Relocation of the NIb Gene in the Tobacco Etch Potyvirus Genome

Eszter Majer,* Zaira Salvador,* Mark P. Zwart,* Anouk Willemsen,* Santiago F. Elena,a,b José-Antonio Darós*

Instituto de Biología Molecular y Celular de Plantas, Consejo Superior de Investigaciones Científicas-Universidad Politécnica de Valencia, Valencia, Spain; The Santa Fe Institute, Santa Fe, New Mexico, USA

Potyviruses express most of their proteins from a long open reading frame that is translated into a large polyprotein processed by three viral proteases. To understand the constraints on potyvirus genome organization, we relocated the viral RNA-dependent RNA polymerase (NIb) cistron to all possible intercistronic positions of the Tobacco etch virus (TEV) polyprotein. Only viruses with NIb at the amino terminus of the polyprotein or in between P1 and HC-Pro were viable in tobacco plants.

Viral genomes are compact assemblies of genes and regulatory sequences, and plant viruses are no exception (1). To compact their genomes (2, 3), plant viruses have evolved mechanisms that include overlapping open reading frames (ORFs), ambisense coding, or translational frameshift and read through (4). One of the most common strategies, however, is coding for a polyprotein that is processed into many different gene products after translation. Potyviruses (genus Potyvirus, family Potyviridae) are one of the largest groups of plant viruses and take this strategy to an extreme. Potyviral genomes are RNA molecules of positive polarity approximately 10,000 nucleotides long, consisting of a long ORF flanked by two short untranslated regions (UTRs) (5). The potyviral ORFs apparently encode 10 mature gene products: P1 proteinase; helper component proteinase (HC-Pro); P3 protein; 6K1 polypeptide; cytidylyl incorporation (CI) protein; 6K2 polypeptide; nuclear inclusion a (NIa) protein, a polyprotein that is further processed to, at least, the viral protein genome-linked (Vpg) and the NIa proteinase (NlaPro); nuclear inclusion b (NIb) protein, the viral RNA-dependent RNA polymerase; and coat protein (CP) (6). These products are released from the viral polyprotein through the activity of the three viral proteinases, P1, HC-Pro, and NlaPro. An additional gene product, P3N-PIPO, results from a translation frameshift in a slippery region of the P3 cistron (7).

The aim of this work was to improve our understanding of the functional, structural, and evolutionary constraints that rendered the actual genome organization of potyviruses. To achieve this goal, we relocated one of the genes of Tobacco etch virus (TEV) to all possible intercistronic positions in the genome and assayed the infectivity of the resulting recombinant viruses in tobacco (Nicotiana tabacum L. cv. Xanthi nc) plants. To avoid further complexity, we added appropriate sequences to the amino and carboxy NIb termini to create NlaPro proteolytic sites that mediate NIb release from the viral polyprotein. To facilitate the monitoring of virus infection, spread, and accumulation, all recombinant viruses were constructed from a TEV clone carrying the Rosea1 (Ros1) visual marker (10, 11). The TEV-Ros1 load correlates with anthocyanin accumulation in tobacco tissues (10). Figure 1 outlines the genomes of the parental (TEV-Ros1) and the derived (TEVDNIb-Ros1-NIb1 to -9) recombinant viruses. Figure S1 in the supplementary material specifies the exact nucleotide sequences of all recombinant viral clones.

The parental and recombinant TEV clones, constructed from the binary plasmid pGTEV-Ros1 (10, 11), were agroinoculated (12) into two leaves of 20 3-week-old wild-type plants and 20 transgenic plants constitutively expressing TEV NIb (8). Plants were grown in a glasshouse at 25°C with 16 h light, and Ros1 expression was visually monitored for 4 weeks. For each clone, systemic leaves from three wild-type and three transgenic plants were harvested at 15 days postinoculation (d.p.i.), photographed, and used to estimate viral load by measuring the anthocyanin accumulation induced by the Ros1 marker (10).

Most of the viral clones tested were not viable in either wild-type or transgenic plants. In other words, systemic tissue 4 weeks after agroinfiltration had no visible anthocyanin accumulation or infection symptoms. Only viruses with NIb relocated to the first two intercistronic positions were viable in wild-type plants. Ros1 activity was detected in systemic leaves of all 20 wild-type plants agroinoculated with the virus carrying NIb in the third intercistronic position (TEVDNIb-Ros1-NIb2, as well as the parental virus TEV-Ros1 (Fig. 2A)). There is therefore a statistically significant effect of the NIb position on viability (test of equal proportions, $\chi^2 = 180.00, 8$ df, $P < 0.001$). An identical, statistically significant result was obtained for transgenic plants expressing NIb ($\chi^2 = 155.077, 8$ df, $P < 0.001$), except that approximately half of the plants agroinoculated with the virus carrying NIb in the third intercistronic position (TEVDNIb-Ros1-NIb3) also showed systemic marker expression (Fig. 2B). The number of Ros1-expressing plants for TEVDNIb-Ros1-NIb3 was intermediate, as it was significantly different from the results for all other clones (pairwise test of equal proportions, $P < 0.001$ for all comparisons).

Next, we considered at what time after agroinfiltration anthocyanin accumulation was first apparent. In wild-type plants, there were significant differences in the median time until visual detec-
tion of Ros1 expression for all three viable viruses (Fig. 2A) (log-rank test, \(P < 0.001\) for all three comparisons between TEV-Ros1, TEV\textsubscript{H9004}NIb-Ros1-NIb1, and TEV\textsubscript{H9004}NIb-Ros1-NIb2). So, TEV\textsubscript{H9004}NIb-Ros1-NIb2 was significantly slower than TEV\textsubscript{H9004}NIb-Ros1-NIb1, while TEV\textsubscript{H9004}NIb-Ros1-NIb1 was significantly slower than TEV-Ros1. In transgenic plants, TEV-Ros1, TEV\textsubscript{H9004}NIb-Ros1-NIb1, and TEV\textsubscript{H9004}NIb-Ros1-NIb2 all had the exact same median time until Ros1 expression was observed visually (Fig. 2B). For TEV\textsubscript{H9004}NIb-Ros1-NIb3, anthocyanin accumulation in systemic tissue was first observed significantly later than for any of the other viable viruses (pairwise log-rank test, \(\chi^2 = 11.600, 1\) df, \(P = 0.001\) for all three comparisons). In summary, TEV\textsubscript{H9004}NIb-Ros1-NIb1 and TEV\textsubscript{H9004}NIb-Ros1-NIb2 were viable in both plant genotypes, although infection appeared to proceed more slowly than for the ancestral virus in wild-type plants. For TEV\textsubscript{H9004}NIb-Ros1-NIb3, viral spread was only seen in some transgenic plants, and when it appeared, it was significantly delayed.

We then analyzed the stability of the relocated NIb in the progeny of the viable viruses. For TEV\textsubscript{H9004}NIb-Ros1-NIb1 and TEV\textsubscript{H9004}NIb-Ros1-NIb2, reverse transcription-PCR (RT-PCR) analysis of the viral progeny 15 d.p.i. confirmed the stability of the relocated NIb gene (Fig. 3A and B). For TEV\textsubscript{H9004}NIb-Ros1-NIb3, which showed delayed anthocyanin accumulation in only some transgenic plants, RT-PCR analysis and sequencing confirmed that the NIb gene was lost (Fig. 3C). Anthocyanin accumulation was quantified in extracts from systemic leaves of three plants inoculated per recombinant clone at 15 d.p.i. (Fig. 4). The data were analyzed with a generalized linear model with full-factorial design, using a log-link function and gamma distributed error structure. Overall, the recombinant clone (\(\chi^2 = 2.383, 451, 10\) df, \(P < 0.001\)) and the plant genotype (\(\chi^2 = 35.356, 1\) df, \(P < 0.001\))
had significant effects on anthocyanin accumulation, and there
was a significant interaction between these two factors as well
(\(F_{2,11005} = 419.276, 10\) df, \(P < 0.001\)). A Tukey post hoc test highlights
the existence of three nonoverlapping groups of recombinant
clones (Fig. 4A): TEV\(\text{Nlb-Ros1-Nlb1}\) and TEV\(\text{Nlb-Ros1-Nlb2}\) showed the same expression level as the wild-type TEV-
Ros1, TEV\(\text{Nlb-Ros1-Nlb3}\) showed an intermediate expression
level that depended on the plant genotype (explaining the signif-
icant interaction term), and the other six recombinant clones were
not significantly different from the mock inoculated plants in ei-
ther plant genotype.

These results suggest the existence of many restrictions to the
organization of the potyviral genome. Even though Nlb can be
provided in \textit{trans} (8, 9), it can only be relocated to the amino
terminus of the polyprotein or in between P1 and HC-Pro without
affecting virus viability. The relocation of Nlb to seven intercistronic
positions rendered nonviable viruses, even in a transgenic plant
constitutively expressing TEV Nlb that can be infected by a TEV
mutant with a complete Nlb deletion (8). The relatively late infec-
tion of approximately half of the transgenic plants inoculated with
TEV\(\text{Nlb-Ros1-Nlb3}\) probably resulted from a sporadic recom-
bination event in which the relocated Nlb was deleted. The result-
ing virus (TEV\(\text{Nlb-Ros1}\)) was then able to infect the plant, but
only when Nlb was provided from a transgene. We amplified by
RT-PCR a cDNA fragment corresponding to the HC-Pro/P3 in-
tertronic region from the TEV\(\text{Nlb-Ros1-Nlb3}\) progeny aris-
ing in the Nlb-expressing transgenic plants and cloned it. Se-
quencing of three independent clones showed that recombination
cleanly restored the wild-type HC-Pro/P3 junction in all cases.
Interestingly, no such recombination events were observed in the
case of any of the other recombinant clones. In fact, the inoculated
leaves of both wild-type and transgenic plants inoculated with
TEV\(\text{Nlb-Ros1-Nlb3}\) showed slight anthocyanin accumulation,
suggesting some replication capacity of this chimera.

The relocation of Nlb to seven intercistronic positions proba-

---

**FIG 3** Stability of the relocated Nlb gene in the viral progeny at 15 d.p.i. RNA from systemic leaves of three representative inoculated plants was purified, cDNAs amplified by RT-PCR using primers flanking the new Nlb location for TEV\(\text{Nlb-Ros1-Nlb1}\) (A), TEV\(\text{Nlb-Ros1-Nlb2}\) (B), and TEV\(\text{Nlb-Ros1-Nlb3}\) (C) were separated by electrophoresis in 1% agarose gels, and stained with ethidium bromide. Lanes 1, DNA marker with the sizes of some of the components indicated on the left; lanes 2, noninoculated negative control; lanes 3, parental virus positive control (TEV-Ros1); lanes 4 to 6, three infected wild-type plants; lanes 7 to 9, three infected transgenic plants. Note that in panel C, lanes 4 to 6 (wild-type plants) are missing because none of these plants was infected. The black and gray arrows indicate the positions of the expected RT-PCR products if the relocated Nlb is present or absent, respectively. Primers used to amplify the different cDNAs were, for RT reactions, 5′-CTTTACA TACITTITTTTCCAACATTTCATG-3′, and for PCRs, 5′-AAAAATAAACAAT
CTGAAACACACATATAC-3′ and 5′-CCTTTGCCATGGTGAGCGCG
AC-3′ (A), 5′-CTCAAACACAAGAAATTTC-3′ and 5′-GTCGCGCGCTCACC
CATGGCAAGAG-3′ (B), and 5′-GTTGCAAAATGCAAGATGTTGC-3′
and 5′-GTGTTACCTTCTTGGCCAAAGGCTGAC-3′ (C).

**FIG 4** Anthocyanin accumulation at 15 d.p.i. in systemic leaves of tobacco plants inoculated with TEV recombinant clones in which the Nlb gene was relocated to different intercistronic positions. (A) The absorbance at 530 nm of extracts from systemic leaves from three inoculated wild-type and transgenic plants constitutively expressing TEV Nlb was plotted. Error bars represent ±1 standard deviation among experimental replicates. Letters a, b, and c over the bars represent groups of inoculations with homogeneous accumulation of anthocyanins according to a Tukey honestly significant difference test. (B) Pictures of systemic leaves harvested at 15 d.p.i. from three wild-type and three
transgenic tobacco plants that were noninoculated and inoculated with
TEV-Ros1 and TEV\(\text{Nlb-Ros1-Nlb1}\) to -3 as indicated.

---

Majer et al.

4588 JVI.ORG Journal of Virology
bly results in nonviable viruses because of the induction of lethal defects in polyprotein processing, disruption of partially processed gene products with distinctive roles during infection, or disruption of regulatory RNA elements in the potyviral genome. A related potyvirus (*Turnip mosaic virus*) was found to support the expression of heterologous proteins between HC-Pro and P3 and between 6K1 and CI in *Nicotiana benthamiana* and *Chenopodium quinoa* (13). These differences may be due to the use of different virus species and host plants in both studies. On the other hand, in the previous work, heterologous genes were inserted into the potyviral genome (13), while here, a potyviral gene was relocated along the genome. Processing of the potyviral polyprotein seems to be a finely regulated process that produces the right amounts of the different gene products in time and space (14). Regulation is mainly based on the specific amino acid sequence recognized by the viral NLaPro, with some processing sites being cleaved faster than others (15–17). The insertion of Nib in some of the intercistronic positions of the polyprotein may have fatal effects on this regulation. Partially processed products from the potyviral polyprotein may have distinctive roles in the infectious cycle, different from those of the final processing products. This seems to be the case for the 6K2/VFg/NiaPro polyprotein that has been suggested to anchor TEV replication complexes to cellular endomembranes while recruiting Nib for replication (18). Insertion of Nib may therefore be lethal because it interrupts functional polyproteins. Finally, the potyviral genome contains regulatory RNA elements overlapping the ORF, including a series of RNA hairpins at the end of the CP cistron and 3′ UTR that are involved in TEV replication (19). Disruption of these elements may also have fatal consequences for the virus. Our results highlight the complexity of the potyviral genome organization, suggesting the existence of many more regulatory elements and functional entities than those currently recognized. Our results seem also to reflect the genomic organization of the different *Potyviridae* genera, apparently consisting of two genome blocks, a more conserved block from P3 to the end and the block including P1 and HC-Pro (in the genus Potyivirus) that is more variable between the different genera and even species.

Potyviruses have been used as expression vectors in plants (20–22). Their expression strategy, mainly based on the production of a large polyprotein, makes them particularly attractive to simultaneously produce equimolar amounts of several heterologous proteins (9, 23–25). Our results show the potentials and limitations inherent in expressing heterologous proteins from potyviral vectors. According to our results, the only positions where sequences coding heterologous proteins can be inserted without completely compromising viral viability are the amino terminal end of the polyprotein, between P1 and HC-Pro, and between Nib and CP. The P1/HC-Pro and Nib/CP intercistronic positions have been used with great frequency to express heterologous proteins in many potyviruses (21, 24, 26–28). However, the outermost amino terminal end of the polyprotein has never been used yet, although green fluorescent protein has been successfully expressed close to the amino terminus of potato potyvirus A polyprotein (29).

**ACKNOWLEDGMENTS**

We thank Verónica Aragonés for excellent technical assistance. This research was supported by grant BIO2011-26741 from the Spanish Ministerio de Economía y Competitividad (MINECO) to J.-A.D., grant PROMETEO/2010/019 from Generalitat Valenciana to S.F.E. and J.-A.D., and grants BFU2012-30805 and 22371 from MINECO and the John Templeton Foundation, respectively, to S.F.E. M.E.M. was supported by a predoctoral fellowship (AP2012-3751) from the Spanish Ministerio de Educación, Cultura y Deporte. M.P.Z. was supported by a Juan de la Cierva postdoctoral contract (JCI-2011-10379) from MINECO and a Rubicon grant from the Netherlands Organization for Scientific Research (www.nwo.nl).

The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation.

**REFERENCES**


