Replication and Expression of the Tospoviral Genome

Frank van Poelwijk,[1] Peter de Haan,[2] Marjolein Kikkert,[1] Marcel Prins,[1] Richard Kormelink,[1] Marc Storms,[1] Jan van Lent,[1] Dick Peters,[1] and Rob Goldbach[1]

[1]Department of Virology, Wageningen Agricultural University, Binnanhaven 11, 6709 PD, Wageningen, The Netherlands; and [2]S & G Seeds, Westeinde 62, 1600 AA, Enkhuizen, The Netherlands

Sequence analysis of the complete, tripartite RNA genomes of two tospoviruses, tomato spotted wilt virus (TSWV) and impatiens necrotic spot virus (INSV), demonstrated that they possess five genes that specify six functional proteins. The negative-stranded L RNA encodes the putative polymerase (TSWV 331.5 kDa; INSV 330.3 kDa), the ambisense M RNA encodes a common precursor to the two glycoproteins (G1 and G2) and a nonstructural protein (NSm), and the likewise ambisense S RNA encodes the nucleocapsid (N) protein and a second nonstructural protein (NSs). These viral proteins are expressed from mRNAs that contain 12–20 nontemplated nucleotides at the 5' ends. This indicates that "cap-snatching" is the mechanism used by the viral polymerase to initiate transcription. Sequence analysis revealed that there was no strict base preference at the endonucleolytic site of the cellular leaders. Whereas the function of NSs, the least conserved tospoviral protein, has remained enigmatic, evidence is accumulating that NSm represents the viral movement protein that is involved in a tubule-guided cell-to-cell movement of nonenveloped nucleocapsids. Protoplasts infected with TSWV or transfected with the NSm gene solely, develop long, NSm-containing tubules that extend from the plasma membrane into the culture medium and are similar to tubules found in plasmodesmata of infected plant tissues. Experiments with transgenic plants confirm that NSm is a plasmodesma-associated protein that can modify intercellular communication in plants.

Although a growing number of distinct tospoviruses are being recognized, most molecular analyses have been performed with TSWV, the most widespread and economically important. This paper discusses current knowledge of the molecular properties of tospoviruses. It focuses mainly on TSWV, but makes reference to other tospovirus species, especially INSV, when appropriate. For both these tospoviruses, complete sequence data are available. This allows comparison of their genome with the genome of animal-infecting bunyaviruses. The most prominent difference between plant- and animal-infecting bunyaviruses resides in the presence of one extra cistron, the NSm cistron, in the M RNA of tospoviruses (which makes this RNA ambisense) and suggests that this gene represents a major evolutionary adaptation of the standard bunyaviral genome to plant hosts. In view of its potential

importance to the establishment of successful infection in plants, special attention is given to the expression of the NSm gene.

Genetic Organization of the Tospoviral Genome

Comparison with animal-infecting bunyaviruses

Like all bunyaviruses, TSWV particles have a lipid envelope that contains two types of envelope glycoprotein, G1 and G2, and encompasses a tripartite RNA genome that is tightly wrapped in N protein subunits and contained 10–20 copies of a large (L) protein (Van Poelwijk et al. 1993), which represents the putative viral polymerase. Computer-assisted sequence analysis of the tospoviral L protein has shown the presence of sequence motifs that are diagnostic for polymerases of negative-strand RNA viruses (de Haan et al. 1991; Tordo et al. 1992). It has also indicated that tospoviruses, in spite of their very distinct host range, are more closely related to the genus *Bunyavirus*, than this latter genus is to the genus *Hantavirus* (de Haan et al. 1991). This may indicate that the introduction of bunyaviruses (tospoviruses) into the plant kingdom is a relatively late evolutionary event. However, on the basis of the ambisense character of the S RNA segment, tospoviruses seem most closely related to the genus *Phlebovirus*, because members of the genera *Bunyavirus*, *Hantavirus*, and *Nairovirus* have a negative-stranded S segment (Elliott 1990; Murphy et al. 1995).

Because tospoviruses most likely derived from animal-infecting ancestors, it is interesting to investigate how these viruses adapted to their new hosts. Comparison of all bunyaviral genomes for which sequence data are available, indicates one very striking difference, i.e., the presence of the extra NSm gene in the M RNA segment. This gene may have a function in the systemic infection of plants.

Comparison among tospoviruses

So far, nucleotide (nt) sequences of the complete genome are available only for TSWV and INSV. For TSWV, a Brazilian isolate (BR-01) has been completely analysed (de Haan et al. 1990, 1991; Kormelink et al. 1992a). For INSV, the sequence of the M RNA was determined for a North American strain (US-01; Law et al. 1992); whereas, the characterized S and L segment were from a Dutch isolate (NL-07; de Haan et al. 1992a; Van Poelwijk et al., unpubl. results). A comparison between the genome of these two viruses as well as of their gene products, is shown in Figure 1 and Table 1. Among the various proteins, the polymerase (L protein) is the most conserved (84% homology in sequence). Both S RNA-coded proteins are the least conserved (N 70% and NSs 69% homology). When groundnut bud necrosis virus (GBNV; Heinze et al. 1995), of which only sequence data for the S RNA are available, is included in the comparison, the NSs protein appears to be by far the most variable protein. This may indicate that this protein is not involved in any basic process like genome replication or transcription. A more detailed

comparison of the L proteins of TSWV and INSV reveals that the INSV polymerase is 10 amino-acid residues shorter because the acidic C-terminal tail found in the TSWV protein is absent (Van Poelwijk et al., unpubl. results).

Transcription of the tospoviral genome

Like all segmented negative-stranded RNA viruses, tospoviruses use cap-snatching to prime the transcription reaction. Primer-extension analysis of partially purified N mRNA showed that this transcript contains 12–20 nontemplated extra nt (Kormelink et al. 1992b). By cloning and sequencing these mRNAs it was recently (Van Poelwijk et al., unpubl. results) shown that these extra sequences are heterogenous and host derived. Thus for the genera *Bunyavirus, Phlebovirus, Nairovirus*, and now *Tospovirus*, viral mRNA synthesis has a common mechanism for initiation. Strikingly, some bunyaviruses, e.g., *Bunyamwera virus* and *Dugbe nairovirus*, show a strict sequence specificity for the 3' terminal nt of the primer; whereas, for others, e.g., *Uukuniemi phlebovirus*, there seems to be no consensus at this position. Our results indicate that TSWV takes an intermediate position in this respect, with some preference (9 out of 20 mRNAs sequenced) for an U residue. This indicates that, though cap-snatching is a commonly used mechanism to prime transcription, there are differences in this process, not only between distinct virus families but even within a given family.

Tospoviral gene products

The five open-reading frames in the tospoviral genome appear to encode six distinct, functional proteins. These are the (structural) N protein, the G1 and G2 (derived from a common precursor, encoded by the M segment), the viral polymerase (L), and the two nonstructural proteins, NSs and NSm (Figure 1, Table 1).

During the viral infection process, it can be anticipated that the N protein not only has a structural function in packaging the genome, but also a regulatory function (as has been found for the influenza virus) in programming the switch between genome transcription and genome replication. The first initiative (Gielen et al. 1991) to create host resistance against TSWV by transgenic expression of the N protein was based on this regulatory function.

The G1 and G2 have their major function during the insect part of the life cycle of the virus. This can be concluded from the observation that tospoviruses tend to loose their envelope and glycoprotein coding capacity during serial, mechanical passage on plants (Resende et al. 1991). The resulting envelope-deficient isolates are still able to infect plants at wild-type speed, but have lost their thrips transmissibility (Wijkamp 1995). Inspection of the primary structure of the G2 glycoprotein reveals the presence of a "cell attachment site" (RGD motif) for both

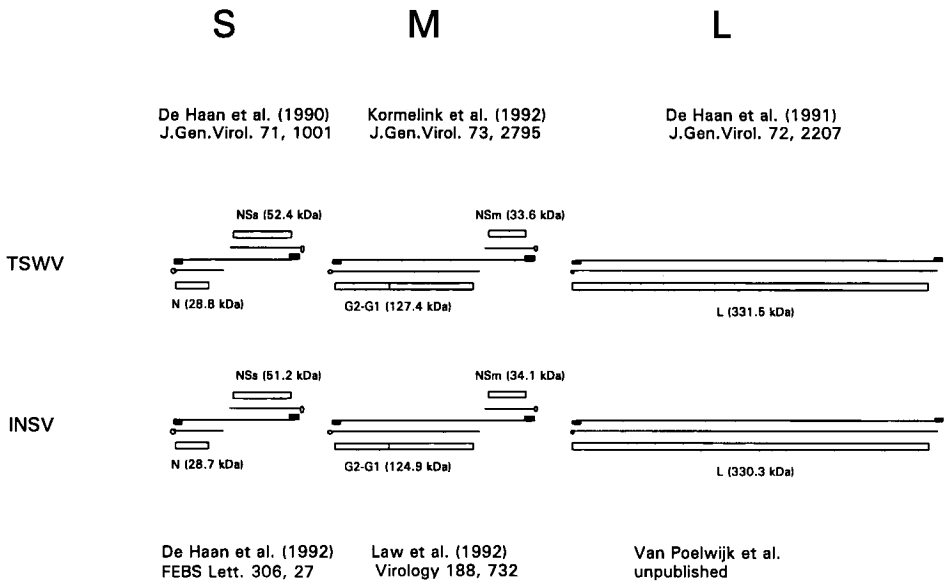

Figure 1. Comparison of the genomes of TSWV and INSV. Genomic RNA segments (with terminal complementary ends indicated as black squares) and viral mRNAs are indicated as single lines, and gene products as open bars. For each genomic RNA species data are taken from the references listed along with the RNA segment. represent the receptor recognition site that is involved in uptake of virus particles by the insect's midgut cells.

INSV (Law et al. 1992) and TSWV (Kormelink et al. 1992). This may represent the receptor recognition site that is involved in uptake of virus particles by the midgut cells of the insect.

The L proteins of all bunyaviruses, and certainly of tospoviruses, are much larger than the core polymerase (protein PB1) of influenza virus. The size of the L protein corresponds with, or even exceeds, the sum of the sizes of the three proteins of the influenza virus that have been implicated in the transcription–replication process (PA, PB1, and PB2). This suggests that the distinct functions of these separate proteins are located in (covalently linked) domains of the single bunyaviral L protein. If this is true, the L protein should possess the endonucleolytic activity involved in the cap-snatching process during transcription initiation. So far, no endonuclease activity has been linked to the tospoviral L protein.

As for the NSs and NSm, hardly any indications for the function(s) of the NSs protein during the viral infection cycle have been obtained. The NSs may occur dispersedly in the cytoplasm of infected plant cells or as fibrous aggregates,

Table 1. Size (kDa) and homology of gene products from TSWV and INSV.

Gene product	TSWV	INSV	Homology (%)	Identity (%)
N	28.8	28.7	70	55
NSs	52.4	51.2	69	52
G2/G1	127.4	124.9	79	65
NSm	33.6	34.1	79	69
L	331.5	330.3	84	70

might indicate that it is not involved in a basic process but rather in a more species-specific event like host-range determination. NSs is also produced during replication in the thrips (Wijkamp et al. 1993), and a function restricted to the insect part of the viral life cycle can not be excluded. There is an accumulating amount of evidence that NSs represents the viral movement protein that is involved in cell-to-cell movement of the virus during systemic infection of the plant.

Involvement of NSm in cell-to-cell movement

By expressing the NSm gene of TSWV in *Escherichia coli*, a specific antiserum can be prepared (Kormelink et al. 1994) for cytological studies. NSm (33.6 kDa) is only expressed for 2–3 days during systemic symptom development (Kormelink et al. 1994). Immunogold studies on ultrathin sections of infected tissues revealed that NSm first associates to free, nonenveloped viral nucleocapsids and later to plasmodesmata, where it assembles into tubular structures similar to the tubules formed by the movement proteins of como-, nepo-, and caulimoviruses (Van Lent et al. 1990, 1991; Perbal et al. 1993; Wieczorek and Sanfacon 1993). It is, therefore, tempting to assume that NSm represents the tospoviral movement protein that is involved in systemic infection of plant tissue. Indeed, among the Bunyaviridae the NSm gene is restricted to tospoviruses, which suggests that it is an adaptation of the prototypic bunyaviral genome to plants. (Some animal bunyaviruses also specify a polypeptide referred to as NSm, but these polypeptides are always part of the glycoprotein precursor and do not represent separate gene products; their function is probably restricted to folding of the glycoproteins.) The viral movement protein is obviously the best candidate for being the product of such gene. The NSm-induced tubules are also formed on the surface of infected protoplasts and on protoplasts transfected with the NSm gene solely (Storms et al. 1995). Because the, basically, negative-stranded nature of the tospoviral genome hampers further studies on the NSm gene by "reverse genetics" (mutational analysis using cDNA clones from which infectivity can be recovered), further attempts to study the biochemical activities of NSm were made by transforming host plants with this gene (Prins and Storms, unpubl. results). Strikingly, it turned out to be very hard to recover NSm-expressing progeny plants. The transformation–regeneration frequency was extremely low, and the few plants that were able to recover had a

stunted phenotype, which indicates that NSm had a great impact on the physiology of the plant. In situ immunogold analysis using anti-NSm serum confirmed a specific accumulation of the transgenically expressed NSm in the plasmodesmata.

Conclusion

Over the past 10 years, an accumulating amount of molecular data on the structure and expression of the tospoviral genome has become available. Not long ago, TSWV was taxonomically set apart as being the sole member of a separate virus taxon, the monotypic "tomato spotted wilt group" (Matthews 1982). Due to increased knowledge of its molecular biology, not only TSWV but also a growing number of previously unidentified plant viruses have been recognized as tospovirus and classified within the family Bunyaviridae. Further molecular studies, have revealed some striking similarities between the tospoviruses and their animal-infecting relatives (e.g., the use of a cap-snatching mechanism for transcription initiation, and the overall structure and genetic organization of their respective genomes).

There are some clear differences at the genus level between the tospoviruses and the other bunyaviruses (e.g., the existence of the extra NSm gene and the ambisense character of two out of the three genome segments). It is tempting to assume that the additional NSm gene represents a genetic adaptation of bunyaviruses to potentiate infection of plant hosts. Indeed, the evidence available points to a function in cell-to-cell movement by facilitating the passage of the virus through plasmodesmata. The NSm-induced plasmodesma-associated tubules resemble the cell-to-cell movement tubules produced by como-, nepo-, and caulimoviruses (Van Lent et al. 1990, 1991; Perbal et al. 1993; Wieczorek and Sanfacon 1993). For these viruses, it has been shown that mature virus particles are transported through the tubules. Although the NSm tubules represent a similar movement structure, this is not likely the mechanism used by the tospoviruses. Three arguments indicate that intercellular movement will occur in the form of nonenveloped nucleocapsid structures. First, the diameter of the enveloped particle (80–110 nm) greatly exceeds that of the tubule (40–45 nm); second, envelope-deficient isolates are able to infect plants systemically at wild-type speed (Resende et al. 1991); and third, NSm binds in situ to nucleocapsids but not to enveloped particles (Kormelink et al. 1994).

This work was supported by the Netherlands Organisation for the Advancement of Pure Science (NWO) with financial aid from its foundations SON and STW.

de Haan, P., Wagemakers, L., Peters, D., and Goldbach, R. 1990. The S RNA segment of tomato spotted wilt virus has an ambisense character. J. Gen. Virol. 71, 1001-1007.

de Haan, P., Kormelink, R., Resende, R. de O., Van Poelwijk, F., Peters, D., and Goldbach, R. 1991. Tomato spotted wilt virus L-RNA encodes a putative RNA polymerase. J. Gen. Virol. 71, 2207-2216.

de Haan, P., de Ávila, A.C., Kormelink, R., Westerbroek, A., Gielen J.J.L., Peters, D., and Goldbach, R. 1992a. The nucleotide sequence of the S RNA of Impatiens necrotic spot virus, a novel tospovirus. FEBS Lett. 306, 27-32.

de Haan, P., Gielen J., Prins, M., Wijkamp, I., Van Schepen, A., Van Grinsven, M.J.M., and Goldbach, R.W. 1992b. Characteristics of RNA-mediated resistance to tomato spotted wilt virus in transgenic tobacco plants. Bio/Technology, 10, 1133-1137.

Elliott, R.M. 1990. Molecular biology of the Bunyaviridae. J. Gen. Virol. 73, 501-522.

Gielen, J., de Haan, P., Kool, A.J., Peters, D., Van Grinsven, M.Q.J.M., and Goldbach, R.W. 1991. Engineered resistance to tomato spotted wilt virus, a negative strand RNA virus. Bio/Technology, 9, 1363-1367.

Heinze, C., Maiss, E., Adam, G., and Casper, R. 1995. The complete nucleotide sequence of the S RNA of a new *Tospovirus* species, representing serogroup IV. Phytopathology, 85, 683-690.

Kitajima, E.W., de Ávila, A.C., Resende, R.de O., Goldbach, R.W., and Peters, D. 1992. Comparative cytological and immunogold labelling studies on different isolates of tomato spotted wilt virus. J. Submicros. Cytol. Path. 24, 1-14.

Kormelink, R., Kitajima, E.W., de Haan, P., Peters, D., and Goldbach, R. 1991. The non-structural protein (NSs) encoded by the ambisense S RNA segment of tomato spotted wilt virus (TSWV) is associated with fibrous structures in infected plant cells. Virology, 181, 459-468.

Kormelink, R. de Haan, P., Meurs, C., Peters, D., and Goldbach, R. 1992a. The nucleotide sequence of the M RNA segment of tomato spotted wilt virus, a bunyavirus with two ambisense RNA segments. J. Gen. Virol. 73, 2795-2804.

Kormelink, R., Van Poelwijk, F., Peters, D., and Goldbach, R. 1992b. Non-viral heterogenous sequences at the 5' ends of tomato spotted wilt virus mRNAs. J. Gen. Virol. 73, 2125-2128.

Kormelink, R. Storms, M., Van Lent, J., Peters, D., and Goldbach, R. 1994. Expression and subcellular location of the NSm protein of tomato spotted wilt virus, a putative viral movement protein. Virology, 200, 56-65.

Law, M.D., Speck, J., and Moyer, J.W. 1992. The M RNA of impatience necrotic spot tospovirus (Bunyaviridae) has an ambisense genomic organization. Virology, 188, 732-741.

Matthews, R.E.F. 1982. Classification and nomenclature of viruses. Intervirology, 17, 1-199.

Murphy, F.A., Fauquet, C.M., Bishop, D.H.L., Ghabrial, S.A., Jarvis, A.W., Martelli, G.P., Mayo, M.A., and Summers, M.D. 1995. Virus Taxonomy. Sixth Rpt. Int. Comm. Taxonomy of Viruses. Arch. Virol. Suppl. 10, 1-586.

Perbal, M.C., Thomas, C.L., and Maule, A.J. 1993. Cauliflower mosaic virus gene

I product forms tubular structures which extend from the surface of infected protoplasts. Virology, 195, 281-285.
Resende, R. de O., de Haan, P., de Ávila, A.C., Kormelink, R., Goldbach, R., and Peters, D. 1991. Generation of envelope and defective interfering RNA mutants of tomato spotted wilt virus by mechanical passage. J. Gen. Virol. 72, 2375-2383.
Storms, M.M.H., Kormelink, R., Peters, D., Van Lent, J.W.M., and Goldbach, R.W. 1995 The nonstructural NSm protein of tomato spotted wilt virus induces tubular structures in plant and insect cells. Virology, 214, 485-493.
Tordo, N., de Haan, P., Goldbach, R., and Poch, R. 1992. Evolution of negative-stranded RNA genomes. Sem. Virol. 3, 341-357.
Van Lent, J., Wellink, J., and Goldbach, R. 1990. Evidence that the Mr 58,000 and 48,000 proteins of cowpea mosaic virus are involved with intercellular movement. J. Gen. Virol. 71, 219-223.
Van Lent, J., Storms, M., Wellink, J., and Goldbach, R. 1991. Tubular structures involved in movement of cowpea mosaic virus are also formed in infected cowpea protoplasts. J. Gen. Virol. 72, 2615-2623.
Van Poelwijk, F., Boye, K., Oosterling, R., Peters, D., and Goldbach, R. 1993. Detection of the L protein of tomato spotted wilt virus. Virology, 197, 468-470.
Van Poelwijk, F., Kolkman, J., and Goldbach, R. 1995. Sequence analysis of the 5' ends of tomato spotted wilt virus N mRNAs. Arch. Virol. 141, 177-184.
Wieczorek, A. and Sanfacon, H. 1993. Characterization and subcellular location of tomato ringspot nepovirus putative movement protein. Virology, 194, 734-742.
Wijkamp, I. 1995. Virus-vector relationships in the transmission of tospoviruses. PhD thesis, Wageningen Agricultural University, Wageningen, The Netherlands.
Wijkamp, I., Van Lent, J.W.M., Kormelink, R., Goldbach, R., and Peters, D. 1993. Multiplication of tomato spotted wilt virus in its vector *Frankliniella occidentalis*. J. Gen. Virol. 74, 341-349.