69:462-474.
- Schulz, M. H., Zerbino, D. R., Vingron, M., and Birney, E. 2012. Oases: Robust de novo RNA-seq assembly across the dynamic range of expression levels. *Bioinformatics* 28:1086-1092.
- Seguin, J., Otten, P., Baerlocher, L., Farinelli, L., and Pooggin, M. M. 2016. MISIS-2: A bioinformatics tool for in-depth analysis of small RNAs and representation of consensus master genome in viral quasispecies. *J. Virol. Methods* 233:37-40.
- Seguin, J., Rajeswaran, R., Malpica-López, N., Martin, R. R., Kasschau, K., Dolja, V. V., Otten, P., Farinelli, L., and Pooggin, M. M. 2014. De novo reconstruction of consensus master genomes of plant RNA and DNA viruses from siRNAs. *PLoS One* 9:e88513.
- Syller, J., and Grupa, A. 2014. The effects of co-infection by different *Potato virus Y* (PVY) isolates on virus concentration in solanaceous hosts and efficiency of transmission. *Plant Pathol.* 63:466-475.
- Syller, J., and Grupa, A. 2016. Antagonistic within-host interactions between plant viruses: Molecular basis and impact on viral and host fitness. *Mol. Plant Pathol.* 17:769-782.
- Takeda, A., Iwasaki, S., Watanabe, T., Utsumi, M., and Watanabe, Y. 2008. The mechanism selecting the guide strand from small RNA duplexes is different among Argonaute proteins. *Plant Cell Physiol.* 49:493-500.
- Thorvaldsdóttir, H., Robinson, J. T., and Mesirov, J. P. 2013. Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. *Brief. Bioinform.* 14:178-192.
- Vanitharani, R., Chellappan, P., Pita, J. S., and Fauquet, C. M. 2004. Differential roles of AC2 and AC4 of cassava geminiviruses in mediating synergism and suppression of posttranscriptional gene silencing. *J. Virol.* 78:9487-9498.
- Wang, X. B., Jovel, J., Udornporn, P., Wang, Y., Wu, Q., Li, W. X., Gascioli, V., Vaucheret, H., and Ding, S. W. 2011. The 21-nucleotide, but not 22-nucleotide, viral secondary small interfering RNAs direct potent antiviral defense by two cooperative Argonautes in *Arabidopsis thaliana*. *Plant Cell* 23:1625-1638.
- Wang, X. B., Wu, Q., Ito, T., Cillo, F., Li, W. X., Chen, X., Yu, J. L., and Ding, S. W. 2010. RNAi-mediated viral immunity requires amplification of virus-derived siRNAs in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* 107:484-489.
- Wu, L., Zhou, H., Zhang, Q., Zhang, J., Ni, F., Liu, C., and Qi, Y. 2010. DNA methylation mediated by a microRNA pathway. *Mol. Cell* 38:465-475.
- Xie, Z., Johansen, L. K., Gustafson, A. M., Kasschau, K. D., Lellis, A. D., Zilberman, D., Jacobsen, S. E., and Carrington, J. C. 2004. Genetic and functional diversification of small RNA pathways in plants. *PLoS Biol.* 2:e104.
- Yifhar, T., Pekker, I., Peled, D., Friedlander, G., Pistunov, A., Sabban, M., Wachsman, G., Alvarez, J. P., Amsellem, Z., and Eshed, Y. 2012. Failure of the tomato *trans*-acting short interfering RNA program to regulate AUXIN RESPONSE FACTOR3 and ARF4 underlies the wiry leaf syndrome. *Plant Cell* 24:3575-3589.
- Zerbino, D. R., and Birney, E. 2008. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* 18:821-829.
- Ziebell, H., and Carr, J. P. 2010. Cross-protection: A century of mystery. *Adv. Virus Res.* 76:211-264.
- Ziebell, H., and MacDiarmid, R. 2017. Prospects for engineering and improvement of cross-protective virus strains. *Curr. Opin. Virol.* 26:8-14.

AUTHOR-RECOMMENDED INTERNET RESOURCES

- Burrows-Wheeler Aligner (BWA 0.7.12): <http://bio-bwa.sourceforge.net>
- Ensembl Genome dataset: ftp://ftp.ensemblgenomes.org/pub/plants/release-28/fast/solanum_lycopersicum/dna/
- Fasteris website: www.fasteris.com
- Integrated Genomics Viewer (IGV) visualization tool: <http://software.broadinstitute.org/software/igv>
- Oases 0.2.09 de novo transcriptome assembler: <https://www.ebi.ac.uk/~zerbino/oases>
- Potato Genome Sequencing Consortium database: http://solanaceae.plantbiology.msu.edu/pgsc_download.shtml
- SIB Swiss Institute of Bioinformatics sciCORE website: <https://scicore.unibas.ch>
- Velvet 1.2.10 sequence assembler: <https://www.ebi.ac.uk/~zerbino/velvet>