The genome of a bunyavirus cannot be defined at the level of the viral particle but only at the scale of the viral population

Michel Yvon, Thomas German, Diane Ullman, Ranjit Dasgupta, Maxwell Parker, Sulley Ben-Mahmoud, Eric Verdin, Patrick Gognalons, Aurélie Ancelin, Joséphine Laï Kee Him, et al.

VERSION ACCEPTED FOR PUBLICATION

Main Manuscript for

The genome of a bunyavirus cannot be defined at the level of the viral particle but only at the scale of the viral population

Authors

Yvon Michel[1], Thomas L. German[2‡], Diane E. Ullman[3], Ranjit Dasgupta[2], Maxwell H. Parker[2], Sulley Ben-Mahmoud[3], Eric Verdin[4], Patrick Gognalons [4], Aurélie Ancelin[5], Joséphine Laï Kee Him[5], Justine Girard[5], Marie-Stéphanie Vernerey[1], Emmanuel Fernandez[1], Denis Filloux[1], Philippe Roumagnac[1], Patrick Bron[5], Yannis Michalakis[6#] and Stéphane Blanc[1#*]

Affiliations

1- PHIM, Univ Montpellier, INRAE, CIRAD, IRD, Institut Agro, Montpellier, France
2- Department of Entomology, University of Wisconsin, Madison, USA
3- Department of Entomology and Nematology, University of California, Davis, USA
4- Pathologie végétale, INRAE, Avignon, France
5- CBS, Univ Montpellier, CNRS, INSERM, Montpellier, France
6- MIVEGEC, Univ Montpellier, CNRS, IRD, Montpellier, France
# These authors equally contributed to this work
‡ This work is dedicated to the memory of Thomas L. German

*** Corresponding author:** stephane.blanc@inrae.fr

This PDF file includes:

Main Text
Figures Legends
(Figures 1-4, Supplemental Table S1 and S2, Figure S1 and Movie S1 are supplied as separate files)

Key words: *Orthotospovirus*, genome packaging, *Bunyavirales*, *Tomato spotted wilt orthotospovirus*, TSWV, genome formula, RNA segment, RNA polarity, strand-specific RT-qPCR, Nanopore Direct RNA sequencing

Competing Interest Statement: The authors of this article declare that they have no financial conflict of interest with the content of this article.

Classification: Biological Sciences / Microbiology / Virology / Population biology

Abstract

Bunyaviruses are enveloped negative or ambisense single stranded RNA viruses with a genome divided into several segments. The canonical view depicts each viral particle packaging one copy of each genomic segment in one polarity named the viral strand. Several opposing observations revealed non-equal ratios of the segments, uneven number of segments per virion, and even packaging of viral complementary strands. Unfortunately, these observations result from studies often addressing other questions, on distinct viral species, and not using accurate quantitative methods. Hence, what RNA segments and strands are packaged as the genome of any bunyavirus remains largely ambiguous. We addressed this issue by first investigating the virion size distribution and RNA content in populations of the tomato spotted wilt virus (TSWV) using microscopy and tomography. These revealed heterogeneity in viral particle volume and amount of RNA content, with a surprising lack of correlation between the two. Then, the ratios of all genomic segments and strands were established using RNA sequencing and RT-qPCR. Within virions, both plus and minus strands (but no mRNA) are packaged for each of the three L, M and S segments, in reproducible non-equimolar proportions determined by those in total cell extracts. These results show that virions differ in their genomic content but together build up a highly reproducible genetic composition of the viral population. This resembles the genome formula described for multipartite viruses, with which some species of the order *Bunyavirales* may share some aspects of the way of life, particularly emerging properties at a supra-virion scale.

Significance Statement

Bunyaviruses infect animals, plants, fungi, and protists. Despite their importance for global health, fundamental aspects of their biology as basic as the definition of their genome remain elusive. The viral genome consists of several negative or ambisense RNA segments, and virions often miss segments and/or package complementary strands. We formally quantify this heterogeneity on the species *Tomato spotted wilt orthotospovirus*. Within individual virus particles, the number, the identity and even the polarity of the segments is widely variable. In contrast, we show that a stable genetic composition is an emerging property of the viral population, each of the RNA segments/polarities accumulating reproducibly at a specific frequency. This resembles the genome formula of multipartite viruses, suggesting that bunyaviruses may also function as multicomponent viral systems.

Main Text

Introduction

The order *Bunyavirales* was established by the ICTV to contain viral species with 2-8 single-stranded RNA genomic components with negative (sequence complementary to coding strand) or ambisense (coding in both orientations) polarity (1), (2). The order is divided into 14 families that include nearly 500 virus species infecting vertebrates, invertebrates, plants, fungi and protists (2, 3). Many of the well-characterized member species of this order are considered emerging viruses and cause serious diseases impacting human health and agriculture productivity globally (4, 5). Except for the family *Hantaviridae*, all bunyaviruses infecting animals and plants are transmitted by insect or tick vectors in which the virus can also replicate, complicating efforts to control disease spread.

Despite the importance of this viral order, some very basic aspects of the biology of the most important species remain mysterious. To illustrate this statement, one can consider the very simple question of what the genome of these viruses is, *i.e.* what is effectively packaged and transmitted to new hosts. The canonical view describes the production of membrane-bound glycoproteins that together delineate a spherical particle. Each particle is considered to enclose one of each of the negative or ambisense RNA segments. Packaged together these segments constitute the complete genome that is transmitted to a new host, putatively as a "one is enough" particle. The current terminology resulting from these assumptions describes the virion contained RNAs as viral sense (vRNAs), and the unencapsidated complementary RNAs generated during genome replication as viral complementary sense (vcRNAs). The literature, including the most prominent Virology web sites, is replete with cartoons depicting packaged genomic segments of vRNA complexed with the N protein that interacts with the inner domain of the glycoproteins associated with the membrane (for example see Refs(6, 7)). A few copies of the polymerase, the L protein, are also shown associated with the genomic RNA segments within particles, as this protein is required for transcription of negative vRNA strands prior to protein production.

Many experimental findings contradict this dogma (8). First, the viral particles have been reported to be of highly variable size (9, 10) and, in the case of Lacrosse virus (LACV), this

103 particle size variation has been speculated to correlate with RNA content variation (9). Second,
104 some semi-quantitative observations of the relative amounts of the distinct RNA segments in
105 purified virus particles of different *Bunyavirales* species have indicated that they are not equi-
106 molar (11–17), ruling out a presumed efficient packaging of exactly one copy of each segment
107 per virion. Third, in numerous instances such as tomato spotted wilt virus (13) (TSWV),
108 Uukuniemi virus (18) (UUKV) and rift valley fever virus (19, 15) (RVFV), viral complementary
109 sense strands of the S and M RNAs have been convincingly detected within virions, though in
110 low amount not formally quantified. Whether the encapsidation of the vcRNAs is functionally
111 relevant or whether it is an exception or the rule in the order *Bunyavirales* is thus far unclear.
112 Fourth, the possibility of packaging some viral mRNA has been envisaged (13, 15), though
113 experimental evidence appears to plead against it. Finally, the identity of RNA segments and
114 their number within individual virus particles of RVFV has been directly studied via an elegant
115 single molecule *in situ* hybridization approach (15, 17). Distinguishing between vRNAs and
116 vcRNAs polarities (15) or not (17), these two seminal studies very convincingly showed, for the
117 first time with a direct approach, that virions differ in their RNA content, some missing segments
118 whereas others having extra copies (20), and that the majority of virus particles within a
119 population are not able to express the complete viral genetic information. It is interesting to
120 note that the same studies demonstrated that particles containing incomplete segment sets are
121 nevertheless capable of entering host cells; hence, they could putatively have a functional role
122 through complementation. Consistent with this, earlier papers described genetic engineering of
123 artificial segments of UUKV that could not ensure the complete replication cycle but were
124 nevertheless packaged and possibly transmitted to neighboring cells (21, 22). Thus, despite
125 considerable evidence that the packaging of genome segments of bunyaviruses is highly
126 heterogeneous, which RNA species are packaged (segment and polarity), in what proportions,
127 whether these proportions are reproducible at the virus population level and whether the
128 packaging process is non-specific, are still open questions that have never been formally and
129 quantitatively addressed together on a single species.

130 Here, using the plant infecting TSWV, we explore these questions with several complementary
131 quantitative approaches. With few variations, the molecular biology of TSWV is similar to many
132 member species of the order *Bunyavirales* infecting animals (5), with three segments (L, M, and

S) of negative (L) or ambisense polarity (M and S). Using (cryo-)electron microscopy and tomography we establish the virus particle size distribution. We further characterize the amount of RNA content of viral particles and demonstrate that RNA content and particle size do not correlate. Using Nanopore direct RNA sequencing we show that both vRNAs and vcRNAs of all three genomic segments are packaged, though in different proportions depending on the segments. In contrast, messenger RNAs are not packaged, not even in minute amounts. The development of a strand-specific Reverse Transcription (RT) qPCR protocol allowed us to thoroughly quantify the number of copies of each segment and polarity in a large number of samples, indicating that each RNA species is packaged in very reproducible proportions across infected plants thereby defining a "genome formula" at the population level. Finally, the genome formula obtained from infected plant total extracts is very similar to that obtained from virus particles isolated from these extracts, formally supporting a non-specific packaging process where all "available" viral genomic RNAs, of both polarities, are packaged with the same efficiency. Together, by demonstrating that the genome is not recapitulated in individual but in a population of virions, our results highlight important issues of the biology of bunyaviruses, at least of TSWV, that are reminiscent of the biology of multipartite viruses.

Results

Virus particle size and RNA content are highly variable and do not correlate

To accurately estimate the diameter of TSWV virions, we recorded cryo-electron microscopy (cryo-EM) images of the frozen-hydrated purified virus preparation (Figure 1A). With this observation method, the sample is immediately frozen/vitrified in a cryogenic liquid and imaged at very low temperature, avoiding distortion of the morphology of biological objects. From cryo-EM images, we estimated the diameter of 4861 purified TSWV particles, which appear as near perfect spheres. The large number of analyzed virions confirmed an important heterogeneity in diameter that ranges from 65 to 135 nm, unimodally and near-symmetrically distributed around an average of ~100 nm (Figure 1B). From the diameter, the estimated volume of the spherical TSWV particles ranges from 1.44×10^{-10} to 1.29×10^{-9} nanoliters, thus an impressive volume difference close to one order of magnitude.

When observing semi-thin sections of TSWV viral particles fixed and embedded into a classical electron microscopy resin (Figure 1C), not only did we consistently observe the heterogeneity of their diameter but, interestingly, also of their inner densities which most likely represent ribonucleoprotein (RNP) segments (also see Supplemental Figure S1 and Supplemental Movie S1). Tomographic analysis allowed the 3D reconstruction of the volumes corresponding to these inner densities (Figure 1C3 and 1C4). Remarkably, virus particles could contain 1, 2, 3 or 4 discrete volumes. We relegate arguments on whether each volume represents a distinct segment or not to the Discussion section.

Summing up all volumes of the inner density within a given virus particle provides an estimate of the amount of packaged RNPs. We extracted tomograms from 25 randomly chosen particles with the only constraint that they are entirely embedded in the resin (not cut through by the section) and plotted their diameter against the calculated total inner density volume. Particle size and the amount of encapsidated genetic material are not correlated (Figure 1 D).

Direct Nanopore RNA sequencing reveals the packaging of both viral and complementary strands for the three genomic segments but not of messenger RNAs

After demonstrating the huge variation in the volume of packaged genetic material within individual virions (four to fivefold variation in Figure 1D), we further explored the potential heterogeneity in the identity/nature of RNAs. In other words, which viral RNA species (identity and polarity of viral segments, viral mRNA) could be packaged in virions and in what proportion? We purified virus particles from a single infected plant and extracted the viral RNA. Direct Nanopore RNA sequencing was performed as indicated in the Methods section and all viral reads, either of viral or complementary polarity, were mapped to the sequence of S, M and L segments of TSWV isolate MR-01 (Figure 2A). According to this technology (23, 24), sequencing is always initiated at the 3' extremity of the RNA and proceeds towards the 5' extremity with an efficiency that decreases with length, *i.e.* the reads can be stopped for many uncontrolled reasons at any time and so regions lying further from the 3' end have less chance to be reached in a given read. We further calculated the sequencing depth that illustrates the number of times a given nucleotide of the RNAs has been read in an experiment, and doing so we could obtain a quantitative estimate as described in the Methods and as further considered in the Discussion

section. The trends within this sample and with this technique were: i) when summing both polarities RNA-L appeared as the least accumulated (~1200 copies for which the sequencing was initiated), RNA-M was slightly more frequent (~2200 copies) and RNA-S clearly dominated (~11,000 copies); ii) contrary to the canonical view outlined in the Introduction, the viral complementary polarity can be highly represented in packaged RNA. While the vcRNA-L is rare and appears here as 3% of total segment L, the vcRNA-S accounts for 10% of segment S and vcRNA-M is even more frequent than the so-called viral strand (52 vs 48%, respectively).

An important feature of this Direct Nanopore RNA Sequencing approach is its potential to detect the presence of mRNA in purified virions. Indeed, because TSWV mRNAs acquire their cap through a cap-snatching process that adds 12-20 nucleotides snatched from host mRNAs (25, 26), it is possible to scrutinize all available 5' sequences of the viral RNA extracted from purified virions and check for the presence of a stretch of 5'-terminal non-viral (host) nucleotides. Recent detailed studies of Direct Nanopore RNA sequencing reported a loss of the processive control of motor proteins when the end of the RNA molecule is released from the pore, resulting in an arrest of the sequencing process 11 nucleotides from the 5' extremity (23, 27, 28). We looked at all reads reaching the 5' region of the TSWV RNAs for all segments and polarities. The number of reads per nucleotide position crashes from about 1200 to 14 between nucleotide position 15 and 10 from the 5' extremity (Figure 2B), indicating an absence of cap-snatching related extension expected in viral mRNAs. Moreover, among the 14 reads that extended beyond the 5'extremity of the viral RNAs, none could be interpreted as a cap-snatched sequence. The first reason is that all but one of these 5'-extensions were much too long, between 56 and 1035 bases; the remaining one was only 6 nucleotides long. The second reason is that, when blasting solely the first 30 nucleotides of these 14 extensions (Figure 2C), we could not find a single significant match which could indicate the presence of a short sequence snatched from a host mRNA (BLASTn searches against Viridiplantae or Solanaceae accessions). Our conclusion is thus that the TSWV virus particles do not package even minute amounts of viral mRNA.

Strand-specific RT-qPCR confirms the presence of viral and complementary strands of the three segments each at a specific frequency in a population of TSWV particles

As further debated in the Discussion section, the quantitative accuracy of Direct Nanopore RNA sequencing is not established, and thus we had to confirm our conclusion with an unrelated and accurate quantitative technology. We developed and thoroughly validated a strand-specific RT-qPCR approach (see Methods section) and first applied it to the viral RNA sample used for Direct Nanopore RNA Sequencing for comparison (Figure 3A). The result that is strikingly consistent across the two techniques is the presence of the viral complementary strands for all three segments, with the lowest proportion for L (~10% of total L), followed by S (~30% of total S) and an astonishing near 60% of total M. Although the general trend is the same, it is noticeable, that the frequency estimates of the distinct viral RNA species differ between Direct Nanopore RNA Sequencing and RT-qPCR. This specific technical point will be addressed in the Discussion section.

The presence of large proportions of viral complementary strands was a surprise. We ascertained that the RNA extracted from virus particles with our protocol does not contain considerable amounts of "free" RNA in solution or RNA externally associated to viral membranes. We repeated the purification of virus particles from another infected plant and treated half of it with RNase prior to RT-qPCR, in order to eliminate all RNA not protected in virus particles. The estimates of the relative proportions of all 6 viral RNA species obtained before and after this treatment are quasi-identical (Figure 3B), confirming that the viral complementary strand of all three segments is indeed packaged. The total amount of RNA was however drastically reduced by this treatment (four times less RNA after treatment), suggesting that the RNase can also degrade packaged RNAs though likely at a slower pace. We repeated this experiment after adding *in vitro* transcripts of a region of the vRNA-M, encompassing the amplified region, to the purified virus suspension. The added amount of soluble RNA should increase the vRNA-M copy number 84 times. This artificially enlarged proportion of vRNA-M totally disappeared after RNAse treatment confirming that the additional free RNA has been completely degraded. In this experimental repeat, the total amount of all packaged RNA species was also decreased to nearly a third, again showing that the RNA packaged into virions can be degraded by prolonged RNAse treatments, but much more slowly than free RNA. Of note is the fact that the degradation of packaged RNA by the RNase treatment is similarly acting on all RNA species, as their relative proportion remains constant (Figure 3B).

The viral and complementary strands of the three TSWV segments are all packaged non-specifically

When comparing the two preparations of purified virus particles in Figure 3A and 3B, each from a distinct infected plant, it is noticeable that the relative frequencies of the six viral RNA species (designated hereafter as the genome formula) are alike, though some slight differences are visible. This observation immediately calls for two questions. Is the genome formula reproducible across virus populations from distinct individual host plants? And if yes, does it stem from deterministic selective/specific packaging mechanisms or does it simply reflect the amount of each of the six viral RNA species accumulated in plant tissues during the infection cycle? To answer these questions, we ground one infected leaf from each of 24 host plants infected in parallel. The resulting 24 crude extracts were then each split in two parts, one for total RNA extraction and the other for virus particles purification followed by RNA extraction. All samples were then submitted to our strand-specific RT-qPCR protocol and the genome formula in total infected plant tissues and in purified virions extracted thereof were compared (Figure 4).

In total plant extracts, when summing up the number of copies of a given segment for both viral and complementary strands, we obtained an L/M/S ratio of approximately 1/1/3. This overall segment ratio can be compared with most semi-quantitative estimates found in the literature for other bunyaviruses (for example see Refs(21),(11–13, 17)), where the viral and complementary RNA strands were not distinguished. In contrast, accurately quantifying the two polarities for each segment is more complex and such a comprehensive genome formula, established in a large set of host individuals infected in parallel, has not been reported before for any species in the order *Bunyavirales*.

In extracts from purified virus, the total number of copies of all 6 viral RNA species was 72 times lower than in leaf extracts. This difference may in part stem from a loss of virions during the purification process, but it also suggests that high amounts of viral RNA may exist in the host plant in a non-packaged form. The genome formula of viral RNAs packaged in virion populations appears strikingly similar to that of total RNA in infected plant tissues. For one out of the six RNA species (for vM), the relative frequency estimated from total plant extracts and from purified virion preparations is not statistically different. For the other five RNA species (vL, vcL, vcM, vS

and vcS) differences are statistically significant but always of very a small magnitude, still consistent with the genome formula being similar in total plant extracts and in purified virions.

Another way to compare the proportion of the viral RNA species in infected cells and in purified virus particles is to test for a possible correlation between the relative frequency of a given viral RNA species in the 24 plant extracts and in the corresponding 24 populations of virions. Remarkably, a positive and statistically significant correlation was revealed in 3 out of 6 cases: vL, vS and vcS. The linear regression equation and the results of the tests for these three RNAs are: **vL**: Freq_vL(particles) = 0.0958832 + 0.7074039*Freq_vL(plants), R^2=0.28, test on slope: t-ratio=2.91, Prob>|t|=0.0081; **vS**: Freq_vS(particles) = 0.2432724 + 0.4215708*Freq_vS(plants), R^2=0.27, test on slope: t-ratio=2.84, Prob>|t|=0.0095; **vcS**: Freq_vcS(particles) = 0.0194624 + 0.4119993*Freq_vcS(plants), R^2=0.84, test on slope: t-ratio=10.68, Prob>|t|<0.0001).

Together, these results indicate that the viral and complementary strands of each of the three TSWV genome segments are accumulated at specific and reproducible levels during the infection of a given host plant species, and that they are packaged non-specifically preserving a similar frequency pattern at the level of the viral population.

Discussion

As detailed in the introduction, numerous previous reports scattered across multiple virus species in the order *Bunyavirales* have provided evidence or hinted at the possibility that virions in this group often deviate from the canonical view by not packaging exactly one copy of each genomic segment, and further that viral complementary strands may be packaged. How often and to what extent the real situations deviate from this canonical view, and how this could affect the understanding of major aspects of the biology of bunyaviruses remained largely unclear. Earlier studies consistently revealed an important heterogeneity of the genetic material contained in individual virions. Hence, it is unclear what is to be considered as the genome of these viruses and whether it should be defined at the virion level or, more meaningfully, at a larger scale. Using TSWV as a model species, we have confirmed and formally quantified the stunning heterogeneity of virus particles and characterized the RNA species that can be packaged and in what proportions. In addition, we have demonstrated that the genetic

heterogeneity at the virion level is structured in a reproducible manner at the population level, extending the concept of genome formula, developed to better describe the biology of multipartite viruses, to the TSWV and perhaps beyond, to the *Bunyavirales* large group of segmented viruses.

The size distribution of viral particles has been earlier estimated for two other bunyaviruses, LACV (9) and UUKV (10). The amplitude of the size variation is comparable in all three bunyaviruses (including TSWV studied here) and higher than that of other segmented viruses which were shown to possess specific packaging mechanisms of their genome segments, such as for example Influenza virus A (29) and rotaviruses (30). Notably the size variation estimated here for TSWV is even more important than that of birnaviruses (31) which have been shown to encapsidate their two genomic segments in a non-specific manner, often packaging more than one copy of each (32). The size distribution appeared unimodal in previous studies on bunyaviruses and, in one case, was qualified as “normal”(10). However, the number of particles used for these estimates was relatively low, respectively 350 and 165, and the corresponding graphs seemed to reveal more particles that were smaller than the modal value. In our experiment, the size distribution of 4861 TSWV particles follows a near-symmetrical distribution, indicating that the number of smaller particles is similar to larger ones, and that both become equally infrequent as the distance to the modal value increases (Figure 1B). Intriguingly our results do not support earlier speculation that virion size variation correlates with the amount of RNA content (9). We observed remarkable variability in the quantity of packaged RNA (volume of electron density within virions), with some particles containing a volume of genetic material four times larger than others, even among particles with the smallest diameters between 70 and 75 nm (Figure 1D). This lack of correlation between particle size and RNA content is consistent with the observation that particles of UUKV can form without encapsidating any genetic material (22, 33). A comparison of the size distribution of empty particles may reveal that diameter and volume of virions are determined by the sole physical properties of the association between membranes and glycoproteins. It is also possible that the packaging of the viral genetic material depends on the amount of RNA that is locally and temporarily available. In any case, this observation further extends the heterogeneity of packaged RNA, as this heterogeneity is also large among particles of the same size class. In some of the observed viral particles, the

tomographic analysis reveals poorly resolved distinct inner density volumes within a single virion, each potentially corresponding to a distinct segment. We observed from one to four such distinct volumes per virion, consistent with the recent demonstration that particles of RVFV can package a very variable number of segments(15, 17). In contrast, we did not observe any particle with a complete absence of inner density, suggesting that empty particles are very rare in TSWV and in our experimental set up while they appear very frequent in RVFV (15). One important caution is that, due to low resolution of our tomographic analysis, the assignment of each distinct inner density volume to one individual segment may not be totally reliable. Heterogeneous local compactions (34, 35) and the extreme flexibility (35) of a single segment may sometimes appear as more than one distinct volume upon tomography 3D reconstruction, and two closely packed/entangled segments may appear as a single larger volume. Higher resolution cryo-EM tomography will be required to solve the question of the spatial arrangement of the segments and to reveal the range of the identity and number of segments potentially co-packaged in individual virions.

The packaging of viral complementary sense RNA has been reported for segment S of UUKV (18), for segment M of LACV (36), for segments M and S of TSWV (13), and more recently for segments S, M and L of RVFV (19),(15). In these papers, the possible biological impacts of packaging the two polarities of a segment have been discussed, the most obvious one being a putative benefit of early expression of some of the genes due to immediate availability of both strands. This would concern primarily ambisense segments, where the ratio of vRNA versus vcRNA may make one or the other of the two genes encoded in opposite orientation expressed early or late (18, 19). It seems less evident, however, for negative sense RNA segments for which the vcRNA strand (which is not capped) cannot be translated before *de novo* transcription from the vRNA where cap-snatching occurs (thus cannot be translated earlier than vRNA). Consistent with this argument, the vcRNA of the negative sense segment L of TSWV could not be detected within virions in earlier studies (13). In contrast, the vcRNA of the negative sense segment M of LACV was detected (36), leading the authors to conclude on a possible artifactual misincorporation. Going beyond detection, some of these studies provided “semi quantitative” estimates. For example, Kormelink and colleagues clearly mentioned that the M segment was the one with the highest amount of packaged vcRNA for TSWV, though still largely lower than

the vRNA. Here with two distinct technologies, Direct Nanopore RNA Sequencing and strand-specific RT-qPCR, we establish that TSWV packages both polarities of its three segments L, M and S. The quantification of all 6 RNA species by these two techniques reveals very consistent trends, with the abundance of packaged vcRNA being M > S > L. However, the exact relative copy number varies considerably depending on the technique and this is further discussed in the next paragraph. Of note here is the general consensus that mRNAs of bunyaviruses are not significantly packaged into virions because they lack the sequence driving their interaction with (and coating by) the nucleoprotein (37),(19),(13, 15, 18). However, with the techniques used earlier for mRNA detection within purified virus preparations, minute amounts of subgenomic mRNA might have been overlooked and hence this was still considered a possibility (15). Here we used the direct RNAseq approach to examine the 5' extremity of the packaged viral RNAs, in search for snatched host sequences. Out of >1200 RNA 5'-extremities analyzed, not even a single mRNA molecule could be identified, confirming that TSWV does not package mRNAs, at least not at a level that could be considered as biologically significant.

Because direct RNAseq and strand-specific RT-qPCR did not reveal the same proportion of vRNA *versus* vcRNA, particularly for segments L and S, we set up an experiment intended to control the accuracy of the strand-specific RT-qPCR (Method section and Figure 3A). Conceptually, it is very simple and consists of producing a solution containing a mix of the 6 RNA species from *in vitro* transcripts in a known ratio, here 1/1/1/1/1/1, and submit it to the RT-qPCR. The resulting estimates were close to the expected relative frequency of 16.6% for each of the six RNAs, the largest deviation being an observed frequency of 20 % (Figure 3A grey horizontal band). We believe these deviations are small insofar as they may arise because the ratios in the initial control mix were not exactly equal. Indeed, quantifying transcripts in solution by spectroscopic methods is poorly accurate due to variable sequence-related secondary structures engendering highly variable proportions of single or double stranded regions. For this reason, and because the quantitative power of the Direct Nanopore RNA Sequencing is not established (24, 27, 28), in this and future work we consider the strand-specific RT-qPCR approach as the more accurate and reliable.

In any case, as revealed with both RNAseq and RT-qPCR, the vcRNA-M is extremely frequent, even more frequent than the vRNA-M. According to RT-qPCR estimates, vcRNA-S and vcRNA-L

400 respectively represent 30% and 8% of the total corresponding segment RNAs. With such high
401 amounts of packaged vcRNAs, our results again question what should be considered the viral
402 genome, v- or vc-strand. More interestingly, they also render unlikely the possibility that
403 packaged vcRNAs are biologically irrelevant. Beyond previous speculations, on a regulatory role
404 for gene expression kinetics at the early stages of infection(13, 18, 19), numerous additional
405 questions are raised: can v and vcRNA of a same segment be co-packaged in the same virion or
406 are they necessarily separated in distinct particles and only together at the population level?
407 Does a virion with the vRNA of a given segment and another virion with the vcRNA of the same
408 segment have similar infectivity? Does this also depend on the polarity of the other segments in
409 the same virion? Do the putative distinct properties of individual virions make any sense at the
410 population level through complementation and/or cooperation? Our findings provide a
411 foundation opening important issues for future research on the order *Bunyavirales*, beyond
412 TSWV and the family *Tospoviridae*.

413 Among the above non-exhaustive list of possible questions, those that appear particularly novel
414 are at the population level. If v- versus vcRNA segment ratios have a functional role, this can
415 only be conceived at the population level. It is indisputable that the RNA content of individual
416 virions is highly heterogeneous in RVFV (15, 17) and as indicated here in TSWV. We therefore
417 made use of the TSWV system to investigate the genetic composition at the population level.
418 First, in all plants of one species, so in one given environment, TSWV populations produce a very
419 similar frequency pattern for the six viral RNA species. This situation is remarkably reminiscent
420 of the genome formula described for multipartite viruses (38–40), where distinct parts of the
421 genetic information are packaged separately in distinct virus particles, but very reproducibly
422 accumulate at the population level. While a phenomenon such as the genome formula with
423 unequal ratios could be anticipated for other bunyaviruses within infected cells with
424 unpackaged genomic segments (14, 41), the situation was less predictable in purified virions as
425 it depends on whether the packaging process constrains these ratios. Some earlier studies
426 reported comparable segment frequency patterns in infected tissues and in purified virions(13,
427 15, 17), whereas others noted a variability of this pattern depending on the virus particle
428 preparation (14, 18). To resolve this issue in TSWV, we used a large set of infected plant
429 replicates to compare the genome formula in infected leaves and in virions from these leaves.

Beyond revealing very similar formulas, more convincing was the demonstration that the small variations of a viral RNA frequency observed across the leaf extracts correlate with those observed across the corresponding virus preparations. Despite an unavoidable experimental "noise" induced by purification steps, reverse transcription and qPCR, such correlations could be detected for three out the six analyzed RNA species (vL, vS and vcS). This experimental noise and the low variations of the less frequent segments/polarities may explain that correlation was not detected in the three other cases. These unique observations empirically confirm that the viral RNA species accumulating during the infection cycle within a host are packaged according to their frequency, thus with no apparent specific mechanisms that would bias the ratios.

In the same view as that currently developing for multipartite viruses(40, 42), one could imagine that bunyaviruses can have emerging properties at the population level (all virions being different) that would be driven by the genome formula. Whether the genome formula controls gene expression (14, 39), whether it is adjustable depending on the environment (15, 43) (including within insect thrips vectors), whether virions with distinct and incomplete genetic information can be transmitted cell-to-cell and host-to-host separately (44–46) , and most of all whether they can proficiently complement in natural situations, are all new research horizons for the biology of bunyaviruses.

Materials and Methods

Plants, growth conditions, inoculation, and sampling

All plants were seeded and maintained in a growth chamber with a 16/8 hours day/night photoperiod, a constant temperature of 20 °C, and 70% hygrometry. The tomato spotted wilt virus isolate used in this study is the TSWV MR-01 (47), maintained as a -80°C frozen stock of *Datura stramonium* (L.) leaves collected from plants initially infected through transmission by the natural thrips vector *Frankliniella occidentalis* (Pergande). *Nicotiana rustica* (L.) plantlets at the sixth leaf stage were mechanically inoculated on their fourth leaf, by rubbing a virus suspension prepared from the frozen stock with 1 g of infected leaves ground in 3 ml of 0.01 M phosphate buffer (pH 7.5) with the addition of Carborundum or Celite abrasive powder. Three

459 weeks post-inoculation, 1 gram of the sixth leaf of each systemically infected plant was collected
460 for further RNA or virus particle purification.

461

462 *Total RNA extraction and virus particle mini-prep protocol*

463 One gram of infected leaf tissue was homogenized for 20 seconds with 4 ceramic beads in 3 ml
464 EB (Extraction Buffer: 0.1 M sodium phosphate, pH 7, containing 0.01 M Na_2SO_3) in a 15 ml
465 Falcon tube, using the FastPrep MP grinding machine. Of these crude extracts, 200 µl could be
466 used for total RNA extraction using the Monarch Total RNA Miniprep kit (NEB #T2010, final
467 elution in 50 µl RNase free water and storage at -80°C) according to manufacturer instructions,
468 and the rest for further virus purification as described hereafter. Crude extracts were filtered
469 through a Luer Terumo syringe stuffed with Miracloth and centrifuged at 4°C and 10,000g for 15
470 minutes. The supernatant was discarded, and the pellet was rinsed with RB buffer (Rinse Buffer:
471 0.01 M sodium phosphate buffer, 0.01 M Na2SO3, pH 7) and rapidly resuspended in 2 ml RB
472 using a tissue grinder. After 15 minutes of 8,000 g clarification at 4°C, the supernatant was
473 collected and passed through a 5 ml Terumo Syringe containing a ClearLine Syringe 30mm AC
474 0.2 µm filter. Two to 2.5 ml of the filtrate were then overlayed on a 10-40% sucrose gradient,
475 prepared in an Ultra Clear Tube for the SW41 rotor, and centrifuged at 70,000 g for 1 hour at
476 4°C in a Beckman SW41 rotor. The opalescent band appearing in the gradient was collected,
477 diluted with RB 1/12 in a 13 ml polycarbonate tube, and pelleted for 1 hour at 70,000 g and 4°C.
478 The pelleted virions were finally resuspended in 150 µl sterile double-distilled water. When
479 larger amounts of purified virions were necessary, several samples could be pooled together in
480 the same 15ml falcon tube after the Miracloth first filtration step. In some instances, viral RNA
481 was extracted from purified virions exactly as from plant crude extracts.

482

483 *Cryo-electron microscopy and estimation of the diameter of virus particles*

484 Three microliters of freshly purified TSWV particles were overlayed onto a glow discharged
485 Quantifoil R 2/2 grid covered with a 2–3 nm ultra-thin carbon layer (Quantifoil Micro tools GmbH,
486 Germany), blotted for one second to eliminate the extra-volume, and then flash frozen in liquid
487 ethane using a CP3 cryo-plunge (Gatan inc.). Before freezing, the humidity rate was stabilized at
488 about 95 %. Cryo-EM was carried out on a JEOL 2200FS FEG microscope operating at 200 kV under

low-dose conditions (total dose of 20 electrons/Å^2), in the zero-energy-loss mode with a slit width of 20 eV. Images were taken at a nominal magnification of 50,000X, corresponding to a calibrated magnification of 45,591X with defocus ranging from 1.5 to 3.0 µm, using a 4k x 4k slow scan CCD camera (Gatan, inc.). The diameter estimation of TSWV particles was performed using ImageJ (48).

Electron tomography and segmentation

Three hundred µL of purified virus particles suspension were fixed by addition of glutaraldehyde (5% final concentration) and incubated for 45 min at 4°C. Fixed virions were then diluted in 8 ml of RB and ultracentrifuged at 31,000 g for 1 hour at 4°C to eliminate the fixing agent. The pellet was resuspended in 50 µl RB and added with a small volume of Chinese ink to dark stain the sample. One hundred µl of 3% low melting agarose (35°C) were added and homogenized before cooling overnight at 4°C. The jellified dark block was then extracted from the tube and post-fixed for 1 hour at 4°C with 1% osmium tetroxide, rinsed three times in 0.1 M Cacodylate buffer, and dehydrated in a series of acetone solutions as follows: two 10-minutes bathes of each of 50, 70, 90, 100% acetone at 4°C. The dehydrated agarose block was imbedded at room temperature in 25% TAAB resin in acetone for 3 hours, in 50% TAAB resin in acetone overnight, in pure resin for at least three hours, and transferred in silicon molds filled with resin. The resin block was finally polymerized one night at 56°C. Thin sections 200nm thickness were prepared using an ultramicrotome (LKB ultratome NOVA), and contrasted for 20 minutes in 2% Uranyl Acetate followed by 3 minutes in 2% Lead Citrate.

Series of tilted images from semi-thin sections of embedded TSWV particles were collected on a JEOL 1400 Flash electron microscope, operating at 120 kV at a magnification of 15,000X, using a Oneview camera (Gatan, inc.) with defocus values ranging from -3 µm to -5 µm. The acquisition was performed semi-automatically using SerialEM(49). Tilt series consisting of about 60 images were collected by tilting the specimen between -60° and 60°, imaging at 2° increments, with a total dose of about 700 electrons/Å^2. Image alignment and three-dimensional reconstructions were performed using Etomo from IMOD package(50). Segmentations of particles were done using Amira software (Thermo Fisher Scientific).

Direct Nanopore RNA sequencing

Total RNA was isolated from 600 µl of purified virions suspension using the Monarch Total RNA Miniprep kit (NEB #T2010) and quantified using the Qubit 3.0 fluorometer (Life Technologies, Carlsbad, CA). RNAs were poly(A) tailed using the *Escherichia coli* Poly(A) Polymerase NEB M0276 kit (New England BioLabs, Beverly, MA, USA), and further purified using paramagnetic beads supplied by the RNAClean XP kit (Beckman Coulter, Brea, CA, USA). The Nanopore sequencing library was constructed using the Direct RNA Sequencing Kit (SQK-RNA002), following manufacturer's instructions. Direct sequencing of the poly(A) tailed RNAs was then performed on one R9.4.1 Flongle flow cell (FLO-FLG001) using MinKNOW software 22.03.5.

Accurate base calling was performed using Guppy (v6.1.2; Oxford Nanopore Technologies, available online: https://nanoporetech.com/) in super accuracy mode (sup) and consensus filtered for a minimum Q-score quality of 7. Adapter removal was then achieved using Porechop v0.2.4 (available online: https://github.com/rrwick/Porechop). Quality of the reads was investigated using NanoPlot v1.40.0(51). Cleaned reads were further mapped in the two orientations with a High/Medium Sensitivity on the three segments of tomato spotted wilt virus isolate MR-01 (segment S, accession number MG593199.1; segment M, accession number MG593198.1; segment L, accession number MG593197.1) using Geneious 5.5.9 (Biomatters Limited). Coverage of each base of each segment was then recovered. Finally, mapped reads were filtered into two groups, one large group of reads with no additional bases at the 5'-end of the viral genomic RNA, and a very small group of only 14 reads having an extension at the 5'-end of the viral genomic RNA.

As a quantitative approach, we simply counted the number of reads that were initiated at the 3'-end of each of the distinct viral RNAs. We considered that this represents the number of copies of a given RNA species entering a sequencing pore, which may reflect its frequency in the sequencing library. To the best of our knowledge, how distinct RNA sequences could bias this quantitative estimate due to secondary structures or to other reasons is thus far undetermined.

Strand-specific reverse transcription (RT) and quantitative real time PCR (qPCR)

From the extracted viral or total RNA, we aimed at quantifying specifically the viral strand (vRNA) and the viral complementary strand (vcRNA) of each of the three genomic segments S, M and L, avoiding the co-quantification of viral mRNAs whenever possible. Classic RT-qPCR primer design revealed an auto-priming problem. We adapted an approach developed earlier to solve a similar issue with the rift valley fever virus(52). This technique involves use of a unique 20 nucleotide-long Tag-sequence fused at the 5' extremity of the RT primer. Because the Tag-sequence is later used as one of the primer pair designed for qPCR, all cDNA generated by "auto-priming" (i.e. cDNA non-specifically primed by secondary structures or by co-purifying small RNAs produced in the infected host) and thus lacking this Tag-sequence will not be amplified. A common Tag-sequence was used for the six specific RT primers (Supplementary Table S1-A), as well as for all forward primers for subsequent qPCR, the specificity at this latter step being therefore determined by the reverse qPCR primers (Supplementary Table S1-A). It is important to note that the distinct primer sets for RT-qPCR target plus or minus sequence strands either in the 5' or 3' halves of the ambisense M and S segments. For a given polarity, the chosen region was that where the subgenomic messenger RNAs do not overlap, avoiding the co-ampification of the M and S segments together with their mRNA. Unfortunately, for segment L, the mRNA is full length and so our approach cannot distinguish the amplification of the vcRNA-L from that of the mRNA, in total RNA extracts from infected plants.

Before RT, the purified viral or total RNA was diluted 1/100 with RNase free water. Four µl of this dilution was used as a template, added with the RT Tagged-primers (10 µM final concentration), and the mixture was heated at 70°C for 5 minutes to eliminate secondary structures that could hinder the RT efficacy. The mixture was then chilled on ice for 5 minutes, complemented with avian myeloblastosis virus (AMV) reverse transcriptase (Promega, Madison, WI, USA), and processed according to the manufacturer's specifications. After RT, qPCR reactions (40 cycles of 95 °C for 10 s, 60 °C for 10 s and 72 °C for 10 s) were carried out using the LightCycler FastStart DNA Master Plus SYBR green I kit (Roche, NJ, USA) in a LightCycler 480 thermocycler (Roche,NJ, USA), following the manufacturer's instructions. Two µl of the RT products diluted 1/10 were used as qPCR template and the primers were adjusted to a final concentration of 0.4 µM. The fluorescence data were analyzed with the LinReg qPCR program (53). Relative frequencies were then calculated in each sample by dividing the estimated

absolute fluorescence value (N0) of a given segment by the sum of the N0 of all six segments (three cDNA viral sense and three cDNA viral complementary sense).

RT is obviously a key step in the quantification of the different segments and polarities, because in contrast to the qPCR, there is no evaluation and correction for possible different efficacies on different RNA sequence targets. We first verified that multiplexing the 6 RT reactions (vRNA and vcRNA for segment L, M and S) in a single tube did not bias the results of subsequent qPCR on the corresponding cDNAs, as compared to the 6 RTs being carried out separately (Supplementary Table S1-B). This possible bias being relaxed, we then experimentally controlled the accuracy of the quantification of the relative copy number of the 6 distinct viral RNA species. We produced *in vitro* transcripts encompassing the 6 amplified sequences, quantified each of these transcripts with a NanoDrop 2000 (RNA mode), mixed them at equimolar ratio with a total RNA final concentration of 2.92 ng/ml, and estimated their relative copy number with our strand-specific RT-qPCR protocol (see Results section Figure 3A and Discussion).

Plasmids construction and in vitro transcription

cDNAs covering the whole TSWV MR-01 genome sequence were synthesized from TSWV-infected *Nicotiana rustica* (L.) plant leaves with random primers using iScript Reverse Transcription Supermix kit (Biorad, CA, USA). From these cDNAs, PCR primers (Supplementary Table S1-A) were designed to amplify about 450 nucleotides from two distinct regions of genomic RNA L (nucleotide position 599-1003 and 7918-8322), RNA M (291-718 and 2844-3322) and RNA S (nt 261-689 and 2106-2522), using the 2x PCR master mix (Promega, Madison, WI, USA), PCR Reactions were done in a thermocycler (Eppendorf) first set at 95°C 2 min, followed by 30 cycles of 95°C 30 sec, 55°C 30 sec, 72°C 30 sec, and finally 72°C 15 min. PCR products were cleaned using Wizard SV Gel and PCR clean up system (Promega) and cloned into pGEM-T Easy vector syst (Promega). All resulting plasmids had one of the TSWV short sequences indicated above flanked on one side by SP6 promoter and on the other by T7 promoter. SP6 or T7 could later be used for *in vitro* transcription depending on the orientation of the cloned TSWV sequence and the desired polarity of the transcript. The orientation and integrity of each of the cloned TSWV sequences was verified by sequencing.

606 Three to five µg of each of these plasmids were linearized by digestion with Sal I or Nco I (New
607 England Biolabs, MA, USA), followed by enzyme inactivation at 80°C for 20 minutes.
608 Transcriptions to generate the desired plus strand or minus strand RNAs were done using SP6
609 (Nco I digests) or T7 (Sal I digests) using Maxiscript kit (Invitrogen), according to the
610 manufacturer recommendation. The template plasmids were then eliminated by DNAse1
611 (RNase free) treatment (Invitrogen, USA), and the RNA finally precipitated with 0.5 M
612 Ammonium Acetate and 70% ethanol, rinsed, and stored in water at -80°C.
613
614 *Statistical analysis*
615 The histogram of the diameter distribution of virus particles was fitted to a normal distribution
616 and the significance of the fit was assessed by the Kolmogorov-Smirnov test.

617 To test whether the frequency of the six viral RNA species (segment and polarity) differed
618 between total plant extracts and purified virus particles we performed an analysis of variance
619 using the model logit(frequency_RNA species) = treatment * RNAspecies + 1|replicate plant,
620 where treatment corresponds to total plant extracts vs purified virus particles. A difference in
621 the frequencies of RNA species across treatment modalities is thus indicated by a statistically
622 significant treatment by RNA species interaction. For each RNA species the statistical
623 significance of the difference in frequency across treatment modalities was assessed by Tukey's
624 HSD test. The identity of the replicate plant is introduced in the model as a random effect to
625 account for the fact that each replicate plant yielded both total plant extracts and purified virus
626 particles which are thus statistically paired.

627 The correlation of each RNA species frequency across treatment modalities was tested using
628 linear regressions on the frequencies of the RNA species (linear regressions on the logits of the
629 frequencies yielded very similar results and are not presented).

630 These three analyses were performed using JMP 13.2.1 software (SAS Institute 2016).

631
632
633

Data availability

All microscopic and PCR data produced are visible in the figures and supplemental material. In particular, Supplemental Table S2 contains all segments' relative frequency calculated from RT-qPCR to build Figure 4. Nanopore RNA sequencing data are available at https://www.ncbi.nlm.nih.gov/sra/PRJNA1020112.

Acknowledgments

We are grateful to MUSE Univ Montpellier for funding this work (project BLANC-MUSE2020-Multivir). MY, EV, PG, MSV and SB acknowledge support from INRAE dpt. SPE; YM from CNRS and IRD; PB, JLKH, AA and JG from INSERM; EF, DF and PR from CIRAD. DEU acknowledges support from the Fulbright Scholar Program and sabbatical leave from University of California Davis. TLG acknowledges funding for this work was provided in part by the USDA-ARS Floriculture and Nursery Research Initiative project# 58-6034-5-029. We thank Jean-Marc Leininger and Norma Ordaz for technical support in providing virus isolate. We dedicate this work to the memory of Thomas L. German.

References

1. E. J. Lefkowitz, *et al.*, Virus taxonomy: the database of the International Committee on Taxonomy of Viruses (ICTV). *Nucleic Acids Research* **46**, D708–D717 (2018).

2. J. H. Kuhn, *et al.*, 2022 taxonomic update of phylum Negarnaviricota (Riboviria: Orthornavirae), including the large orders Bunyavirales and Mononegavirales. *Arch Virol* **167**, 2857–2906 (2022).

3. N. S. Akopyants, L.-F. Lye, D. E. Dobson, J. Lukeš, S. M. Beverley, A Novel Bunyavirus-Like Virus of Trypanosomatid Protist Parasites. *Genome Announc* **4**, e00715-16 (2016).

4. H. Boshra, An Overview of the Infectious Cycle of Bunyaviruses. *Viruses* **14**, 2139 (2022).

5. R. Kormelink, J. Verchot, X. Tao, C. Desbiez, The Bunyavirales: The Plant-Infecting

662 Counterparts. *Viruses* **13**, 842 (2021).

663 6. R. M. Elliott, Orthobunyaviruses: recent genetic and structural insights. *Nat Rev*
664 *Microbiol* **12**, 673–685 (2014).

665 7. , Viralzone Tospoviridae.

666 8. P. J. Wichgers Schreur, R. Kormelink, J. Kortekaas, Genome packaging of the
667 Bunyavirales. *Curr Opin Virol* **33**, 151–155 (2018).

668 9. Y. Talmon, *et al.*, Electron microscopy of vitrified-hydrated La Crosse virus. *J Virol* **61**,
669 2319–2321 (1987).

670 10. A. K. Overby, R. F. Pettersson, K. Grünewald, J. T. Huiskonen, Insights into bunyavirus
671 architecture from electron cryotomography of Uukuniemi virus. *Proc Natl Acad Sci U S A* **105**,
672 2375–2379 (2008).

673 11. R. F. Pettersson, M. J. Hewlett, D. Baltimore, J. M. Coffin, The genome of Uukuniemi
674 virus consists of three unique RNA segments. *Cell* **11**, 51–63 (1977).

675 12. R. Pettersson, L. Kääriäinen, The ribonucleic acids of Uukuniemi virus, a noncubical tick-
676 borne arbovirus. *Virology* **56**, 608–619 (1973).

677 13. R. Kormelink, P. de Haan, D. Peters, R. Goldbach, Viral RNA synthesis in tomato spotted
678 wilt virus-infected Nicotiana rustica plants. *J Gen Virol* **73 (Pt 3)**, 687–693 (1992).

679 14. N. Gauliard, A. Billecocq, R. Flick, M. Bouloy, Rift Valley fever virus noncoding regions of
680 L, M and S segments regulate RNA synthesis. *Virology* **351**, 170–179 (2006).

681 15. E. Bermúdez-Méndez, E. A. Katrukha, C. M. Spruit, J. Kortekaas, P. J. Wichgers Schreur,
682 Visualizing the ribonucleoprotein content of single bunyavirus virions reveals more efficient
683 genome packaging in the arthropod host. *Commun Biol* **4**, 345 (2021).

684 16. S. Murakami, K. Terasaki, K. Narayanan, S. Makino, Roles of the Coding and Noncoding
685 Regions of Rift Valley Fever Virus RNA Genome Segments in Viral RNA Packaging. *J Virol* **86**,
686 4034–4039 (2012).

687 17. P. J. Wichgers Schreur, J. Kortekaas, Single-Molecule FISH Reveals Non-selective
688 Packaging of Rift Valley Fever Virus Genome Segments. *PLoS Pathog* **12**, e1005800 (2016).

689 18. J. F. Simons, U. Hellman, R. F. Pettersson, Uukuniemi virus S RNA segment: ambisense
690 coding strategy, packaging of complementary strands into virions, and homology to members of
691 the genus Phlebovirus. *J Virol* **64**, 247–255 (1990).

692 19. T. Ikegami, S. Won, C. J. Peters, S. Makino, Rift Valley Fever Virus NSs mRNA Is
693 Transcribed from an Incoming Anti-Viral-Sense S RNA Segment. *J Virol* **79**, 12106–12111 (2005).

694 20. P. J. Wichgers Schreur, N. Oreshkova, R. J. M. Moormann, J. Kortekaas, Creation of Rift
695 Valley fever viruses with four-segmented genomes reveals flexibility in bunyavirus genome
696 packaging. *J Virol* **88**, 10883–10893 (2014).

697 21. K. Flick, *et al.*, Functional analysis of the noncoding regions of the Uukuniemi virus
698 (Bunyaviridae) RNA segments. *J Virol* **78**, 11726–11738 (2004).

699 22. A. K. Overby, V. Popov, E. P. A. Neve, R. F. Pettersson, Generation and analysis of
700 infectious virus-like particles of uukuniemi virus (bunyaviridae): a useful system for studying
701 bunyaviral packaging and budding. *J Virol* **80**, 10428–10435 (2006).

702 23. R. E. Workman, *et al.*, Nanopore native RNA sequencing of a human poly(A)
703 transcriptome. *Nat Methods* **16**, 1297–1305 (2019).

704 24. T. Wongsurawat, *et al.*, Rapid Sequencing of Multiple RNA Viruses in Their Native Form.
705 *Front Microbiol* **10**, 260 (2019).

706 25. R. Kormelink, F. van Poelwijk, D. Peters, R. Goldbach, Non-viral heterogeneous
707 sequences at the 5’ ends of tomato spotted wilt virus mRNAs. *J Gen Virol* **73 (Pt 8)**, 2125–2128
708 (1992).

709 26. S. Olschewski, S. Cusack, M. Rosenthal, The Cap-Snatching Mechanism of Bunyaviruses.
710 *Trends Microbiol* **28**, 293–303 (2020).

711 27. M. T. Parker, *et al.*, Nanopore direct RNA sequencing maps the complexity of
712 Arabidopsis mRNA processing and m6A modification. *Elife* **9**, e49658 (2020).

28. L. Mulroney, *et al.*, Identification of high-confidence human poly(A) RNA isoform scaffolds using nanopore sequencing. *RNA* **28**, 162–176 (2022).

29. F. P. Booy, R. W. Ruigrok, E. F. van Bruggen, Electron microscopy of influenza virus. A comparison of negatively stained and ice-embedded particles. *J Mol Biol* **184**, 667–676 (1985).

30. D. Asensio-Cob, J. M. Rodríguez, D. Luque, Rotavirus Particle Disassembly and Assembly In Vivo and In Vitro. *Viruses* **15**, 1750 (2023).

31. J. Pous, *et al.*, Structure of birnavirus-like particles determined by combined electron cryomicroscopy and X-ray crystallography. *J Gen Virol* **86**, 2339–2346 (2005).

32. D. Luque, *et al.*, Infectious bursal disease virus is an icosahedral polyploid dsRNA virus. *Proc Natl Acad Sci U S A* **106**, 2148–2152 (2009).

33. A. K. Overby, R. F. Pettersson, E. P. A. Neve, The glycoprotein cytoplasmic tail of Uukuniemi virus (Bunyaviridae) interacts with ribonucleoproteins and is critical for genome packaging. *J Virol* **81**, 3198–3205 (2007).

34. K. Komoda, M. Narita, K. Yamashita, I. Tanaka, M. Yao, Asymmetric Trimeric Ring Structure of the Nucleocapsid Protein of Tospovirus. *Journal of Virology* **91**, 10.1128/jvi.01002-17 (2017).

35. F. R. Hopkins, *et al.*, The Native Orthobunyavirus Ribonucleoprotein Possesses a Helical Architecture. *mBio* **13**, e01405-22.

36. R. Raju, D. Kolakofsky, The ends of La Crosse virus genome and antigenome RNAs within nucleocapsids are base paired. *J Virol* **63**, 122–128 (1989).

37. M. Snippe, R. Goldbach, R. Kormelink, Tomato spotted wilt virus particle assembly and the prospects of fluorescence microscopy to study protein-protein interactions involved. *Adv Virus Res* **65**, 63–120 (2005).

38. A. Sicard, *et al.*, Gene copy number is differentially regulated in a multipartite virus. *Nature communications* **4**, 2248 (2013).

39. Y. Michalakis, S. Blanc, Editorial overview: Multicomponent viral systems. *Curr Opin Virol*

739 **33**, vi–ix (2018).

740 40. Y. Michalakis, S. Blanc, The Curious Strategy of Multipartite Viruses. *Annu. Rev. Virol.* **7**,
741 203–218 (2020).

742 41. C. Rossier, R. Raju, D. Kolakofsky, LaCrosse virus gene expression in mammalian and
743 mosquito cells. *Virology* **165**, 539–548 (1988).

744 42. A. Sicard, Y. Michalakis, S. Gutiérrez, S. Blanc, The Strange Lifestyle of Multipartite
745 Viruses. *PLoS Pathog* **12**, e1005819 (2016).

746 43. R. Gallet, *et al.*, Gene copy number variations at the within-host population level
747 modulate gene expression in a multipartite virus. *Virus Evol* **8**, veac058 (2022).

748 44. A. Sicard, *et al.*, A multicellular way of life for a multipartite virus. *Elife* **8** (2019).

749 45. E. Bermúdez-Méndez, *et al.*, Incomplete bunyavirus particles can cooperatively support
750 virus infection and spread. *PLOS Biology* **20**, e3001870 (2022).

751 46. J. Di Mattia, *et al.*, Nonconcomitant host-to-host transmission of multipartite virus
752 genome segments may lead to complete genome reconstitution. *Proc Natl Acad Sci U S A* **119**,
753 e2201453119 (2022).

754 47. N. A. Ordaz, *et al.*, The Sw-5b NLR immune receptor induces earlier transcriptional
755 changes in response to thrips and mechanical modes of inoculation of Tomato spotted wilt
756 orthotospovirus. *Mol Plant Microbe Interact* (2023) https:/doi.org/10.1094/MPMI-03-23-0032-R.

757 48. C. A. Schneider, W. S. Rasband, K. W. Eliceiri, NIH Image to ImageJ: 25 years of image
758 analysis. *Nat Methods* **9**, 671–675 (2012).

759 49. D. N. Mastronarde, Automated electron microscope tomography using robust prediction
760 of specimen movements. *J Struct Biol* **152**, 36–51 (2005).

761 50. J. R. Kremer, D. N. Mastronarde, J. R. McIntosh, Computer visualization of three-
762 dimensional image data using IMOD. *J Struct Biol* **116**, 71–76 (1996).

763 51. W. De Coster, S. D'Hert, D. T. Schultz, M. Cruts, C. Van Broeckhoven, NanoPack:

764 visualizing and processing long-read sequencing data. *Bioinformatics* **34**, 2666–2669 (2018).

765 52. B. Tercero, K. Terasaki, K. Nakagawa, K. Narayanan, S. Makino, A strand-specific real-
766 time quantitative RT-PCR assay for distinguishing the genomic and antigenomic RNAs of Rift
767 Valley fever phlebovirus. *J Virol Methods* **272**, 113701 (2019).

768 53. J. M. Ruijter, *et al.*, Amplification efficiency: linking baseline and bias in the analysis of
769 quantitative PCR data. *Nucleic Acids Research* **37** (2009).

770

771

Figures legends

Figure 1: Particles of TSWV are highly heterogeneous in size and RNA content

A. Cryo-EM image of frozen-hydrated purified virus particles.

B. Histogram of the diameter distribution of virus particles, n = 4861. The red line illustrates the best fit to a normal distribution. Although not passing the Kolmogorov-Smirnov test (p-value = 0.035), because of very slight over-representation of large and under-representation of small values, the distribution is unimodal and approximately symmetrical around the average of 100 nm. This histogram results from two independent virion preparations where the diameter of 1964 and 2897 particles was respectively estimated.

C. Electron tomography and segmentation of virus particles embedded in resin and semi-thin sectioned. The arrow indicates one reconstituted virus particle visible in the semi-thin section. The other particles, each constituting one line of the lower panel, were chosen from other regions of the semi-thin section. Micrographs labeled C1: Central section of the electron tomogram reconstruction of a complete viral particle. Micrographs labeled C2, C3 and C4 are different representations of the segmentation of 3D reconstructions of viral particles. C2 shows the surface of reconstructed virus particles, C3 combines the central section of the viral particles with the segmentation of their inner density volumes (IDV) which are isolated and slightly magnified in C4. Extensive views of viral particles from the semi thin section in C are presented as Supplemental Figure S1 and Supplemental Movie S1.

D. Analysis of the relationship between the diameter of the particles and the total volume of their inner density (the latter most likely corresponding to packaged ribonucleoproteins). Pearson's correlation coefficient (at the top) is low, and a linear regression is not statistically significant indicating an absence of correlation.

Figure 2: Direct RNAseq of viral RNAs extracted from purified virus particles

A: Sequencing depth (scale is indicated within each individual panel) per nucleotide position for L, M and S segments. The three genome segments in the viral orientation are illustrated by a colored thick line under the graphs where nucleotide numbering is indicated. Sequencing reads for the viral (v) and viral complementary (vc) strands are respectively shown in blue (top panels) and red (bottom panels). Row data of the RNAseq are available at https://www.ncbi.nlm.nih.gov/sra/PRJNA1020112

B: Sum of the number of reads at nucleotide positions from 50 to 0 from the 5' extremity of all six RNA species shown in A. The sequencing proceeds from the 3' towards the 5' end of the molecule, so from right to left on the graph. The sequencing depth per nucleotide position appears stable when the sequencing progresses from position 50 to position 15 towards the 5' extremity, but totally crashes between position 15 and position 10, demonstrating an arrest of the sequencing process when approaching the 5' extremity.

C: Among the > 1200 reads reaching the 5'-terminal regions of any of the six viral RNA species (B), only 14 extended beyond position 0 (extension length from 6 to 1035 bases). This panel shows trimmed 30 nucleotide-long sequences extended immediately upstream of the 5' end of the viral segment in these 14 cases. Note that the bottom sequence had only 6 nucleotides while all others had more than 30.

Figure 3: Relative frequency of each RNA segment and polarity packaged in purified virions

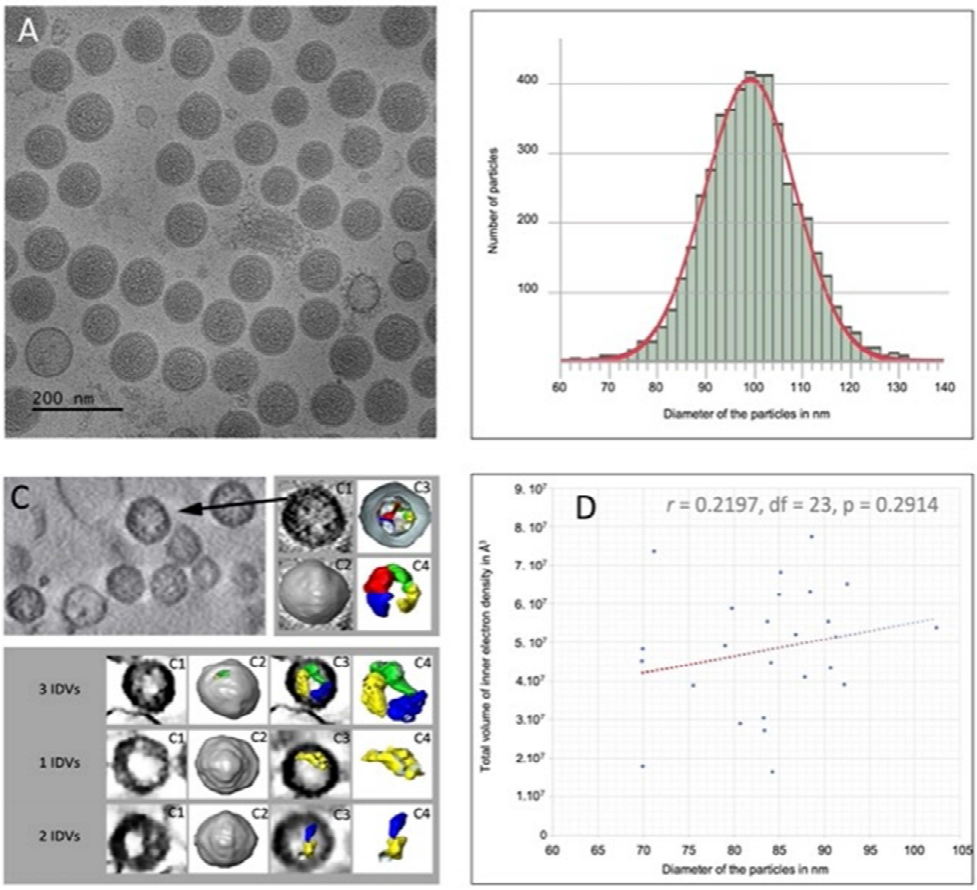
A: The relative proportion of each of the viral RNA segments and polarities was estimated by strand specific RT-qPCR in a purified TSWV particles preparation (the very same preparation as that sequenced in Figure 2). The estimate was repeated with four distinct dilutions of the same purified virions preparation: 1 - 1/10 - 1/100 - 1/1000 and the variance engendered by repeated measures and/or dilutions is presented as boxplots. The grey band indicates the range of values within which we estimated the relative frequency of 6 artificial *in vitro* transcripts, each corresponding to the RT-qPCR targeted region of a segment-polarity, after attempting to mix them at equal ratio. The respective estimate for each of the transcripts within this artificial mix is indicated by a colored dot.

B: The relative proportion of each of the viral RNA segments and polarities was estimated in another purified TSWV particles preparation, obtained from a different infected host plant, either with (pink) or without (blue) RNAse treatment as indicated in the Results section.

Figure 4: Comparison of the relative frequency of each RNA segment and polarity in packaged and total viral RNA

The relative proportion of each of the viral RNA segments and polarities was estimated by strand specific RT-qPCR either in total RNA extracts from infected plant leaves (Total; n = 24) or in virions purified from these leaves (Virions; n = 24). The results are represented as box plots with the horizontal bar indicating the median value of the distribution. The identity of the viral RNA is indicated above the graph. Statistical analysis was performed as indicated in the Methods section. The grey shaded panel indicates the RNA species for which the difference in frequency observed in total RNA extracts and in purified virions is not statistically significant. The segments frequencies calculated from RT-qPCR on these 48 samples are available in Supplemental Table S2

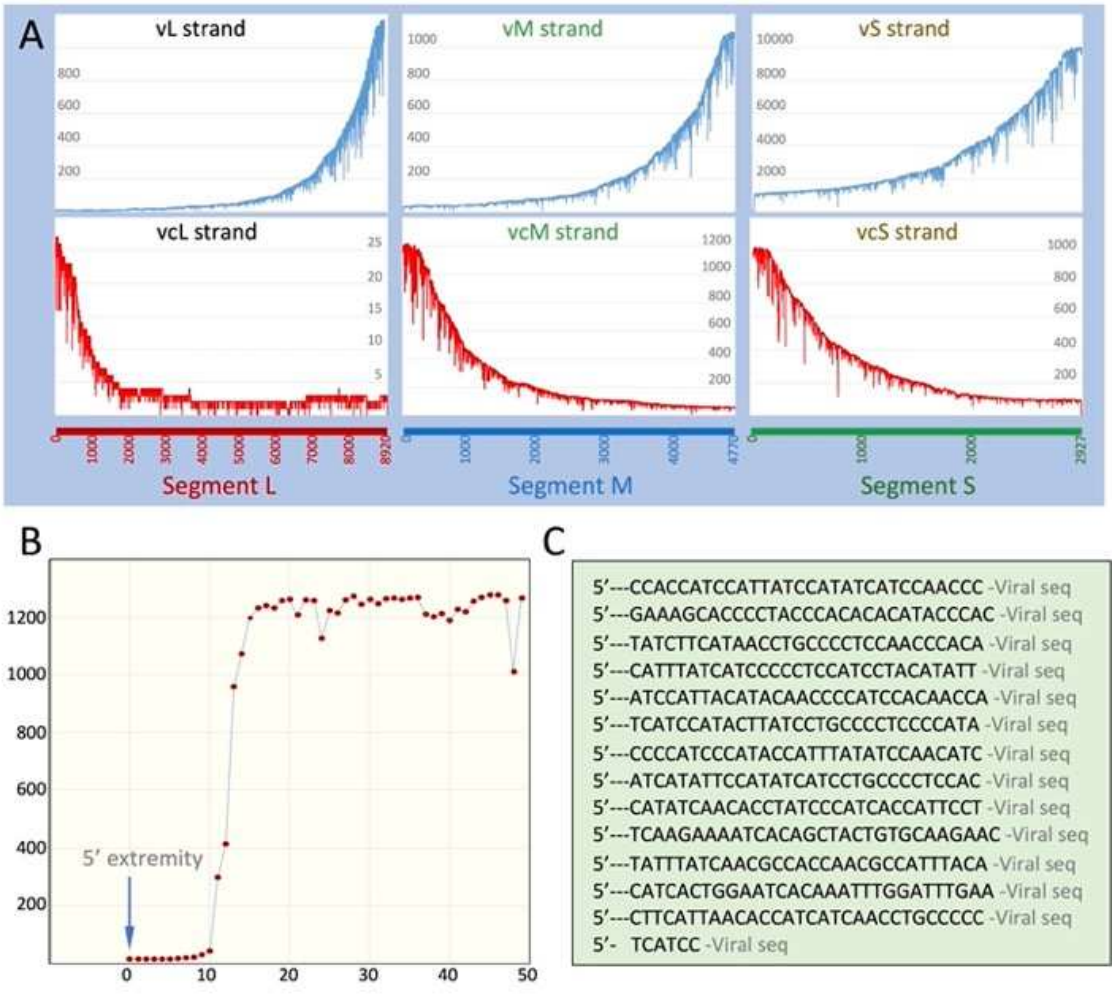
843 **Figure 1**



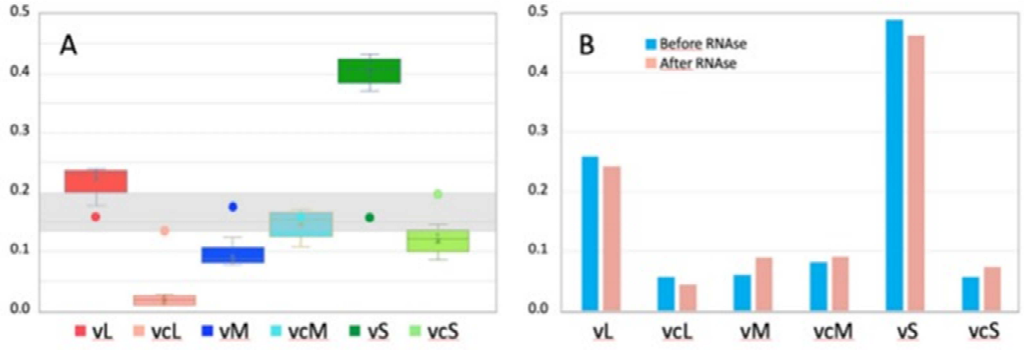
844

845

Figure 2


A
vL strand
vM strand
vS strand
vcL strand
vcM strand
vcS strand
Segment L
Segment M
Segment S
B
5' extremity
C
5'---CCACCATCCATTATCCATATCATCCAACCC -Viral seq
5'---GAAAGCACCCCTACCCACACACATACCCAC -Viral seq
5'---TATCTTCATAACCTGCCCCTCCAACCCACA -Viral seq
5'---CATTTATCATCCCCCTCCATCCTACATATT -Viral seq
5'---ATCCATTACATACAACCCCATCCACAACCA -Viral seq
5'---TCATCCATACTTATCCTGCCCCTCCCCATA -Viral seq
5'---CCCCATCCCATACCATTTATATCCAACATC -Viral seq
5'---ATCATATTCCATATCATCCTGCCCCTCCAC -Viral seq
5'---CATATCAACACCTATCCCATCACCATTCCT -Viral seq
5'---TCAAGAAAATCACAGCTACTGTGCAAGAAC -Viral seq
5'---TATTTATCAACGCCACCAACGCCATTTACA -Viral seq
5'---CATCACTGGAATCACAAATTTGGATTTGAA -Viral seq
5'---CTTCATTAACACCATCATCAACCTGCCCCC -Viral seq
5'- TCATCC -Viral seq

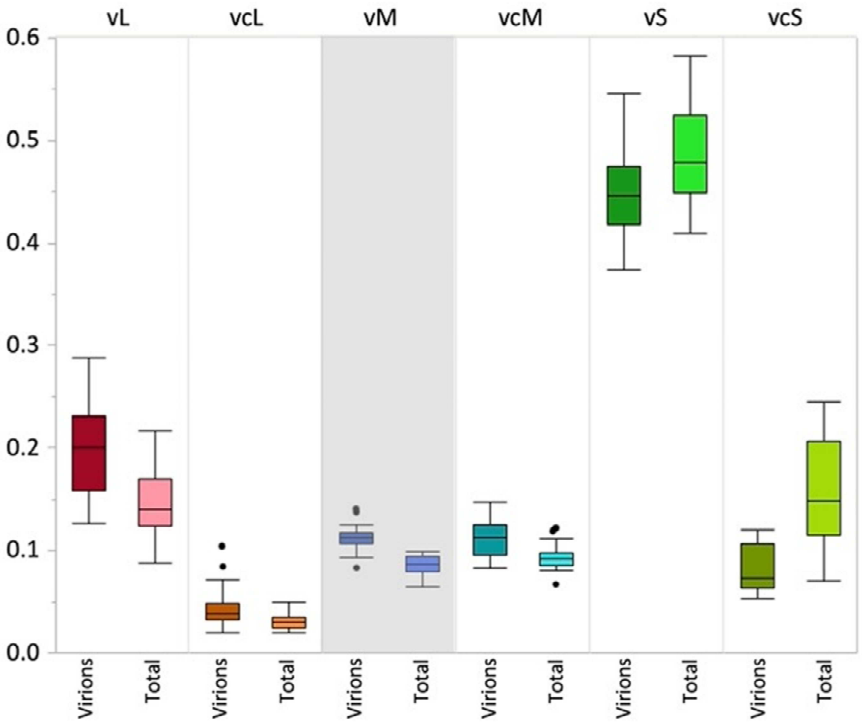
849 **Figure 3**



850

851

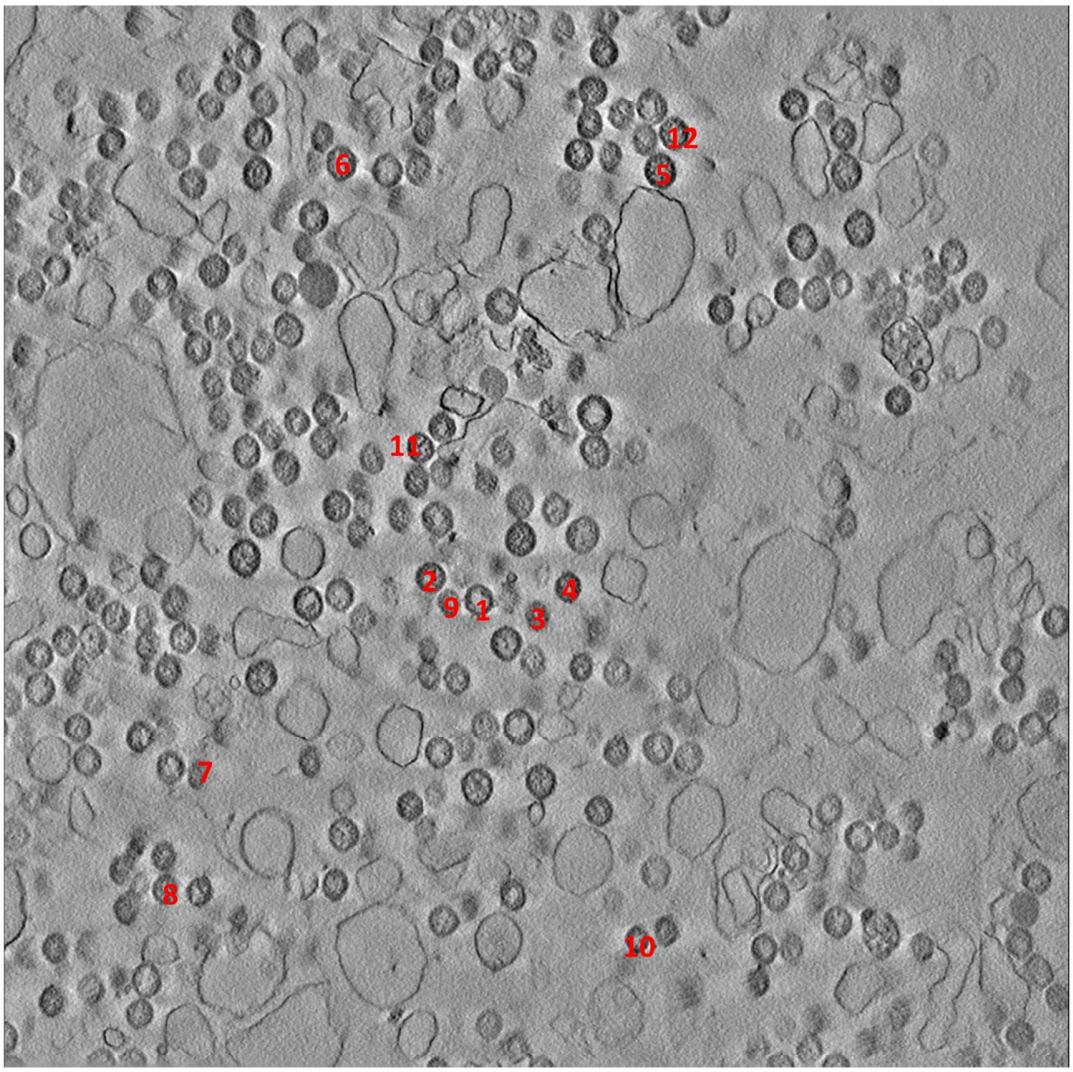
852 **Figure 4**



853

854

Supplemental Figure 1



856

857 **Supplemental Table S1**

Table S1-A: Sequence of primers used for cloning, reverse transcription, PCR, and qPCR

Oligonucleotide Name	Purpose	Nucleotide sequence (5'-3')
TSWV-SRNA-F2106	cloning	GGATTGCTGGAGCCAAGTATAG
TSWV-SRNA-R2522	cloning	GATTGGAGCCACTGACATGA
TSWV-SRNA-F261	cloning	CACTGCAAAGGTAGGGAATCT
TSWV-SRNA-R689	cloning	GCTTGAAACAGTACAGCCATTC
TSWV-MRNA-F2844	cloning	GCAGAATGTTCTCTGGTATGT
TSWV-MRNA-R3322	cloning	GTCTCAAATGCCCATGTCTAT
TSWV-MRNA-F291	cloning	CATCTTCCTTTGGAACCTATGA
TSWV-MRNA-R718	cloning	GTACTCTTGGCTGCACATCAA
TSWV-LRNA-F599	cloning	CTGTCGATGCCAGGACTGAA
TSWV-LRNA-R1003	cloning	TGCTGGTCCAGTGACTAA
TSWV-LRNA-F7918	cloning	TGCTGGTCCAGTGACTAA
TSWV-LRNA-R8322	cloning	CTGTCGATGCCAGGACTGAA
TSWV-vS-RT-Tag	cDNA/RT	**GGCCGTCAACGTGGCGCCTAA**GCTTCTCACCCTTTGATTC
TSWV-vcS-RT-Tag	cDNA/RT	**GGCCGTCAACGTGGCGCCTAA**TTTGTCTCCTGCTCAGCTCC
TSWV-vM-RT-Tag	cDNA/RT	**GGCCGTCAACGTGGCGCCTAA**TGCGTTGCACACCCATCTAT
TSWV-vcM-RT-Tag	cDNA/RT	**GGCCGTCAACGTGGCGCCTAA**TTTTGTTGGCAACGGGAAGC
TSWV-vL-RT-Tag	cDNA/RT	**GGCCGTCAACGTGGCGCCTAA**CTGTGGGTGTGGTTCCAACT
TSW-vcL-RT-Tag	cDNA/RT	**GGCCGTCAACGTGGCGCCTAA**CCTATTGCTGGGATGCCGAT
Tag-Forward-qPCR	qPCR	**GGCCGTCAACGTGGCGCCTAA**
TSWV-vS-qPCR	qPCR	ATAGCCAAGACAACACTGATC
TSWV-vcS-qPCR	qPCR	GAGCCCTTGTTTGCAGCTTC
TSWV-vM-qPCR	qPCR	GCCTCATACAGACCAACCGA
TSWV-vcM-qPCR	qPCR	GGGTTTGGTATAGTGGGGCA
TSWV-vL-qPCR	qPCR	CCTATTGCTGGGATGCCGAT
TSWV-vcL-qPCR	qPCR	CTGTGGGTGTGGTTCCAACT

The non-viral Tag sequences are shown in bold face
Primers where used :
(i) to generate PCR products that have later been cloned to produce plasmids for in vitro transcription (cloning)
(ii) for the production of cDNAs through reverse transcription (cDNA/RT)
(iii) for qPCR

Table S1-B: comparison of single RT and multiplexed RT

	vRNA-L	vcRNA-L	vRNA-M	vcRNA-M	vRNA-S	vcRNA-S
RT made with individual primers in individual tubes	0.31	0.29	0.07	0.09	0.21	0.04
RT made with all six primers in a single tube	0.28	0.30	0.07	0.09	0.22	0.04

The same RNA extract from infected plants was used as a template for RT-qPCR in all cases
Values represent the relative proportion of each viral RNA species in the RNA extract as estimated by RT-qPCR