Citrus tristeza virus infection induces the accumulation of viral small RNAs (21- 24-nt) mapping preferentially at the 3'-terminal region of the genomic RNA and affects the host small RNA profile

Susana Ruiz-Ruiz[1,4], Beatriz Navarro[2], Andreas Gisel[3], Leandro Peña[4], Luis Navarro[4], Pedro Moreno[4], Francesco Di Serio[2], Ricardo Flores[1]

[1]Instituto de Biología Molecular y Celular de Plantas (Universidad Politécnica de Valencia-Consejo Superior de Investigaciones Científicas), Avenida de los Naranjos, 46022 Valencia, Spain

[2]Istituto di Virologia Vegetale (Consiglio Nazionale delle Ricerche), Via Amendola 165/ A, 70126 Bari, Italy

[3]Istituto di Tecnologie Biomediche (Consiglio Nazionale delle Ricerche), Via Amendola 122/ D, 70126 Bari, Italy

[4]Instituto Valenciano de Investigaciones Agrarias, 46113 Moncada, Valencia, Spain

Author for correspondence:
Ricardo Flores
E-mail: rflores@ibmcp.upv.es
Phone: 34-963877861
Fax: 34-963877859

The two first authors have contributed equally to this work

Keywords: closteroviruses, microRNAs, RNA silencing, small interfering RNAs

Number of words: 6785 (excluding Abstract and Legends to Figures)

Number of Figures: 9

Number of Supplementary Figures: 3

Number of Supplementary Tables: 3

Abstract

To get an insight into the host RNA silencing defense induced by *Citrus tristeza virus* (CTV) and into the counter defensive reaction mediated by its three silencing suppressors (p25, p20 and p23), we have examined by deep sequencing (Solexa-Illumina) the small RNAs (sRNAs) in three virus-host combinations. Our data show that CTV sRNAs: i) represent more than 50% of the total sRNAs in Mexican lime and sweet orange (where CTV reaches relatively high titers), but only 3.5% in sour orange (where the CTV titer is significantly lower), ii) are predominantly of 21- 22-nt, with a biased distribution of their 5' nucleotide and with those of (+) polarity accumulating in a moderate excess, and iii) derive from essentially all the CTV genome (ca. 20 kb), as revealed by its complete reconstruction from viral sRNA contigs, but adopt an asymmetric distribution with a prominent hotspot covering approximately the 3'-terminal 2500 nt. These results suggest that the citrus homologues of Dicer-like (DCL) 4 and 2 most likely mediate the genesis of the 21 and 22 nt CTV sRNAs, respectively, and show that both ribonucleases act not only on the genomic RNA but also on the 3' co-terminal subgenomic RNAs and, particularly, on their double-stranded forms. The plant sRNA profile, very similar and dominated by the 24-nt sRNAs in the three mock-inoculated controls, was minimally affected by CTV infection in sour orange, but exhibited a significant reduction of the 24-nt sRNAs in Mexican lime and sweet orange. We have also identified novel citrus miRNAs and determined how CTV influences their accumulation.

Citrus tristeza virus infection induces the accumulation of viral small RNAs (21- 24-nt) mapping preferentially at the 3'-terminal region of the genomic RNA and affects the host small RNA profile

Susana Ruiz-Ruiz[1,4], Beatriz Navarro[2], Andreas Gisel[3], Leandro Peña[4], Luis Navarro[4], Pedro Moreno[4], Francesco Di Serio[2], Ricardo Flores[1]

[1]Instituto de Biología Molecular y Celular de Plantas (Universidad Politécnica de Valencia-Consejo Superior de Investigaciones Científicas), Avenida de los Naranjos, 46022 Valencia, Spain

[2]Istituto di Virologia Vegetale (Consiglio Nazionale delle Ricerche), Via Amendola 165/ A, 70126 Bari, Italy

[3]Istituto di Tecnologie Biomediche (Consiglio Nazionale delle Ricerche), Via Amendola 122/ D, 70126 Bari, Italy

[4]Instituto Valenciano de Investigaciones Agrarias, 46113 Moncada, Valencia, Spain

Author for correspondence:
Ricardo Flores
E-mail: rflores@ibmcp.upv.es
Phone: 34-963877861
Fax: 34-963877859

The two first authors have contributed equally to this work

Keywords: closteroviruses, microRNAs, RNA silencing, small interfering RNAs

Number of words: 6785 (excluding Abstract and Legends to Figures)

Number of Figures: 9

Number of Supplementary Figures: 3

Number of Supplementary Tables: 3

Abstract

To get an insight into the host RNA silencing defense induced by *Citrus tristeza virus* (CTV) and into the counter defensive reaction mediated by its three silencing suppressors (p25, p20 and p23), we have examined by deep sequencing (Solexa-Illumina) the small RNAs (sRNAs) in three virus-host combinations. Our data show that CTV sRNAs: i) represent more than 50% of the total sRNAs in Mexican lime and sweet orange (where CTV reaches relatively high titers), but only 3.5% in sour orange (where the CTV titer is significantly lower), ii) are predominantly of 21- 22-nt, with a biased distribution of their 5' nucleotide and with those of (+) polarity accumulating in a moderate excess, and iii) derive from essentially all the CTV genome (ca. 20 kb), as revealed by its complete reconstruction from viral sRNA contigs, but adopt an asymmetric distribution with a prominent hotspot covering approximately the 3'-terminal 2500 nt. These results suggest that the citrus homologues of Dicer-like (DCL) 4 and 2 most likely mediate the genesis of the 21 and 22 nt CTV sRNAs, respectively, and show that both ribonucleases act not only on the genomic RNA but also on the 3' co-terminal subgenomic RNAs and, particularly, on their double-stranded forms. The plant sRNA profile, very similar and dominated by the 24-nt sRNAs in the three mock-inoculated controls, was minimally affected by CTV infection in sour orange, but exhibited a significant reduction of the 24-nt sRNAs in Mexican lime and sweet orange. We have also identified novel citrus miRNAs and determined how CTV influences their accumulation.

Introduction

Plants, and other higher eukaryotes, have developed an RNA-based antiviral silencing response triggered by double-stranded (ds) and highly-structured single-stranded (ss) RNA that elicits its processing by enzymes of the RNase-III class (DICER-like, DCL) (Molnar et al. 2005; Qi et al. 2005; Moissiard and Voinnet 2006). The resulting primary virus-derived small RNAs (vsRNAs) of 21 to 24 nt are amplified into secondary vsRNAs by host RNA-dependent RNA polymerases (RDRs) (Dalmay et al. 2000) and, together, they load and guide an RNase H-like enzyme (Argonaute, AGO) forming the core of the RNA-induced silencing complex (RISC) (Hamilton and Baulcombe 1999; Vaucheret 2008), against the viral ssRNAs (Omarov et al. 2007; Pantaleo et al. 2007). Viruses have reacted to this defense mechanism by encoding suppressors of RNA silencing in their genomes (Csorba et al. 2009; Ding, 2010). Because RNA silencing also regulates plant development through small RNAs (sRNAs) of endogenous origin, namely microRNAs and small interfering RNAs (miRNAs and siRNAs, respectively), with both pathways overlapping to some extent, developmental defects incited by viruses have been regarded as a side effects of their suppressors converging in both pathways, although other data indicate that this may not necessarily occur in all cases (Díaz-Pendón and Ding 2008).

Citrus tristeza virus (CTV), family *Closteoviridae,* has the largest monopartite plant virus genome, composed by a (+) ssRNA of about 19.3 kb organized in 12 open reading frames (ORFs) potentially encoding at least 17 protein products flanked by 5' and 3' untranslated regions (UTRs) (Bar-Joseph and Dawson 2008; Karasev et al. 1995). The two 5' proximal ORFs are directly translated from the genomic RNA (gRNA) and encode components of the replicase complex, while the 3' proximal ORFs encoding ten proteins are expressed via a nested set of 3' co-terminal subgenomic RNAs (sgRNAs) (Hilf et al. 1995). Three of these proteins (p25, p20 and p23) are silencing suppressors in *Nicotiana* spp. (Lu et al. 2004). Besides, p23, an RNA-binding protein (López et al. 2000) that controls the asymmetrical accumulation of (+) and (-) CTV RNAs (Satyanarayana et al. 2002), is a pathogenicity determinant when expressed ectopically in several *Citrus* spp. (Fagoaga et al. 2005; Ghorbel et al. 2001) and a likely determinant of the seedling yellows syndrome incited by CTV in

sour orange *(Citrus aurantium* L.) and grapefruit *(C. paradisi* Macf.) (Albiach-Martí et al. 2010). In nature CTV infection is restricted to phloem cells of some species of two genera of the family *Rutaceae* (Moreno et al. 2008), with the virus reaching variable titers depending on the strain and the host: the titer in Mexican lime *(C. aurantifolia* (Christm.) Swing.) and sweet orange *(Citrus sinensis* L. Osb.) can be ten-fold higher than that in sour orange (Folimonova et al. 2008). Moreover, the citrus relative trifoliate orange *(Poncirus trifoliata* L. Raf.) is resistant to CTV at the whole plant level, with the resistance being conferred by the single dominant locus *Ctv*, which most likely blocks virus movement because CTV replicates and forms virions in protoplasts of this species (Albiach-Martí et al. 2004). In addition to the gRNA and sgRNAs, CTV propagation occurs with the accumulation of vsRNAs (Fagoaga et al. 2006), but there is no detailed information about their physical properties, relative abundance and distribution along the viral genome, as well as whether CTV infection has any effect on plant sRNAs. Moreover, it is unclear not only for CTV but in general, which are the substrates for antiviral DCLs because (+) ssRNA viruses may produce dsRNAs during replication of the gRNA as well as during transcription of the sgRNAs (Aliyari et al. 2008), and highly-structured regions of the ss (+) gRNA and sgRNAs may also be targeted by DCLs (Molnar et al. 2005; Szittya et al. 2010). Supporting this latter view, the highly-structured leader region of a DNA virus transcript has been reported as the major hotspot of vsRNAs (Moissiard and Voinnet 2006).

To fill this gap we have analyzed by Northern-blot hybridization, deep sequencing and computational approaches the sRNAs from young stem bark of Mexican lime, sweet orange and sour orange infected by a severe CTV isolate, as well as their corresponding mock-inoculated controls. Our data show that the CTV sRNAs of 21- 22-nt are extremely abundant in Mexican lime and sweet orange, wherein the accumulation pattern of plant sRNAs is significantly altered, but only represent a minor fraction in sour orange. Moreover, CTV sRNAs adopt an asymmetric distribution along the viral gRNA, with a major hotspot covering approximately the 3'-terminal 2500 nt. Such a pronounced uneven distribution has not been reported in other RNA viruses, including one member of the same family (Kreuze et al. 2009).

Materials and methods

Plant material and CTV isolate

Samples of young stem bark (3 g) from seedlings mock-inoculated or CTV-infected with the severe isolate T318A (GenBank accession number DQ151548) were collected 3 months post-inoculation (for Mexican lime and sweet orange) or 12 months post-inoculation (for sour orange). Plants were grown in an artificial potting mix (50% peat moss, 50% sand) and maintained in an insect-proof, temperature-controlled greenhouse (26-23ºC/ 18-15ºC day/ night).

RNA extraction, purification, and Northern-blot hybridization

Total nucleic acid preparations were obtained by phenol-chloroform extraction (Ancillo et al. 2007) and clarified with methoxyethanol (Bellamy and Ralph 1968). Aliquots were fractionated by electrophoresis in either denaturing 1% agarose/ formaldehyde gels (for CTV gRNA and sgRNAs) or in 17% polyacrylamide/ urea gels (for sRNAs), stained with ethidium bromide, and transferred to Hybond-N+ membranes (Roche Diagnostics). The membranes were hybridized with digoxigenin-labeled riboprobes, homologous and complementary to the 3'-terminal gene *p23*, obtained by *in vitro* transcription under the control of the promoter of the T7 RNA polymerase. Hybridizations were performed at 68°C (for CTV gRNA and sgRNAs) or 42°C (for sRNAs) in the ULTRAhyb hybridization buffer (Ambion). Equal loading was assessed by UV spectrophotometry, and by the intensity of 4S and U6 RNA bands after electrophoresis on 5% polyacrylamide/ urea gels and ethidium bromide staining or hybridization with a digoxigenin-labeled riboprobe complementary to U6 RNA. The relative abundance of some representative CTV sRNA was independently assessed by hybridization with the 5'-radiolabeled probes CTV-cld (5'-TTAAACTCAGGATAAGCTCTAGTGAGCATCA-3') and CTV-chd (5'-GGAGAACTTCTTTGGTTCACGCATACGTTAAG-3')

complementary to positions 3224-3254 and 19163-19194, respectively, of the T318A isolate (Ruiz-Ruiz et al. 2006). Similarly, the relative abundance of two host miRNAs and one sRNA was also independently assessed by hybridization with their complementary 5'-radiolabeled DNA probes. After overnight hybridization the membranes were washed twice with 2X SSC plus 0.1% SDS for 10 min at room temperature, and twice with 0.1X SSC plus 0.1% SDS for 15 min at 42°C. Membranes hybridized with digoxigenin-labeled probes were revealed with the chemiluminescent substrate CSPD (Roche Diagnostics) and exposure to X-ray film, and those hybridized with radiolabeled probes with a bioimage analyzer (Fujifilm FLA-5100).

Quantitative real-time RT-PCR

The viral load in the bark of each host plant was estimated by a real-time RT-PCR protocol that uses the general primer set PM197F-PM198R and SYBR Green I for specific and reliable quantitative detection of the CTV gRNA (Ruiz-Ruiz et al. 2007). The LightCycler software provided a plot of the fluorescence intensity against the number of cycles, as well as the threshold cycle (Ct) value using the automatic method. The mean Ct and the standard deviation (SD) for each plant sample were calculated. Synthesis of the DNA product of the expected size was confirmed by melting curve analysis and by electrophoresis in a 2% agarose gel, and the number of CTV gRNA copies in each sample was estimated by interpolating the individual Ct values in an external standard curve.

Deep sequencing of CTV and plant sRNAs

The protocol for purifying the sRNAs, adapter ligation, RT-PCR amplification, library purification, and high-throughput DNA sequencing on the Illumina Genome Analyzer EAS269-GAII (FASTERIS SA, Plan-les-Ouates, Switzerland), has been reported elsewhere (Di Serio et al. 2010). Four bar-coded samples, corresponding to mock-inoculated and CTV-infected Mexican lime and sour orange sRNAs, were analyzed in a single read channel. On the other hand, two samples corresponding to

mock-inoculated and CTV-infected sweet orange sRNAs, were analyzed in single channels.

Sequence analysis of sRNAs

The resulting sequences were examined for the presence of the adapters and, after their trimming, they were sorted into separate files according to their length. For further analysis the 18-26 nt reads were pooled and each set was analyzed by BLAST (Altschul et al. 1990) against the nucleotide sequence of the CTV isolate T318A and plant sequences (particularly of citrus and close relatives) deposited in databases. Data were normalized with respect to 18-26 nt reads from CTV-infected sour orange sample. A set of perl scripts was developed to filter, analyze and visualize the mapping data. Sequence assembly of the CTV gRNA from the overlapping CTV sRNAs was performed with the program Velvet (Zerbino and Birney 2008).

Results and discussion

CTV sRNAs (21- 22-nt) constitute a major fraction of the total sRNAs from Mexican lime and sweet orange

Deep sequencing of gel-purified sRNAs generated approximately 14.200.000 reads, 99% of which corresponded to the four bar-coded samples (from mock-inoculated and CTV-infected Mexican lime and sour orange) run in the same channel, with the fraction of each sample representing 24-27% of the total number of reads (Supplementary Table S1). Dissection of the reads from CTV-infected Mexican lime revealed that, within the range of 18-26 nt, 56.3% were CTV sRNAs matching perfectly the parental sequence, while this fraction was reduced to 3.5% in sour orange (Fig. 1). To further confirm these data and discard sequencing artifacts, two additional sRNA samples (from mock-inoculated and CTV-infected sweet orange) were subjected to deep sequencing but in independent channels, generating about 12.850.000 and 9.933.000 reads, respectively, more than 92% of which contained the

bar-coded bases in 5' and the adapter in 3' termini (Supplementary Table S2). Within the range of 18-26 nt, the fraction of reads from the CTV-infected sweet orange that corresponded to CTV sRNAs matching perfectly the parental sequence was 53.8% (Fig. 1), a value very similar to that obtained for Mexican lime.

The quantitative differences observed in the CTV sRNA profiles are most likely a consequence of the higher virus titer in Mexican lime and sweet orange than in sour orange, as estimated by Northern-blot hybridization: the signals generated by the gRNA and the three more abundant sgRNAs (coding for p23, p20 and p25) were clearly less intense in sour orange than in the other two hosts (Fig. 2A).

Estimation of the gRNA levels by quantitative RT-PCR confirmed this view. More specifically, the threshold cycle (Ct) values obtained for Mexican lime, sweet orange and sour orange were 16.79±0.04, 15.09±0.09 and 23.81±0.07, respectively. The amplification products showed a single melting peak with the Tm value characteristic of severe isolates (Ruiz-Ruiz et al. 2007), supporting the amplification specificity. Moreover, gel electrophoresis analysis of the amplification products showed a single band of the expected size with no primer-dimer formation (data not shown). The mean number of CTV gRNA copies, estimated using an external standard, was in the same range for Mexican lime and sweet orange (2.25×10^6 and 6.99×10^6 molecules per ng of total RNA, respectively), but significantly lower for sour orange (2.11×10^4 molecules per ng of total RNA). The mock-inoculated plants did not yield any amplification product and their melting curve profiles did not differ from negative controls (without template or reverse transcriptase).

Considering that virus titers in woody plants are usually low, the high level of vsRNAs found in Mexican lime and sweet orange was unexpected, even in samples from young stem bark where the phloem-restricted CTV accumulates preferentially. This abundance is not an artifact of the amplification step preceding deep sequencing because the Northern-blot hybridization signals generated by the CTV sRNAs were significantly more intense in Mexican lime and sweet orange than in sour orange (Fig. 2B). In contrast, the dominant 21- and 22-nt sRNAs derived from *Sweet potato chlorotic stunt virus* (SPCSV), another member of the family *Closteroviridae,* represent a minor fraction of the total sRNAs extracted from SPCSV-infected sweet potato, although the starting material was leaf tissue (Kreuze et al. 2009).

The CTV sRNAs of 21 and 22 nt were the most prevalent in the infected samples, with those of (+) polarity being in a moderate excess with respect to their (-) counterparts (Fig. 3). Assuming that citrus DCLs behave as their Arabidopsis homologues, DCL4 and DCL2 are the best candidates for generating the 21- and 22-nt vsRNAs, respectively. Although a surrogate defense role against *Turnip crinkle virus* has been proposed for DCL2 with respect to DCL4 in Arabidopsis (Deleris et al. 2006; Qu et al. 2008), recent data indicate that the effect of DCL4 could be indirect and that DCL2 has a prominent contribution (Azevedo et al. 2010). DCL2 also seems particularly relevant in infected Mexican lime and sweet orange, considering that the CTV sRNAs of 22 nt are more abundant than their 21-nt counterparts (Fig. 1 and 2); the high levels of the vsRNAs of 21 and 22 nt, however, were not accompanied by a similar decrease in the levels of the plant sRNAs of the same size (Fig. 1).

CTV sRNAs map preferentially at the 3'-terminal region of the gRNA

Mapping the CTV sRNAs of both polarities revealed in the three citrus species extensive targeting of the gRNA, but with a clear asymmetric distribution resulting in a major hotspot that covers approximately the 3'-terminal 2500 nt (starting at *p25* and gradually increasing up to *p23*); however, the hotspot did not extend into the 3'-UTR (Fig. 4). Furthermore, the hotspot profiles for the 21- and 22-nt (+) CTV sRNAs were very similar, as also were those of the 21- and 22-nt (-) CTV sRNAs (data not shown), indicating that specific regions of the RNA substrates are targeted by the same DCLs and their auxiliary dsRNA-binding proteins (Curtin et al. 2008); a comparable situation has been recently reported for some representative plant viruses (Donaire et al. 2009) and viroids (Di Serio et al. 2009; Di Serio et al. 2010). To get additional independent confirmation of our deep sequencing data, ruling out the possibility that the differences observed in the CTV sRNA profiles could result from amplification artifacts, we performed two Northern-blot hybridization controls. First, we assessed the relative abundance of (+) vsRNAs along the virus genome by hybridizing aliquots of the RNA preparation from CTV-infected sweet orange with eight equalized ≅1500-nt digoxigenin-labeled riboprobes spanning the complete viral genome; as expected, the strongest signals resulted from the 3'-terminal 1500 nt and

then from the 5'-terminal 1500 nt (Fig. 4, lower panel). And second, we estimated the relative abundance of (+) and (-) vsRNAs derived from specific regions of the CTV genome. The digoxigenin-labeled (+) and (-) riboprobes derived from gene *p23* generated intense signals in the positions expected for the 21- and 22-nt vsRNAs of both polarities in CTV-infected sweet orange and Mexican lime, but only weak signals in CTV-infected sour orange and no signal at all in the mock-inoculated control (Fig. 5A). Moreover, while hybridization with a 5' radioactively-labeled oligonucleotide (CTV-cld) complementary to genomic positions 3224-3254 (with a low density of vsRNAs) failed to provide any signal, hybridization with a 5' radioactively-labeled oligonucleotide (CTV-chd) complementary to genomic positions 19163-19194 (with a relatively high density of vsRNAs) generated clear signals corresponding to 21- and 22-nt vsRNAs only in CTV-infected sweet orange and Mexican lime (Fig. 5B).

To further dissect the CTV sRNA profiles, we analyzed the density of vsRNAs along the ORFs: they showed the same pattern in the three hosts, peaking at *p23* and then declining at the 3'-UTR (Fig. 6). Widespread and efficient targeting of plant viral genomes has been observed previously (Donaire et al. 2009), with the vsRNAs derived from the 3'-terminal region of the largest RNA of *Tobacco rattle virus* (encompassing two genes that are expressed via two sgRNAs) being somewhat more abundant (Donaire et al. 2008). However, the pronounced asymmetric accumulation of CTV sRNAs derived from the 3'-moiety of the CTV gRNA is unprecedented, even in one member (SPCSV) of the same family (Kreuze et al. 2009).

The most plausible explanation for this asymmetric distribution is that the ten 3'-co-terminal CTV sgRNAs —and particularly their dsRNAs detected in extracts of infected tissues (Aramburu et al. 1991; Bar-Joseph and Dawson 2008; Moreno et al. 1990) and presumably accumulating *in vivo* at least transiently as reported for some animal viruses (Weber et al. 2006)— are substrates for the citrus homologues of DCL2 and DCL4. The moderate excess of (+) over (-) CTV sRNAs basically results from the regions coding for p20 and p23 (Fig. 4 and 6) and most likely reflects: i) the excess of (+) over (-) strands in infected cells (Satyanarayana et al. 2002), and ii) the higher accumulation of p20 and p23 sgRNAs detected by Northern-blot analysis in the same RNA preparations used for purifying the sRNAs for deep sequencing (Fig.

2), in line with previous results (Navas-Castillo et al. 1997). We believe that the surplus of (+) CTV sRNAs reflects the *in vivo* situation and not the sequestration of their (-) counterparts by the viral (+) gRNA and sgRNAs (Smith et al. 2010), because the hybridization signals generated by the (-) CTV sRNAs did not change following PAGE under fully denaturing conditions (7 M urea supplemented with 40% formamide) (data not shown). Yet, the relative decrease in the density of vsRNAs derived from the 3'-UTR (273 nt) demands alternative explanations, including protection by proteins involved in transcription of (-) strands or in translation of the sgRNAs that may have special affinity for this region. Supporting this view, very few CTV sRNAs were retrieved from the most 3'-terminal 50 nt (Supplementary Fig. S1). On the other hand, a slight over-accumulation of CTV sRNAs mapping along the 5'-terminal 1000 nt was detected (particularly in sweet orange) (Fig. 4).

CTV sRNAs of 21 and 22 nt display a bias in their 5'-terminal nucleotide

In Mexican lime, the population of the most abundant (+) and (-) CTV sRNAs of 21 and 22 nt was dominated by species with a 5'-terminal C and U (about 40-45 and 30%, respectively), with those having a 5'-terminal G being underrepresented (7-13%). In sour and sweet orange the situation was more balanced, although most 21- and 22-nt vsRNAs with a 5'-terminal G were also underrepresented (15%) (Fig. 7). This disproportionate distribution does not result from the base composition of the CTV gRNA or the three most abundant sgRNAs, which have a G content around 25% (data not shown). On the other hand, some of the ten most abundant CTV sRNAs detected in the three hosts were coincidental, but inspection of their base composition and sequence did not reveal any common trait (data not shown). Potential factors that may shape the CTV sRNA distribution and the lack of a clear phasing include the dissimilar accessibility of regions of the corresponding RNA precursors to DCLs and their auxiliary proteins, the preference of certain AGOs for the 5'-terminal nucleotide of their guide sRNAs (Mi et al. 2008; Montgomery et al. 2008), and their differential targeting by one or more exoribonucleases.

Complete reconstruction of the CTV gRNA from the assembly of the CTV sRNAs

To further assess that most of the CTV gRNA (the largest of a monopartite plant virus) is extensively targeted by DCLs, an attempt was made to reconstruct its sequence from the overlapping CTV sRNAs of the sweet orange library, which we first mapped without mismatches onto the sequence of isolate T318A. Computer-assisted assembly of contigs revealed that the complete CTV genome of 19252 nt could indeed be assembled, with the exception of position 5985 and a short stretch between positions 15739 and 15764. Searching the sweet orange library for reads that with one mismatch could cover these gaps we found that position 5985, annotated as U in the sequence of isolate T318A (GenBank DQ151548), appears as A in 100% of the vsRNAs mapping around this position; database examination revealed that all deposited CTV sequences have an A at this position, strongly suggesting an annotation error. Also, positions 15739 and 15764, annotated as C and G respectively, appear as A in 100% of the corresponding CTV sRNAs. Finally, the U at position 15761 appears as A in 10% of the CTV sRNAs, indicating some sequence heterogeneity. RT-PCR amplification and sequencing of this region of isolate T318A has indeed corroborated our deep sequencing data (Navarro and Ambrós, personal communication). Altogether, these results extend those previously obtained for another closterovirus with a genome approximately half-size (Kreuze et al. 2009), and highlight the potential of the deep sequencing approach used here to identify virus and even virus variability, because the intrinsic heterogeneity generated by the technique is lower than the genetic variability resulting from replication.

Effect of CTV infection on the miRNAs from the three citrus hosts

To complement these data, we also searched our deep sequencing libraries first for miRNAs, and then for virus-induced alterations of their relative abundance. In regard to the first point, the only previous data on citrus miRNAs have been obtained by searching citrus EST databases for miRNA precursors of 26 miRNAs conserved in Arabidopsis and other plant species, with 13 precursors found adopting fold-back structures similar to those of Arabidopsis. Northern hybridization revealed the expression of these 13 citrus microRNAs and of the other 13 putative miRNAs,

with 41 potential mRNA targets identified by nearly-perfect complementarity with 15 citrus miRNAs (and four of these targets experimentally validated by detection of the miRNA-mediated cleavage products) (Song et al. 2007).

By comparative analysis of our deep-sequencing citrus libraries with miRNAs from other plant species and searching for their fold-back precursors in EST citrus databases, we identified 42 miRNAs from 27 different conserved families (Supplementary Table S3); further support for their existence was provided by the identification of some citrus miRNAs* (located in the opposite arm of pre-miRNA hairpins). Out of these 42 miRNAs, 24 are new in citrus, 17 of which belong to 13 conserved families with no previous citrus member. Additional examination of our sRNAs libraries showed the presence of all miRNAs previously detected by Northern-blot hybridization in citrus with the exception of miR161, miR163, miR170, miR173 and miR391 (Song et al. 2007), which have been identified only in Arabidopsis (excluding miR161 also found in *Brassica rapa)*.

Concerning the effect of virus infection on the miRNA relative abundance, we restricted our study to the 28 (of the 42) citrus miRNAs with more than 10 reads in at least one of the three hosts (Supplementary Table S3). The analysis revealed a complex pattern in qualitative and quantitative terms (Fig. 8). Some miRNAs were upregulated (i.e. miR168 and miR398) or downregulated (i.e. miR171a and miR403) in the three hosts, while others displayed a host-specific change (i.e. miR156 with an increase of six-fold log units in Mexican lime, and miR166b with a four-fold increase in sour orange and a three-fold decrease in sweet orange). The changes in miRNA profiles in response to CTV infection were validated by Northern-blot hybridization for two of the most abundant miRNAs (Supplementary Fig. S2). Most of the 28 citrus miRNAs examined target mRNAs coding for transcription factors (Song et al. 2007 and data not shown). It is worth noting that miR168 (upregulated by CTV in the three hosts, see above) targets the mRNA coding for AGO1, which mediates miRNA- (Baumberger et al. 2005) and vsRNA-directed gene silencing (Morel et al. 2002; Qu, et al. 2008). Transcriptional co-regulation of miR168 and AGO1 mRNA has been described (Vaucheret et al., 2006), with recent data indicating that induction of AGO1 mRNA and the ensuing accumulation of AGO1 is a component of the host defense response, while the induction of miR168 observed in several plant-virus interactions

is a counter-defense reaction of the virus (Varallyay et al. 2010). It will be interesting to examine whether any of the three silencing suppressors of CTV also targets AGO1.

Mapping plant sRNAs reveals a hotspot at the locus *Ctv* of trifoliate orange containing the putative resistance gene

The host sRNA profile was similar in the three mock-inoculated controls, being dominated by the 24-nt species, in common with other angiosperms (Dolgosheina et al. 2008; Morin et al. 2008) (Fig. 1). CTV infection induced minor changes on the relative size distribution of total (plant and virus) sRNAs in sour orange, but had a profound effect in Mexican lime and sweet orange: while a minor 21-nt peak and a major 24-nt peak were detected in the mock-inoculated controls, essentially the reverse situation occurred in their CTV-infected counterparts, with a major peak of 21- 22-nt and a minor peak of 24 nt. This relative increase in 21- 22-nt sRNAs might result from their specific binding and stabilization by one or more of the CTV-encoded silencing suppressors, as proposed for other virus-encoded silencing suppressors (Csorba et al., 2009; Ding, 2010), although there are alternative explanations. Focusing on plant sRNAs of 24 nt, virus infection produced in Mexican lime and sweet orange a 2.5-3 fold reduction whereas the effects on the plant sRNAs of 21 and 22 nt were less pronounced and essentially absent, respectively (Fig. 1). Intriguingly, co-infection of sweet potato with SPCSV and *Sweet potato feathery mottle virus* (SPFMV, family *Potyviridae),* producing the synergistic sweet potato virus disease (SPVD), also results in a severe relative reduction of the 24-nt sRNAs (essentially derived from the plant) that has not been observed in singly-infected plants (Kreuze et al. 2009).

To get a more precise insight, we first examined the genomic location of the 24-nt plant sRNAs most deeply affected by CTV infection. Unexpectedly, one of these sRNAs (5'-CUUAGAAUUGAUUGCAAAGCUGCA-3'), overexpressed in CTV-infected Mexican lime and sweet orange but not in sour orange, was homologous (with a single mismatch) to a fragment of the 282-kb region from the trifoliate orange genome containing locus *Ctv* for CTV resistance, and more specifically to the putative gene CTV.20 located within this region (see below). Given the functional

relevance of the 282-kb region, we looked for other plant sRNAs mapping within it. Remarkably, 1-4% of all plant sRNAs (depending on the species and on whether it was infected or not) derived from this region, a value higher than expected considering the size of the citrus genome (367 Mb) as well as the sequence differences existing between the trifoliate orange and the three citrus hosts here assayed. Moreover, the accumulation of 24-nt plant sRNAs with sequences homologous to the 282-kb region of the trifoliate orange was increased by CTV infection in sour orange, while in Mexican lime and, particularly, in sweet orange was decreased (Supplementary Fig. S3). The changes in sRNA profiles resulting from CTV infection were validated by Northern-blot hybridization for the most abundant 24-nt sRNA of locus *Ctv* (Supplementary Fig. S2). Further analysis revealed an uneven distribution of the 24-nt plant sRNAs along the 282-kb region, with major peaks mapping at some intergenic regions (between the putative genes CTV.5 and CTV.6, and CTV.10 and CTV.11) and at the putative gene CTV.20 (Fig. 9). Searching for an explanation for this profile we found: i) an inverted repeat (plus/ minus) of more than 1100 nt and 89% sequence identity between positions 88901-90066 (intergenic region CTV.5-CTV.6) and positions 154117-155302 (intergenic region CTV.10-CTV.11), and ii) a direct repeat (plus/ plus) of more than 6000 nt and 99% sequence identity within CTV.20 between positions 233519-239777 and 241051-247107. Interestingly, there are ESTs in databases from different citrus species with up to 93 and 97% sequence identity with these two repeats (data not shown).

Sequence analysis of the 282-kb region of the trifoliate orange genome containing locus *Ctv* has predicted 22 putative genes, including a cluster of seven disease resistance (R) genes, two transposons and eight retrotransposons (Fig. 9) (Yang et al. 2003). How these genes, and particularly the gene(s) providing resistance to CTV, are regulated is not known, but our data suggest that CTV infection might affect differentially their expression in highly susceptible hosts (like Mexican lime and sweet orange) and in a host like sour orange with lower virus accumulation. There is evidence that RNA silencing modulates the expression of an increasing number of endogenous genes, and that defense mechanisms can be activated in response to attack by pathogens that disturb RNA silencing (Yi and Richards 2007). CTV has this potential through its three encoded silencing suppressors, but whether the resistance

gene(s) is regulated in such a way remains to be determined.

Conclusions

Deep sequencing has revealed that the CTV sRNAs from young stem bark represent a minor fraction of total sRNAs in sour orange, but more than 50% in Mexican lime and sweet orange, in which the profile of the 24-nt plant sRNAs and some miRNAs is significantly altered. On the other hand, the abundance of CTV sRNAs in Mexican lime and sweet orange suggests that it is their function, rather than their synthesis, the target of the RNA silencing suppressors encoded by this virus (p25, p20 and p23). Actually, accumulation of siRNAs derived from a CTV transgene in Mexican lime is not sufficient for eliciting protection against the virus (López et al. 2010). It is conceivable that some of the intracellular suppressors (p23 and p20) may block load of CTV sRNAs into RISC, thus preventing or attenuating its saturation and the ensuing deleterious effects, or interfere with RISC function through other indirect routes. Finally, the preferential accumulation of CTV sRNAs mapping at the 3'-terminal 2500-nt of the RNA genome most likely results from the concurrent accumulation of this fragment present in the ten 3' co-terminal sgRNAs and their dsRNAs; this excess would overpass a threshold level and trigger —as proposed with the over-expression of certain transgenes— the plant RNA silencing response, including the activation of RDRs (Voinnet 2008; Wassenegger and Krczal 2006) ultimately resulting in the generation of secondary vsRNAs. Therefore, our results establish that vsRNAs from CTV, and presumably from other (+) RNA viruses expressing sgRNAs, result from DCLs acting on multiple viral substrates generated in the course of gRNA replication and sgRNA transcription.

Acknowledgments This research was supported by a grant (Prometeo/ 2008/ 121) from the Generalitat Valenciana, Spain, and by an aid (PAID-02-10/ 2180) from the Program for Research and Development of the Universidad Politécnica de Valencia. We are grateful to Jaime Piquer and Pablo Lemos for technical support in the greenhouse and with the illustrations, respectively.

Supplementary material The online version of this article (doi pending) contains supplementary material, which is available to authorized users.

References

Albiach-Martí MR, Grosser JW, Gowda S, Mawassi M, Tatineni S, Garnsey S, Dawson WO (2004) Citrus tristeza virus replicates and forms infectious virions in protoplasts of resistant citrus relatives. Mol Breed 14:117-128.

Albiach-Martí MR, Robertson C, Gowda S, Tatineni S, Belliure B, Garnsey SM, Folimonova SY, Moreno P, Dawson WO (2010) The pathogenicity determinant of citrus tristeza virus causing the seedling yellows syndrome maps at the 3′-terminal region of the viral genome. Mol Plant Pathol 11:55-67.

Aliyari R, Wu QF, Li HW, Wang XH, Li F, Green LD, Han CS, Li WX, Ding SW (2008) Mechanism of induction and suppression of antiviral immunity directed by virus-derived small RNAs in *Drosophila*. Cell Host Microbe 4:387-397.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. Mol Biol 215:403-410.

Ancillo G, Gadea J, Forment J, Guerri J, Navarro L (2007) Class prediction of closely related plant varieties using gene expression profiling. J Exp Bot 58: 1927-1933.

Aramburu J, Navas-Castillo J, Moreno P, Cambra M (1991) Detection of double-stranded RNA by ELISA and dot immunobinding assay using an antiserum to synthetic polynucleotides. J Virol Methods 33:1-11.

Azevedo J, García D, Pontier D, Ohnesorge S, Yu A, García S, Braun L, Bergdoll M, Hakimi MA, Lagrange T, Voinnet O (2010) Argonaute quenching and global changes in Dicer homeostasis caused by a pathogen-encoded GW repeat protein. Genes Dev 24:853-856.

Bar-Joseph M, Dawson WO (2008) Citrus tristeza virus. In Mahy BWJ, Van Regenmortel MHV (eds) Encyclopedia of virology, 3rd edn. Elsevier, Oxford, pp 520-525.

Baumberger N, Baulcombe DC (2005) Arabidopsis ARGONAUTE1 is an RNA slicer that selectively recruits microRNAs and short interfering RNAs. Proc Natl Acad Sci USA 102:11928-11933.

Bellamy AR, Ralph RK (1968) Recovery and purification of nucleic acids by means of cetyltrimethylammonium bromide. Methods Enzymol. 12B:156-160.

Csorba T, Pantaleo V, Burgyán J (2009) RNA silencing: an antiviral mechanism. Adv Virus Res 75:35-71.

Curtin SJ, Watson JM, Smith NA, Eamens AL, Blanchard CL, Waterhouse PM (2008) The roles of plant dsRNA-binding proteins in RNAi-like pathways. FEBS Lett 582:2753-2760.

Dalmay T, Hamilton A, Rudd S, Angell S, Baulcombe DC (2000) An RNA-dependent RNA polymerase gene in Arabidopsis is required for posttranscriptional gene silencing mediated by a transgene but not by a virus. Cell 101:543-553.

Deleris A, Gallego-Bartolomé J, Bao J, Kasschau KD, Carrington JC, Voinnet O (2006) Hierarchical action and inhibition of plant Dicer-like proteins in antiviral defense. Science 313:68-71.

Díaz-Pendón JA, Ding SW (2008) Direct and indirect roles of viral suppressors of RNA silencing in pathogenesis. Annu Rev Phytopathol 46:303-326.

Ding SW (2010) RNA-based antiviral immunity. Nature Rev Immunol 10:632-644.

Di Serio F, Gisel A, Navarro B, Delgado S, Martínez de Alba AE, Donvito G, Flores R (2009) Deep sequencing of the small RNAs derived from two symptomatic variants of a chloroplastic viroid: implications for their genesis and for pathogenesis. PLoS ONE 4:e7539.

Di Serio F, Martínez de Alba AE, Navarro B, Gisel A, Flores R (2010) RNA-dependent RNA polymerase 6 delays accumulation and precludes meristem invasion of a nuclear-replicating viroid. J Virol 84:2477-2489.

Dolgosheina EV, Morin RD, Aksay G, Sahinalp SC, Magrini V, Mardis ER, Mattsson, J, Unrau PJ (2008) Conifers have a unique small RNA silencing signature. RNA 14:1508-1515.

Donaire L, Barajas D, Martínez-García B, Martínez-Priego L, Pagán I, Llave C (2008) Structural and genetic requirements for the biogenesis of tobacco rattle virus-derived small interfering RNAs. J Virol 82:5167-5177.

Donaire L, Wang Y, González-Ibeas D, Mayer KF, Aranda MA, Llave C (2009) Deep-sequencing of plant viral small RNAs reveals effective and widespread targeting of viral genomes. Virology 392:203-214.

Fagoaga C, López C, Moreno P, Navarro L, Flores R, Peña L (2005) Viral-like symptoms induced by the ectopic expresión of the p23 of citrus tristeza virus are citrus specific and do not correlate with the patogenicity of the virus strain. Mol Plant-Microbe Interact 18:435-445.

Fagoaga C, López C, Hermoso de Mendoza AH, Moreno P, Navarro L, Flores R, Peña L (2006) Post-transcriptional gene silencing of the p23 silencing suppressor of citrus tristeza virus confers resistance to the virus in transgenic Mexican lime. Plant Mol Biol 66:153-165.

Folimonova SY, Folimonov AS, Tatineni S, Dawson WO (2008). Citrus tristeza virus: survival at the edge of the movement continuum. J Virol 82:6546-6556.

Ghorbel R, López C, Fagoaga C, Moreno P, Navarro L, Flores R, Peña L (2001) Transgenic citrus plants expressing the citrus tristeza virus p23 protein exhibit viral-like symptoms. Mol Plant Pathol 2:27-36.

Hamilton AJ, Baulcombe DC (1999) A species of small antisense RNA in posttranscriptional gene silencing in plants. Science 286:950-952.

Hilf ME, Karasev AV, Pappu HR, Gumpf DJ, Niblett CL, Garnsey SM (1995) Characterization of citrus tristeza virus subgenomic RNAs in infected tissue. Virology 208:576-582.

Karasev AV, Boyko VP, Gowda S, Nikolaeva OV, Hilf ME, Koonin EV, Niblett CL, Cline K, Gumpf DJ, Lee RF, Garnsey SM, Lewandowski DJ, Dawson WO (1995). Complete sequence of the citrus tristeza virus RNA genome. Virology 208:511-520.

Kreuze JF, Pérez A, Untiveros M, Quispe D, Fuentes S, Barker I, 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.

López C, Navas-Castillo J, Gowda S, Moreno P, Flores, R. (2000) The 23 kDa protein coded by the 3′-terminal gene of citrus tristeza virus is an RNA-binding protein. Virology 269:462-470.

López C, Cervera M, Fagoaga C, Moreno P, Navarro L, Flores R, Peña L (2010) Accumulation of transgene-derived siRNAs is not sufficient for RNAi-mediated protection against citrus tristeza virus (CTV) in transgenic Mexican lime. Mol Plant Pathol 11:33-41.

Lu R, Folimonov A, Shintaku M, Li WX, Falk BW, Dawson WO, Ding SW (2004) Three distinct suppressors of RNA silencing encoded by a 20-kb viral RNA genome. Proc Natl Acad Sci USA 101:15742-15747.

Mi S, Cai T, Hu Y, Chen Y, Hodges E, Ni F, Wu L, Li S, Zhou H, Long C, Chen S, Hannon GJ, Qi Y (2008) Sorting of small RNAs into Arabidopsis argonaute complexes is directed by the 5' terminal nucleotide. Cell 133:116-127.

Moissiard G, Voinnet O (2006) RNA silencing of host transcripts by cauliflower mosaic virus requires coordinated action of the four Arabidopsis Dicer-like proteins. Proc Natl Acad Sci USA 103:19593-19598.

Molnar A, Csorba T, Lakatos L, Varallyay E, Lacomme C, Burgyan J (2005). Plant virus-derived small interfering RNAs originate predominantly from highly structured single-stranded viral RNAs. J Virol 79:7812-7818.

Montgomery TA, Howell MD, Cuperus JT, Li D, Hansen JE, Alexander AL, Chapman EJ, Fahlgren N, Allen E, Carrington JC (2008) Specificity of ARGONAUTE7-miR390 interaction and dual functionality in TAS3 trans-acting siRNA formation. Cell 133:128-141.

Morel JB, Godon C, Mourrain P, Beclin C, Boutet S, Feuerbach F, Proux F, Vaucheret H (2002) Fertile hypomorphic ARGONAUTE (ago1) mutants impaired in posttranscriptional gene silencing and virus resistance. Plant Cell 14:629-639.

Moreno P, Guerri J, Muñoz N (1990) Identification of Spanish strains of citrus tristeza virus (CTV) by analysis of double-stranded RNAs (dsRNA). Phytopathology 80:477-482.

Moreno P, Ambrós S, Albiach-Martí MR, Guerri J, Peña L (2008) Citrus tristeza virus: a pathogen that changed the course of the citrus industry. Mol Plant Pathol 9:251-268.

Morin RD, Aksay G, Dolgosheina E, Ebhardt HA, Magrini V, Mardis, ER, Sahinalp SC, Unrau PJ (2008) Comparative analysis of the small RNA transcriptomes of *Pinus contorta* and *Oryza sativa*. Genome Res 18:571-584.

Navas-Castillo J, Albiach-Martí MR, Gowda S, Hilf ME, Garnsey SM, Dawson WO (1997) Kinetics of accumulation of citrus tristeza virus RNAs. Virology 228:92-97.

Omarov RT, Cioperlik JJ, Sholthof HB (2007) RNAi-associated ssRNA-specific ribonucleases in tombusvirus P19 mutant-infected plants and evidence for a discrete siRNA-containing effector complex. Proc Natl Acad Sci USA 104:1714-1719.

Pantaleo V, Szittya G, Burgyán J (2007) Molecular bases of viral RNA targeting by viral small interfering RNA-programmed RISC. J Virol 81:3797-3806.

Qi Y, Denli AM, Hannon GJ (2005) Biochemical specialization within Arabidopsis RNA silencing pathways. Mol Cell 19:421-428.

Qu F, Ye X, Morris TJ (2008) Arabidopsis DRB4, AGO1, AGO7, and RDR6 participate in a DCL4-initiated antiviral RNA silencing pathway negatively regulated by DCL1. Proc Natl Acad Sci USA 105:14732-14737.

Ruiz-Ruiz S, Moreno P, Guerri J, Ambrós S (2006). The complete nucleotide sequence of a severe stem pitting isolate of citrus tristeza virus from Spain: comparison with isolates from different origins. Arch Virol 151:387-398.

Ruiz-Ruiz S, Moreno P, Guerri J, Ambrós S (2007) A real-time RT-PCR assay for detection and absolute quantitation of citrus tristeza virus in different plant tissues. J Virol Methods 145:96-105.

Satyanarayana T, Gowda S, Ayllón MA, Albiach-Martí MR, Rabindram R, Dawson WO (2002) The p23 protein of citrus tristeza virus controls asymmetrical RNA accumulation. J Virol 76:473-483.

Smith NA, Eamens AL, Wang MB (2010) The presence of high-molecular-weight viral RNAs interferes with the detection of viral small RNAs. RNA 16:1062-1067.

Song C, Fang J, Li X, Liu H, Thomas-Chao C (2007) Identification and characterization of 27 conserved microRNAs in citrus. Planta 230:671-685.

Szittya G, Moxon S, Pantaleo, V, Toth, G, Rusholme-Pilcher RL, Moulton V, Burgyan J, Dalmay T (2010) Structural and functional analysis of viral siRNAs. PLoS Pathog 6:e1000838.

Varallyay E, Valoczi A, Agyi A, Burgyan J, Havelda Z (2010) Plant virus-mediated induction of miR168 is associated with repression of ARGONAUTE1 accumulation. EMBO J 29:3507-3519.

Vaucheret H (2008) Plant ARGONAUTES. Trends Plant Sci 13:350-358.

Vaucheret H, Mallory AC, Bartel DP (2006) AGO1 homeostasis entails coexpression of miR168 and AGO1 and preferential stabilization of miR168 by AGO1. Mol Cell 22:129-136.

Voinnet O (2008) Use, tolerance and avoidance of amplified RNA silencing by plants. Trends Plant Sci 13:317-328.

Wassenegger M, Krczal G (2006) Nomenclature and functions of RNA-directed RNA polymerases. Trends Plant Sci 11:142-151.

Weber F, Wagner V, Rasmussen SB, Hartmann R, Paludan SR (2006) Double-stranded RNA is produced by positive-strand RNA viruses and DNA viruses but not in detectable amounts by negative-strand RNA viruses. J Virol 80:5059-5064.

Yang ZN, Ye XR, Molina J, Roose ML, Mirkov TE (2003) Sequence analysis of a 282-kilobase region surrounding the citrus tristeza virus resistance gene *(Ctv)* locus in *Poncirus trifoliata* L. Raf. Plant Physiol 131:482-492.

Yi K, Richards EJ (2007) A cluster of disease resistance genes in arabidopsis is coordinately regulated by transcriptional activation and RNA silencing. Plant Cell 19:2929-2939.

Zerbino DR, Birney E (2008) Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res 18:821-829.

Legends to Figures

Figure 1. Size distribution of CTV and plant sRNAs (orange and green, respectively) in CTV-infected and mock-inoculated Mexican lime, sour orange and sweet orange. The histograms compare the distribution of 18- to 26-nt total sRNA reads. A minor contamination of CTV-sRNAs observed in the mock-inoculated Mexican lime (representing 0.1 % of the total sRNAs of this size) has not been represented.

Figure 2. Northern-blot hybridization of CTV RNAs from mock-inoculated and CTV-infected Mexican lime, sour orange and sweet orange. Total RNA preparations were fractionated by electrophoresis in 1% agarose/ formaldehyde gels (for CTV gRNA and sgRNAs) **(A)**, or in 17% polyacrylamide/ urea gels (for CTV and plant sRNAs) **(B)**, electrotransferred to membranes, and hybridized with a digoxigenin-labeled riboprobe complementary to the 3'-terminal gene *p23*. Lanes (1) and (2) mock-inoculated and CTV-infected sweet orange, lanes (3) and (4) mock-inoculated and CTV-infected Mexican lime, and lanes (5) and (6) mock-inoculated and CTV-infected sour orange, respectively. The positions of the CTV gRNA and the most abundant sgRNAs are indicated in panel **(A)**, with the left lane M of panel **(B)** corresponding to a 23-nt RNA marker homologous to positions 18692 to 18714 of the CTV gRNA. Equal loading was assessed by ethidium bromide staining of 4S RNA **(A)** and by hybridization with a digoxigenin-labeled riboprobe complementary to U6 RNA **(B)**.

Figure 3. Polarity distribution of the 18- to 26-nt reads perfectly matching plus (blue) and minus (red) CTV sRNAs in the three citrus hosts assayed.

Figure 4. Location and frequency in the CTV gRNA of the 5' termini of (+) and (-) vsRNA reads from the three citrus hosts (bars above and below the x-axis, respectively). The preferential distribution of the reads at the 3'-terminal moiety of the gRNA is clearly visible. The ORFs, and 5'- and 3'-UTRs, are indicated within boxes and with lines, respectively. Note that the read scale is different, and that the same numbering is used in the plus polarity (5'→3' orientation is from left to right) and in the minus polarity (5'→3' orientation is from right to left). In the lower panel

are shown the signals generated by the vsRNAs from CTV-infected sweet orange when hybridized with eight equalized ≅1500-nt digoxigenin-labeled riboprobes spanning the complete viral genome. The nucleotide coordinates of the probes along the CTV genome are indicated, and the number of (+) CTV sRNAs reads within each of the eight regions analyzed are between parenthesis. Prior to hybridization, aliquots of the RNA preparation from CTV-infected sweet orange were fractionated by electrophoresis in 17% polyacrylamide/ urea gels and blotted to membranes that were washed and then revealed with the chemiluminescent substrate CSPD and exposure to X-ray film. A 23-nt RNA marker, homologous to positions 18692 to 18714 of the CTV gRNA, was also included (lane M).

Figure 5. Confirmation by Northern-blot hybridization of the existence and relative abundance of (+) and (-) CTV sRNAs in CTV-infected sweet orange (lanes 1), CTV-infected Mexican lime (lanes 2), CTV-infected sour orange (lanes 3), and mock-inoculated Mexican lime (lanes 4). **(A)** Analysis of total RNA preparations hybridized with digoxigenin-labeled riboprobes complementary and homologous to the *p23* gene (upper and middle panels, respectively). A 23-nt RNA marker, homologous to positions 18692 to 18714 of the CTV gRNA, was also included (lane M). **(B)** Analysis of the same preparations hybridized with 5'-radiolabeled oligonucleotides CTV-chd and CTV-cld for detecting abundant and rare (+) CTV sRNAs (upper and middle, respectively). In both instances RNAs were previously fractionated by electrophoresis in 17% polyacrylamide/ urea gels. After hybridization with equalized probes, membranes were revealed with the chemiluminescent substrate CSPD and exposure to X-ray film **(A)**, or with a bioimage analyzer **(B)**. Equal loading was assessed by hybridization with a digoxigenin-labeled riboprobe complementary to U6 RNA.

Figure 6. Density (reads per nucleotide) of the (+) and (-) CTV sRNA reads (18 to 26 nt) along the ORFs and the 5'- and 3'-UTRs in the three citrus hosts assayed.

Figure 7. Frequency of the 5' nucleotide in the CTV sRNAs. The histograms compare in the three citrus hosts assayed the total reads corresponding to the 21- 22-nt CTV

sRNAs of (+) and (-) polarity (upper and lower panels, respectively) with distinct 5' termini.

Figure 8. Effect of CTV infection on the accumulation of specific miRNA from Mexican lime, sour orange and sweet orange. Only miRNAs with more that 10 reads in at least one of the three citrus hosts are considered, and the expression levels of miRNAs are plotted as $\log_2$ of fold change (reads of CTV-infected versus mock-inoculated samples). Some miRNA sequences predicted by comparative analyses (Song et al. 2007) exhibited minor variations with respect to those retrieved from our deep sequencing libraries. In particular: i) the miR156 and miR393 most frequently sequenced in Mexican lime, sweet and sour orange have an additional nucleotide at the 5'-end when compared with the sequence deposited in the miRbase and predicted, respectively, ii) the most abundant miR166 and miR172 have their 5'-termini shifted one or two nucleotides downstream with respect to those predicted and deposited in the miRbase, respectively, and iii) the prevalent miR166 in the three citrus species analyzed is 21-nt long and does not contain the additional 3' nucleotide proposed in some members of this family (csi-miR166a, csi-miR166b, ctv-mir166, crt-miR166a, crt-miR166b) deposited in the miRbase (see also Supplementary Table S3).

Figure 9. Mapping of the 24-nt sRNAs from the three mock-inoculated and CTV-infected citrus host assayed along the 282-kb region containing the locus *Ctv* of trifoliate orange. The search allowed up to two mismatches. Abbreviations of the 22 putative genes (CTV.1 to CTV.22) are according to Yang et al. (2003) and the corresponding intergenic regions are named as Igr.1 to Igr.21).

Supplementary Figures and Tables

Figure S1. Location and frequency in the 3' terminal 275 nt of the CTV gRNA of the 5' termini of (+) and (-) vsRNA reads from the three citrus hosts. Note that essentially no reads are observed mapping at the 3'-terminal 50 nt. Other details as in the legend of Fig. 4.

Figure S2. Northern-blot hybridization showing the effect of CTV infection on the accumulation of two miRNAs (miR166 and miR319) (Supplementary Table S3) and one sRNA (5'-ATGTACTGTTGATGGTGACGTGGC-3', positions 154866-154899 and 89318-89337 in the minus and plus strands of the trifoliate orange locus *Ctv)* from mock-inoculated and CTV-infected Mexican lime, sour orange and sweet orange. Total RNA preparations were fractionated by electrophoresis in 17% polyacrylamide/ urea gels, electrotransferred to membranes and hybridized with 5'-radiolabeled probes complementary to the miRNAs and sRNA. Lanes (1) and (2) mock-inoculated and CTV-infected sweet orange, lanes (3) and (4) mock-inoculated and CTV-infected Mexican lime, and lanes (5) and (6) mock-inoculated and CTV-infected sour orange, respectively. The reads for the miRNAs and the sRNA in the six samples are between parenthesis. Equal loading was assessed by hybridization with a digoxigenin-labeled riboprobe complementary to U6 RNA and by ethidium bromide staining of 4S RNA.

Figure S3. Effect of CTV infection on the accumulation of the 24-nt sRNA from the three citrus hosts (Mexican lime and sweet and sour orange) with homology to the 282-kb region of the trifoliate orange genome containing locus *Ctv* (Yang et al. 2003). The expression levels of the sRNAs are plotted as $\log_2$ of reads per million (rpm). The search allowed up to two mismatches. Black line represents the diagonal (y = x).

Table S1. Length and reads of sequenced inserts from mock-inoculated and CTV-infected Mexican lime and sour orange.

Table S2. Length and reads of sequenced inserts from mock-inoculated and CTV-

infected sweet orange.

Table S3. Characteristics of the miRNAs retrieved from the deep sequencing libraries of mock-inoculated and CTV-infected Mexican lime, sour orange and sweet orange.

Screen/black and white figure

Click here to download Screen/black and white figure: Fig 1 (Ruiz) (31111).ppt

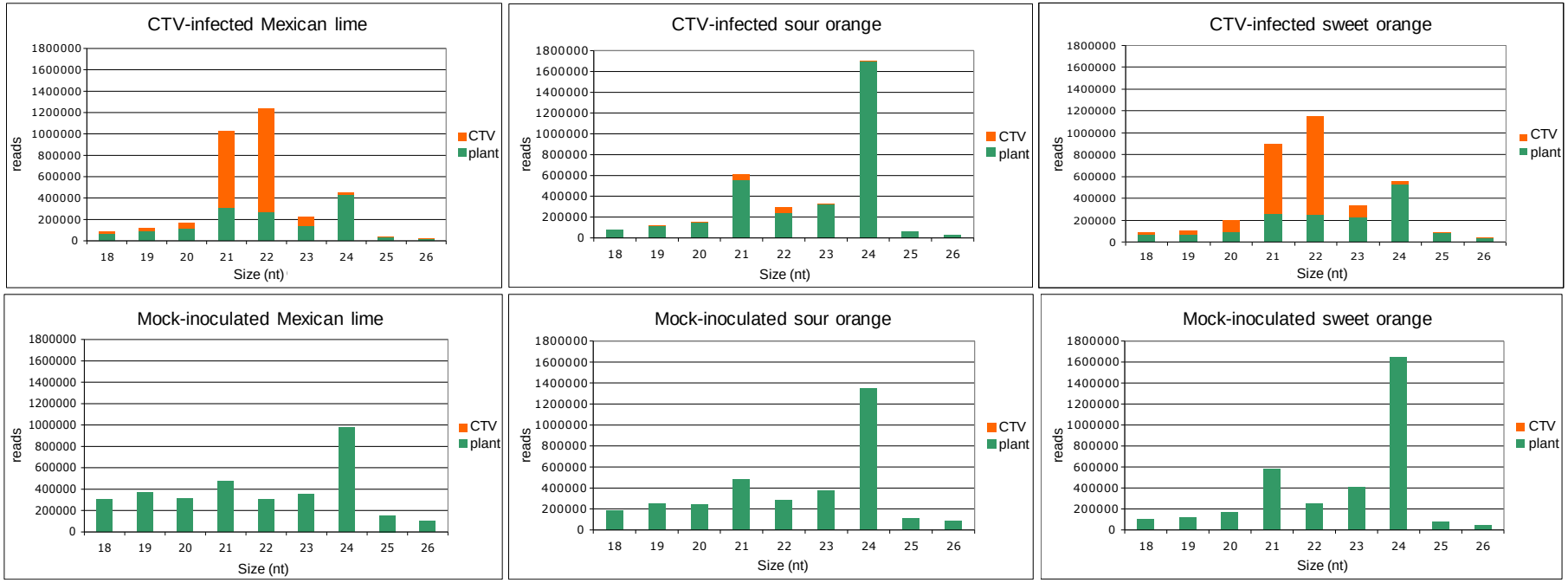


Figure 1

Screen/black and white figure

Click here to download Screen/black and white figure: Fig 2 (Ruiz) (31111).ppt

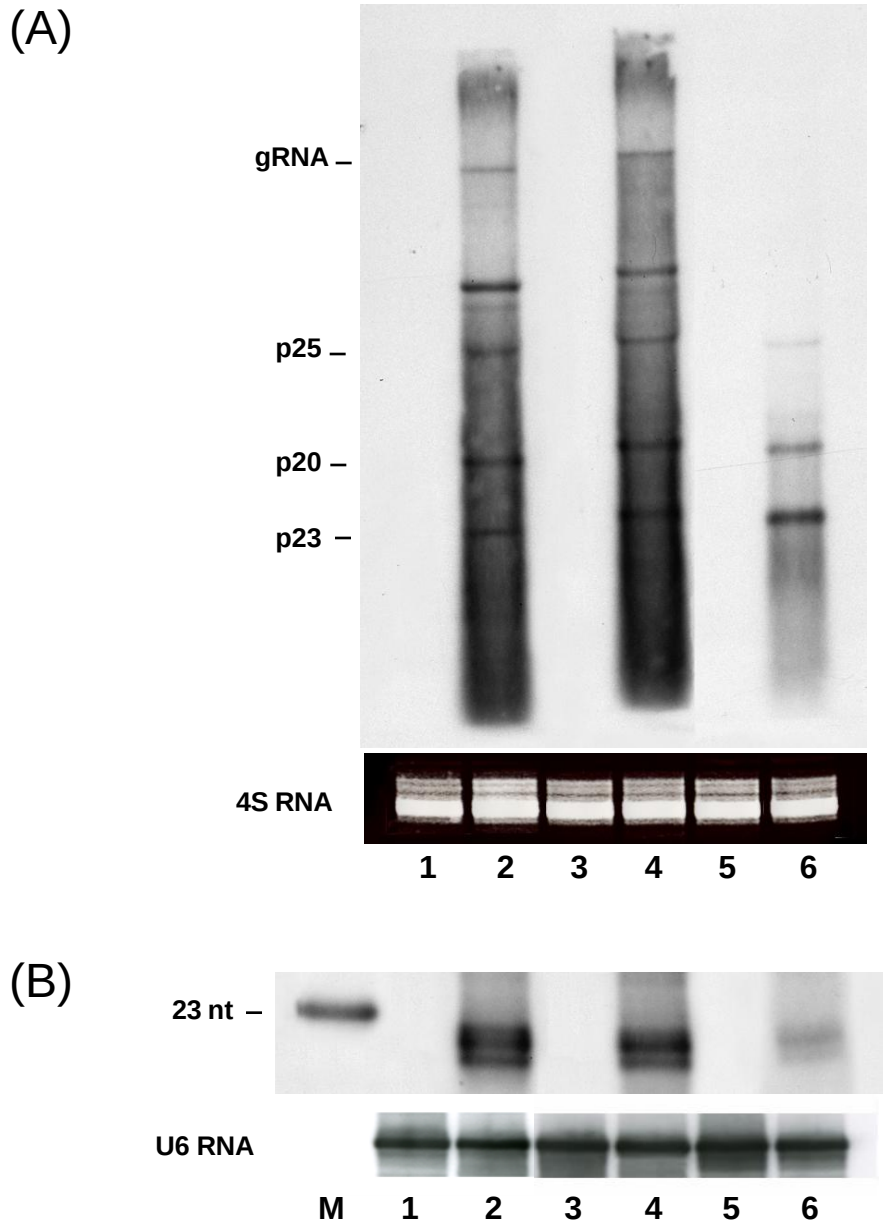


Figure 2

Screen/black and white figure

Click here to download Screen/black and white figure: Fig 3 (Ruiz) (31111).ppt

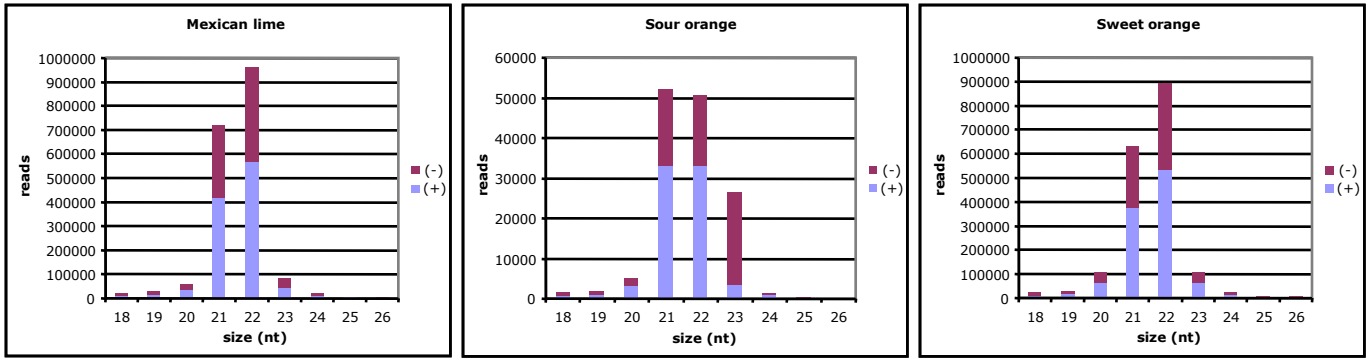


Figure 3

Screen/black and white figure

Click here to download Screen/black and white figure: Fig 4 (Ruiz) (31111).ppt

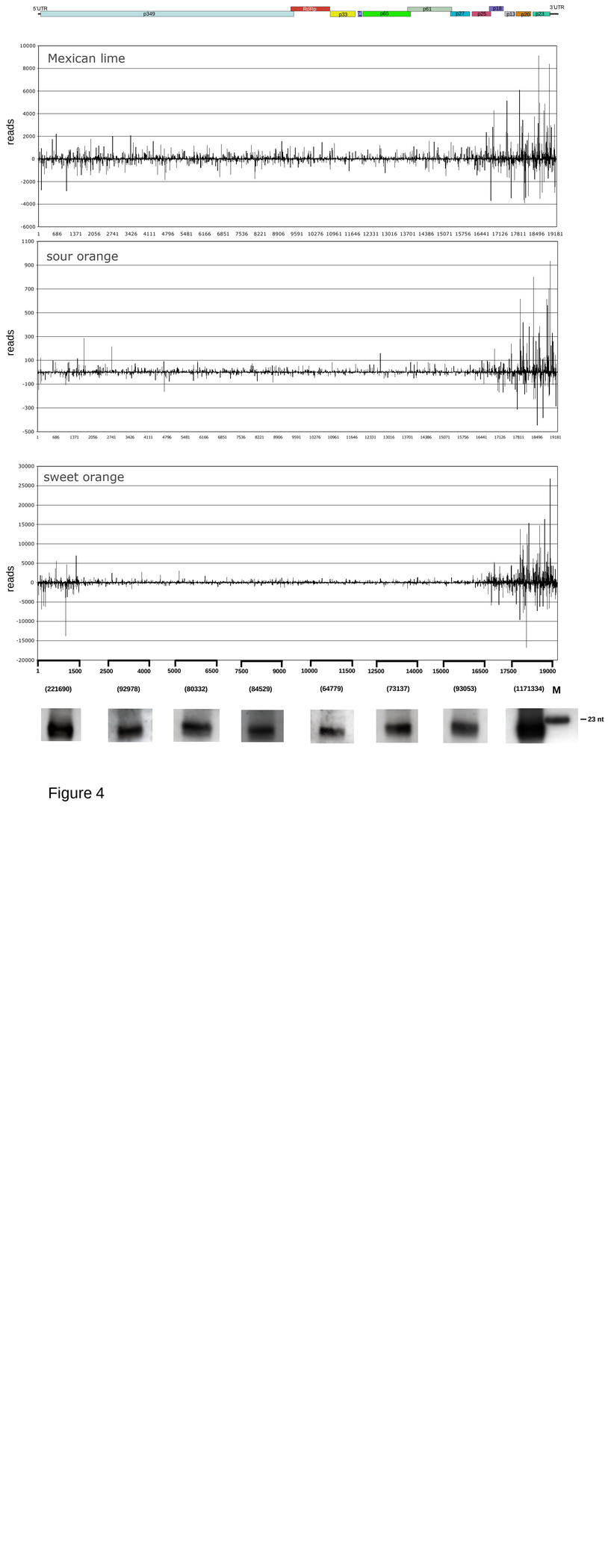
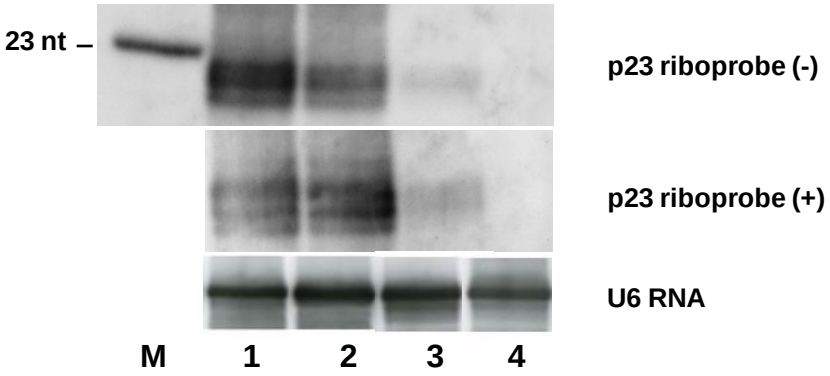


Figure 4

Screen/black and white figure

Click here to download Screen/black and white figure: Fig 5 (Ruiz) (31111).ppt

(A)



(B)

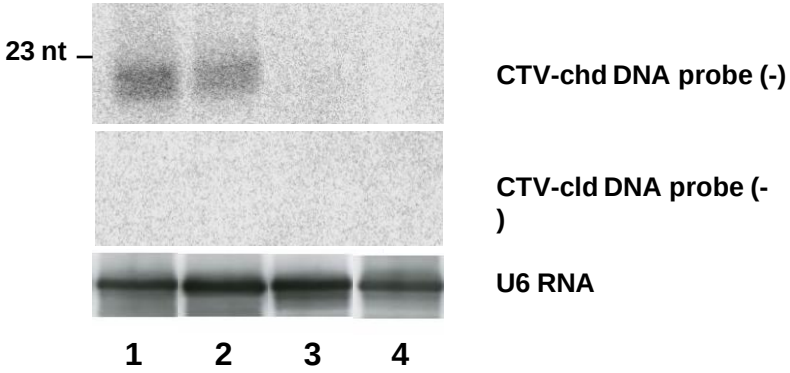


Figure 5

Colour figure

Click here to download Colour figure: Fig 6 (Ruiz) (31111).ppt

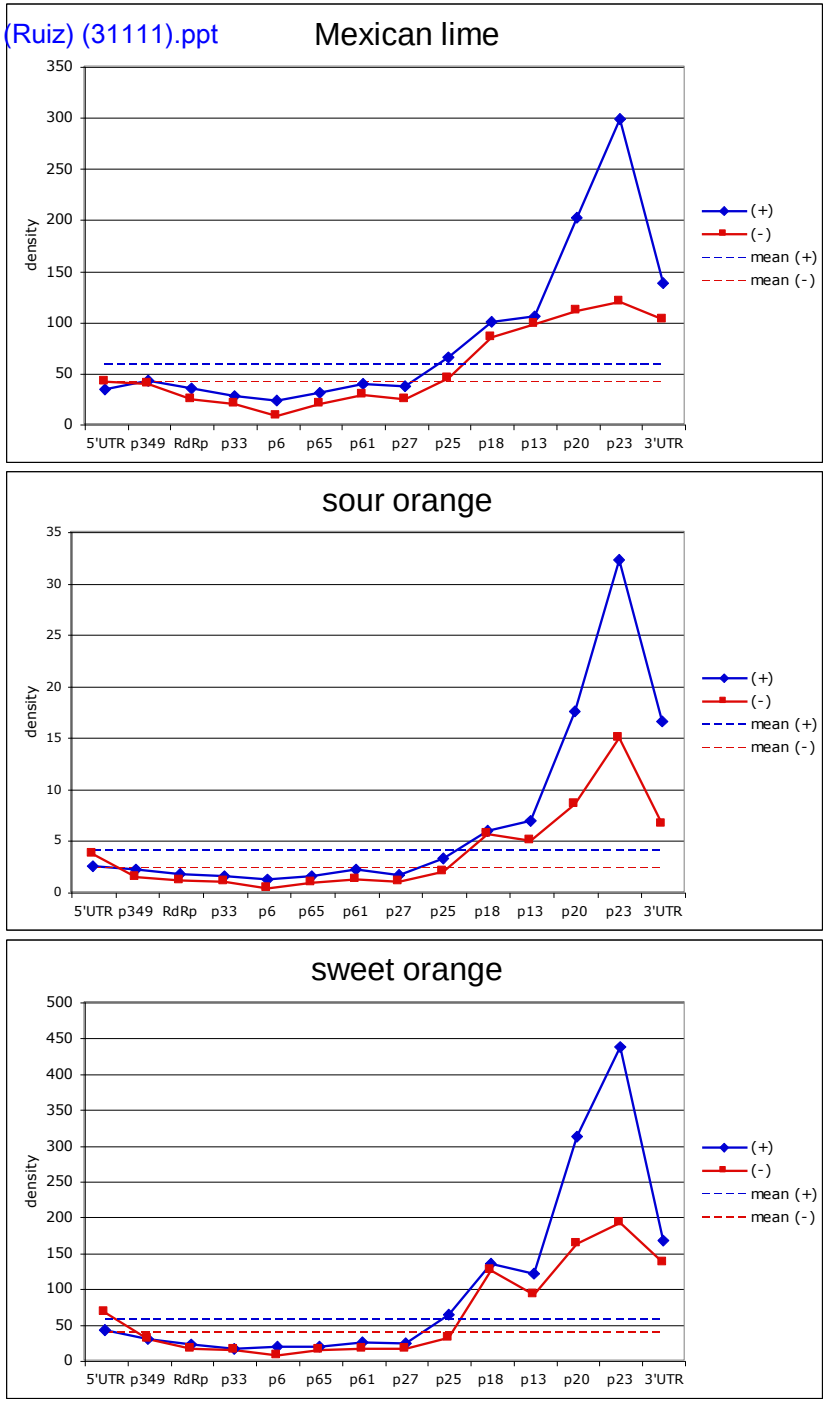


Figure 6

Screen/black and white figure

Click here to download Screen/black and white figure: Fig 7 (Ruiz) (31111).ppt

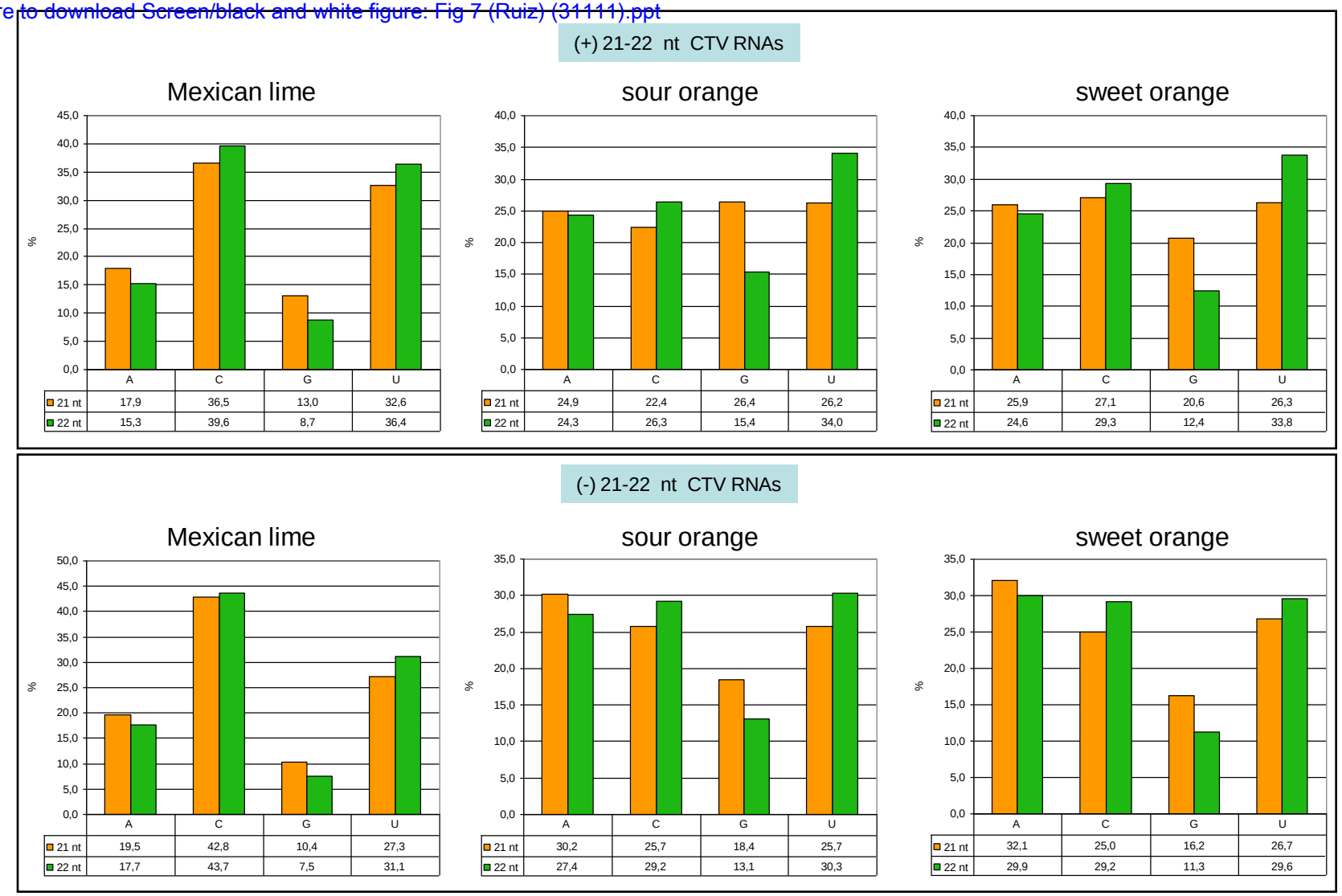


Figure 7

Colour figure

Click here to download Colour figure: Fig 8 (Ruiz) (31111).ppt

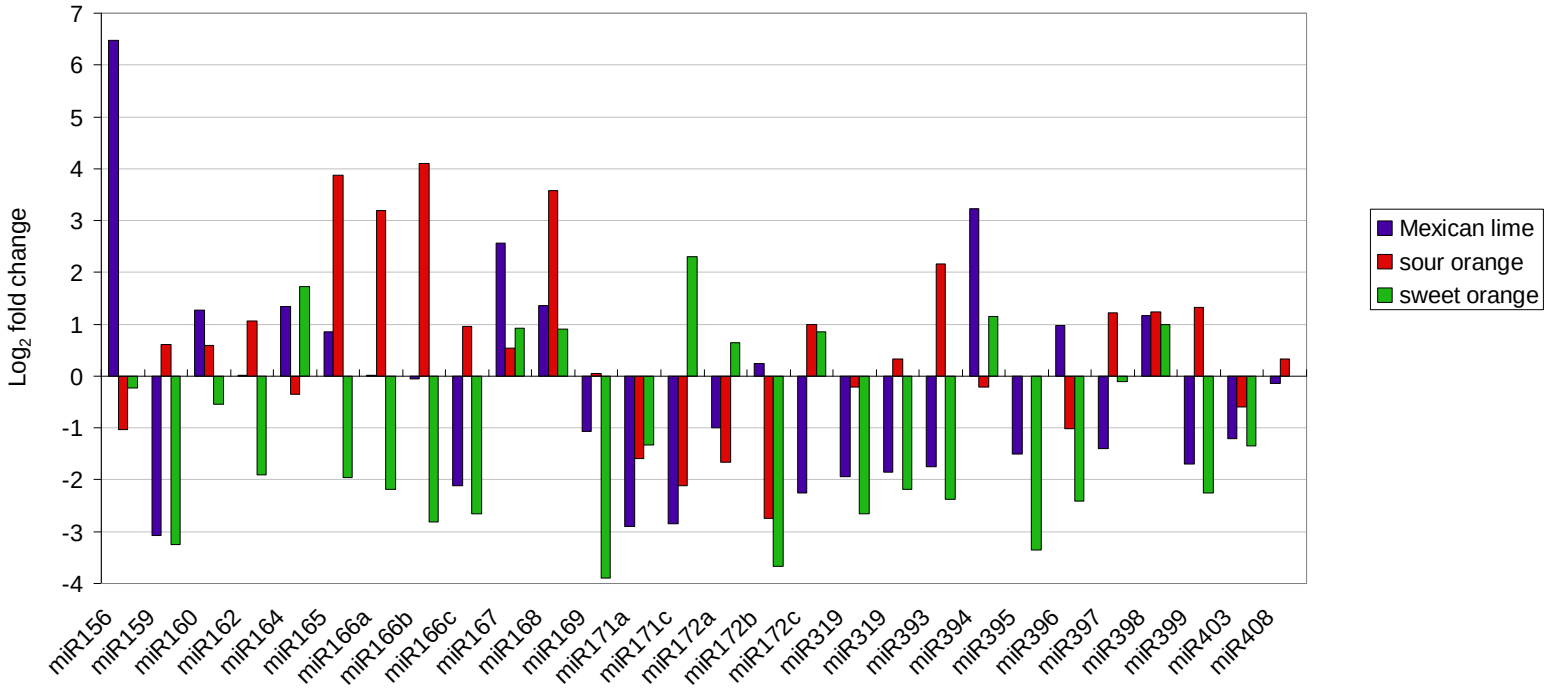


Figure 8

Colour figure

Click here to download Colour figure: Fig 9 (Ruiz) (31111).ppt

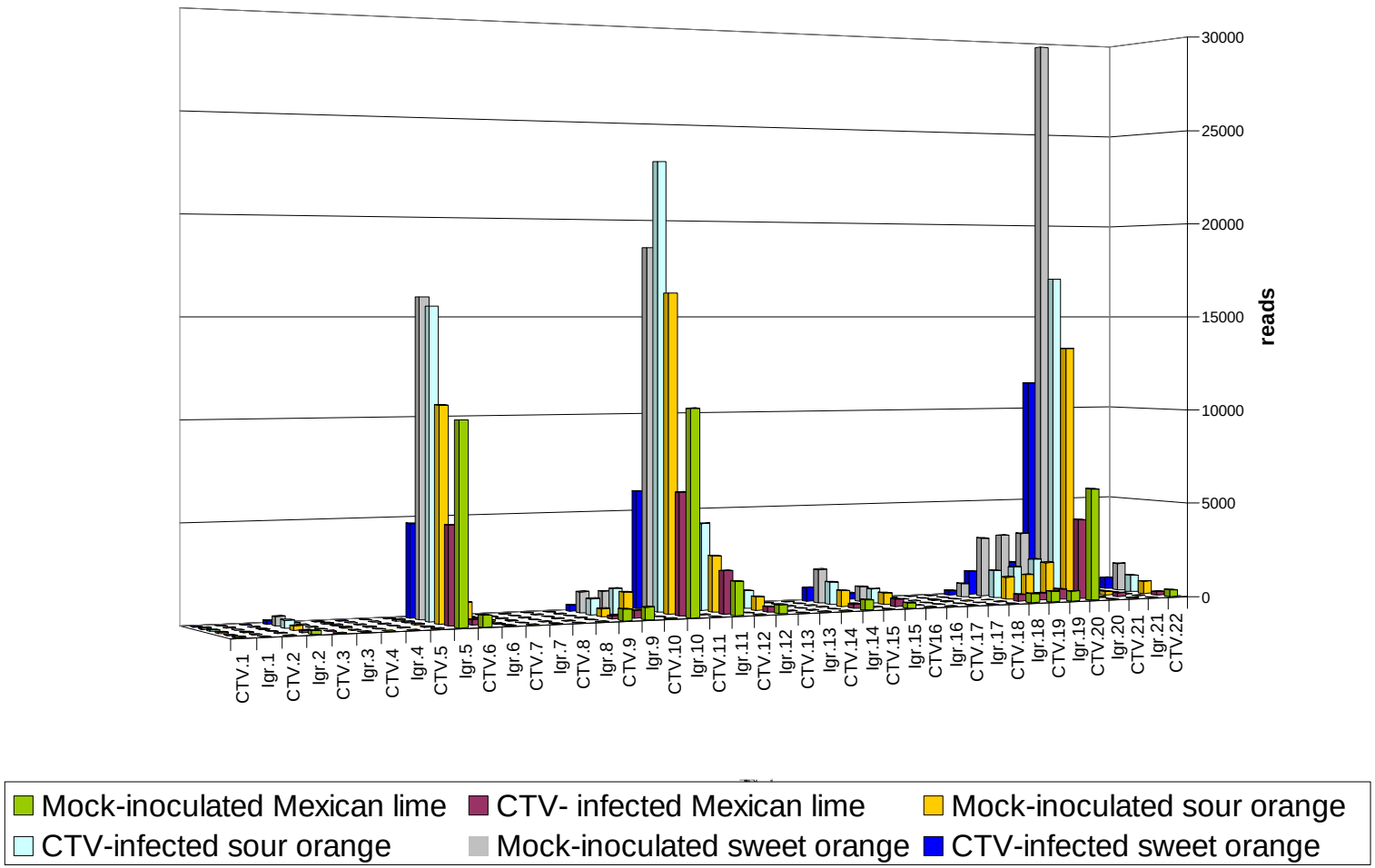


Figure 9