

Complete Sequencing of Potato Virus X New Strain Genome and Construction of Viral Vector for Production of Target Proteins in Plants

N. V. Ravin^{1*#}, E. S. Mardanova^{1#}, R. Yu. Kotlyarov^{1#},
V. K. Novikov², J. G. Atabekov², and K. G. Skryabin¹

¹Center of Bioengineering, Russian Academy of Sciences, pr. 60-letiya Oktyabrya 7/1,
117312 Moscow, Russia; fax: (495) 135-0571; E-mail: nravin@biengi.ac.ru

²Faculty of Biology, Lomonosov Moscow State University, 119991 Moscow, Russia;
fax: (495) 938-0601; E-mail: atabekov@genebee.msu.su

Received June 16, 2007

Revision received July 11, 2007

Abstract—The complete nucleotide sequence of the genome of a new potato virus X (PVX) strain Tula isolated by us has been determined. Based on comparison of the PVX Tula nucleotide sequence with the sequences of 12 other PVX strains, this strain was assigned to the European cluster of PVX strains. Phylogenetic analysis revealed the same phylogeny for both full genome sequences and nucleotide sequences of polymerase and coat protein genes, suggesting that the PVX evolution did not involve recombination between different strains. The full-size cDNA copy of the PVX Tula genome was cloned and the accumulation of the viral coat protein in infected *Nicotiana benthamiana* was shown to be about twofold higher than for the PVX strain UK3. Based on the PVX Tula genome, a new vector which contained the target gene instead of the removed triple transport gene block and the coat protein gene has been constructed for expression of target proteins in plants. The productivity of the new vector was about 1.5-2-fold higher than the productivity of the vector of the same structure based on the standard PVX strain genome. The new viral vector can be used for superproduction of recombinant proteins in plants.

DOI: 10.1134/S0006297908010069

Key words: potato virus X, plants as biofactories, viral vector

Production in plants of different proteins for medicine, agriculture, and industry is a promising line of bioengineering. Plants as “biofactories” have some advantages compared to animal cells or microorganisms: plants are much easier to cultivate, decreasing the final price of the product, and there is no common plant and human pathogens, and therefore “plant” proteins are quite safe. The use of self-replicating recombinant viral vectors seems to be the most efficient approach to rapidly produce in plants significant amounts of target proteins. Such vectors are usually constructed based on the tobac-

co mosaic virus or potato virus X, which have rather small genomes and contain plus-strand RNA [1-3]. With viral vectors, target proteins can be produced in plants within a few days to 20-30% of the total soluble protein, which corresponds to 1-2 g per kg plant biomass [4]. Vectors based on phytoviral genomes are successfully used to produce in plants proteins for medicine, including vaccine proteins of agents of rabies [5], plague [6], the thermolabile enterotoxin B-subunit [7], antibodies [8], etc. (for review see [9]).

The efficiency of a viral expression system is determined by the synthesis of its own virus proteins during the infection. A new strain Tula of potato virus X isolated by researchers of the Department of Virology, Faculty of Biology, Lomonosov Moscow State University, effectively replicates in plants and can be used as a base for a new highly efficient vector for expression of target proteins.

The purpose of the present work was to determine the complete nucleotide sequence of the PVX Tula strain, prepare the full cDNA copy of its genome, construct on

Abbreviations: CP) coat protein; GUS) β -glucuronidase; PVX) potato virus X; RDRP) RNA-dependent RNA polymerase; Sgp) subgenomic RNA promoter; 35S) the cauliflower mosaic virus RNA 35S promoter; 35S-T) the cauliflower mosaic virus RNA 35S terminator; *uidA*) *Escherichia coli* β -glucuronidase gene.

* To whom correspondence should be addressed.

These authors have equally contributed to this work.

its basis an expressing vector, and test its efficiency on a model reporter gene.

MATERIALS AND METHODS

Potato virus X strains. The PVX strain Tula was isolated by researchers of the Department of Virology, Faculty of Biology, Lomonosov Moscow State University from potato plants in the Tula district of the Russian Federation. From the initial PVX isolate prepared from an infected plant, some strains were isolated and one of them was named "Tula" and used in the present work. For phylogenetic analysis, full nucleotide sequences of another 11 PVX strains were used: Russian (GenBank M38480 [2, 3]), Korean (GenBank AF373782), X3 (GenBank D00344 [10]), Taiwan (GenBank AF272736), OS (GenBank AB056718 [11]), UK3 (GenBank M95516 [12]), Roth1 (GenBank AF111193 [13]), BS (GenBank AB056719 [14]), HB (GenBank X72214 [15]), CP (GenBank X55802 [16]), CP4 (GenBank AF172259 [13]). Viruses of the CP4, HB, and CP strains were isolated in South America; the Korean, Taiwan, OS, and BS strains were isolated in Eastern Asia; and the Roth1, Russian, UK3, and X3 strains were prepared in Europe.

Media, reagents, enzymes, and synthetic oligonucleotides. Bacteria were grown in LB broth or plates with LB agar at 37°C (*Escherichia coli*) or 30°C (*Agrobacterium tumefaciens*). If required, the media were supplemented with antibiotics: ampicillin (100 µg/ml), kanamycin (50 µg/ml), rifampicin (50 µg/ml), or gentamicin (25 µg/ml). The following synthetic oligonucleotides were used:

1F, GAAAACTAAACCATACACCAACAAC;
 1348F, GAAAC(T/C)TGGGACTTCAGATTCC;
 1651R, GCGGTTT(G/A)GCTTTCTTT(G/A)C(G/A)G;
 2245R, GA(G/A)AGC(T/C)T(G/A)(C/A)C(A/T)GA-(A/G)AGGAAAGT(G/T)GC;
 2566R, GTAGAC(G/A)CT(T/C)TGTCTGCTATCTCC;
 2941F, GCGCGAATGTGATGTACACGG;
 3130F, GT(G/C)GAGAATGAGGAGTC;
 4040R, CGGCYGGGATGTCAAACCTTTG;
 4321F, GC(C/T)TA(C/T)GACCACAA(G/A)GACTC;
 4495R, CCTAAACTTTTCAAACANTGATGAG;
 5476F, CAATCATAGCAGT(C/A)ATTAG;

5711R, CGAGAGTTGCGCCTGCAG;

6415R, ATTTATATTATTCATACAATC.

Agroinfiltration of plants. *Agrobacteria* containing recombinant binary vectors were grown for 12 h on a shaker at 30°C. The cells (1.5 ml) were precipitated by centrifugation (4000g, 5 min), and the precipitate was resuspended in 1.5 ml buffer containing 10 mM Mes (pH 5.5) and 10 mM MgCl₂. Leaves of the plant *Nicotiana benthamiana* were injected with suspension of *agrobacteria* (optical density of the solution OD₆₀₀ = 0.2) using a syringe without a needle.

Isolation of proteins from plant tissue. Leaf matter (100 mg) was triturated in extracting buffer (0.4 M sucrose, 50 mM Tris (pH 8.0), 5 mM MgCl₂, 10% glycerol, 5 mM β-mercaptoethanol) to a homogenous suspension to perform dot-blot analysis or in GUS-buffer (50 mM sodium phosphate buffer, 10 mM β-mercaptoethanol, 10 mM EDTA, 0.1% Triton X-100) to obtain protein preparations for determination the activity of β-glucuronidase (GUS). The resulting mixture was centrifuged at 14,000g for 15 min, and the protein-containing supernatant was collected. The protein concentrations were determined by the Bradford method.

Determination of β-glucuronidase activity in agroinfiltrated plants. Different portions of the protein extract isolated from the agroinfiltrated zone were supplemented with 100 µl of GUS-buffer and 1 µl of 1 mM MUG (4-methylumbelliferyl β-D-glucuronide) solution. The reaction was performed for 1 h at 37°C and stopped by addition of four volumes of 0.2 M Na₂CO₃. The GUS activity was determined fluorometrically (excitation at 360 nm, emission at 450 nm).

Preparation of PVX Tula virus cDNA. The PVX Tula cDNA preparation was obtained on genomic RNA isolated from the virus particles by reverse transcription using the 6415R primer complementary to the PVX genome 3'-end. At the first stage the RNA preparation (2 µg) was supplemented with 20 pmol primer, the volume was adjusted to 11 µl, incubated for 5 min at 70°C, and the mixture was placed on ice. Then the reaction buffer, dNTP (to 1 mM), and 20 U of ribonuclease inhibitor (Promega, USA) were added, and the volume was adjusted to 19 µl. The resulting mixture was incubated for 5 min at 37°C, then 200 U of reverse transcriptase H Minus M-MuLV Reverse Transcriptase (Fermentas, Lithuania) was added, and cDNA was synthesized for 60 min at 42°C. The reaction was stopped by heating to 70°C for 10 min. The resulting preparation of cDNA was used as a template for PCR.

Phylogenetic analysis of nucleotide sequences. The nucleotide sequences were leveled using the Clustal X program [17]. Phylogenetic trees were designed as described in [18] using the Treecon program package [19]. The significance of phylogenetic trees was estimated by the bootstrap method [20].

RESULTS AND DISCUSSION

Determination of complete nucleotide sequence of potato virus X strain Tula genome. To determine the complete sequence, the PVX Tula genome was divided into six overlapping fragments, which were amplified by PCR on the PVX Tula cDNA. The primer sequences corresponded to conserved motifs of the PVX different strain genomes. The following primer pairs were used: 1F–1651R, 1348F–2566R, 2245R–4040R, 3130F–4495R, 4321F–5711R, and 5476F–6415R. The resulting PCR fragments were purified by electrophoresis in agarose gel and sequenced. The resulting sequences were

combined and, thus, the PVX Tula genome was completely sequenced (GenBank EU021215).

Analysis of the resulting nucleotide sequence of the genome indicates a similarity between the PVX strain Tula and PVX strain Russian [3]. The Tula strain genome also contains five genes: the RNA-dependent RNA polymerase gene, three genes of the so-called triple gene block responsible for the intercellular movement of the virus, and the coat protein gene. The genomes of these two strains have a length of 6435 nucleotides and include in total 310 inconsistent nucleotides (4.8%) and 37 amino acids (1.8%).

To study evolutionary relationships between different PVX strains, phylogenetic trees were designed by the

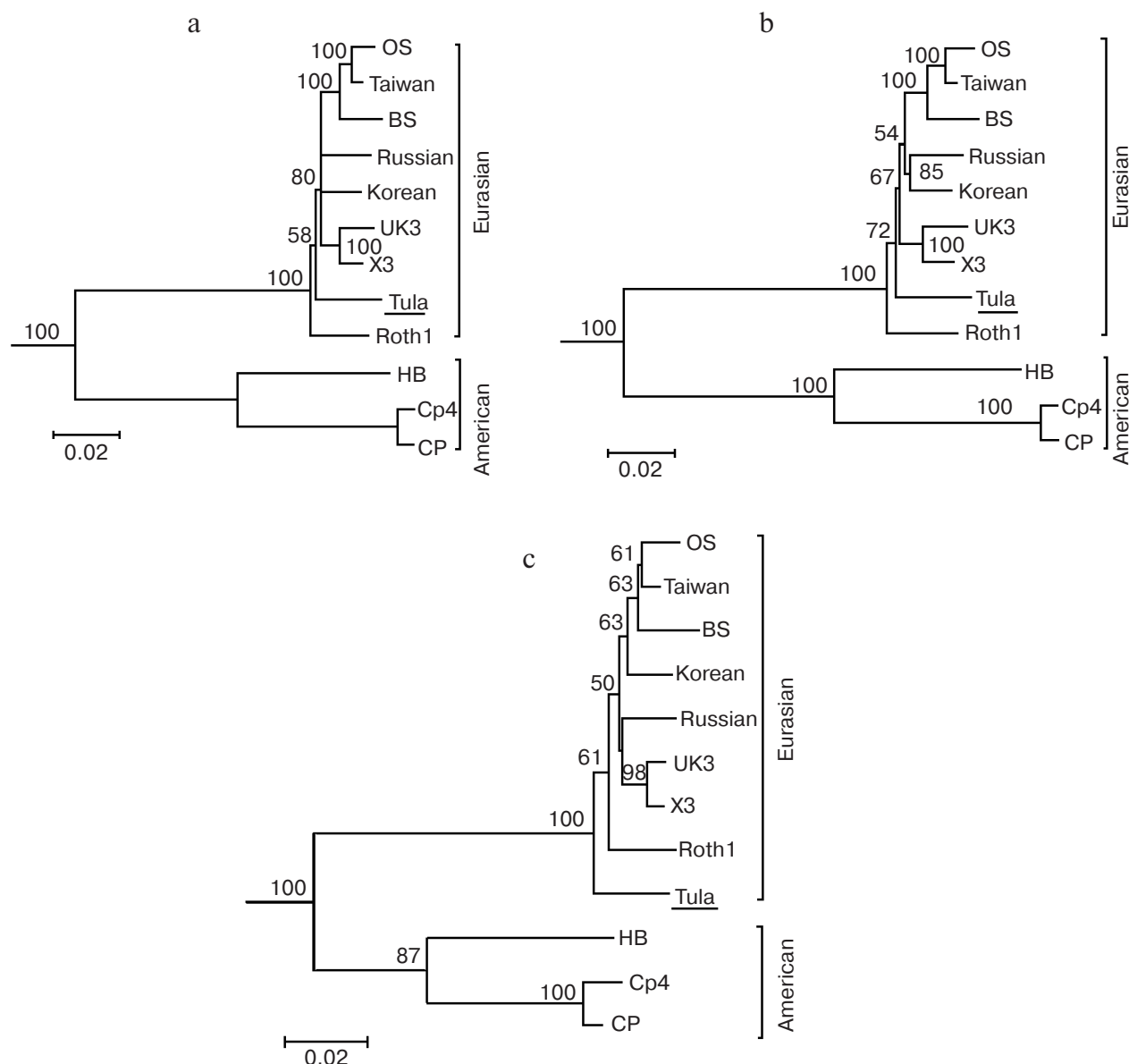


Fig. 1. Phylogenetic trees designed by the neighbor-joining method on the base of the complete nucleotide sequences of genomes (a), polymerase genes (b), or coat protein genes (c). Numbers indicate values of the bootstrap support of each branch.

neighbor-joining method based on complete nucleotide sequences of different PVX strain genomes and also on nucleotide and amino acid sequences of the polymerase and coat protein genes. On designing the trees, the papaya mosaic virus genome sequence (GenBank D13957.1) was used as an external group. Because of a high similarity of many strains, the trees based on the amino acid sequences were less informative than the trees based on the nucleotide sequences; therefore, they are not considered further. The trees based on nucleotide sequences of the triple gene block genes did not reveal significant differences between the majority of Eurasian strains (data not presented).

Analysis of the phylogenetic trees (Fig. 1) shows that the likeness between different PVX strains correlates with their geographical origin. Thus, two main groups consist of "American" (HB, CP4, and CP) and "Eurasian" strains of PVX, and the Eurasian strains OS, BS, and Taiwan isolated in the South-Eastern Asia form a separate cluster. The Tula strain belongs to the "Eurasian" group. Similar data were obtained in the work [13] as a result of analysis of eight different PVX strain genomes.

It is the most interesting that the phylogenetic trees based on both the complete nucleotide sequence and separately on the polymerase and coat protein genes display no significant differences. This suggests that evolution of the PVX genomes is a result of a gradual accumulation of replacements throughout the whole genome and not caused by recombinations between the strains producing "mosaic" genomes as often occurs in viruses of prokaryotes [21].

The absence of recombinant strains can be caused, in particular, by a low efficiency of recombination. The recombination between viral RNA molecules is supposed to be determined by RNA polymerase [22], and RNA polymerase mutants with an affected recombination but retained replication function were described for the brome mosaic virus [23]. And finally, recombinations can disturb the functionally important secondary structure of RNA.

Preparation of full infective cDNA copy of the PVX strain Tula genome. To clone full DNA copy of the PVX Tula genome, we amplified two overlapping genome fragments using two primer pairs, 1F–4040R and 2941F–6415R; a cDNA preparation from viral RNA was used as a template. Then the resulting fragments were cloned and combined into a complete copy of the PVX Tula genome in the plasmid vector pGEM-T (Promega). In the next stage the full copy of the PVX Tula genome was cloned in the plasmid vector under the 35S promoter control, and then the whole genetic construct involving the 35S promoter, the PVX Tula cDNA, and 35S-T terminator was cloned in T-DNA of the binary vector pBIN19 region. The structure of the resulting vector pBIN_PVX_Tula is shown in Fig. 2.

Accumulation of the PVX Tula and UK3 strains in infected plants. To compare infectivity of the PVX Tula and UK3 strains, the coat protein accumulation was deter-

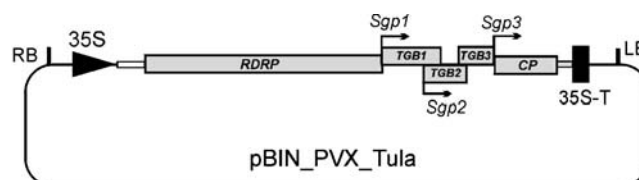


Fig. 2. Structure of the pBIN_PVX_Tula vector. The cDNA sequences of the PVX Tula genome copy are shown in gray color; rectangles indicate the RNA polymerase genes (RDRP), triple gene block (25K, 12K, and 8K), and the coat protein gene (CP). Promoters of subgenomic RNAs are shown by bent arrows. 35S and 35-T are promoter and terminator, respectively, of the cauliflower mosaic virus 35S RNA; RB and LB are the right and left limits of the T-DNA binary vector region.

mined on infecting *N. benthamiana* plants. The infecting conditions and dose of the infecting virus were standardized by agroinfiltrating the plants with PVX genome instead of mechanical inoculation with viral particles. Binary vectors pBIN_PVX201 [24] and pBIN_PVX_Tula were used containing between the 35S promoter and 35-T terminator full copies of the PVX UK3 and Tula genomes, respectively. Recombinant binary vectors were injected into the bacterium *A. tumefaciens*, which was used for infecting *N. benthamiana* plants (agroinfiltration).

On infecting plant cells, the T-DNA region of the binary vector, which involves the PVX genome cDNA, is transferred into the cell nucleus. The transcription initiated from the 35S promoter active in plants results in synthesis of viral genomic RNA, which is later replicated and translated, as occurs during infection. This leads to the virus reproduction and infection of the whole plant. An important advantage of the agroinfiltration method is a high efficiency of transfection providing for infection of virtually all cells of the leaf. The infecting agent dose is determined by quantity of agrobacteria, which allows two different PVX strains to be infected under maximally similar conditions.

The plants were infected with the agrobacteria containing plasmids pBIN_PVX_Tula or pBIN_PVX201, and for each strain several plants were taken. The infection signs were evident two weeks after the infection on both agroinfiltrated and other leaves—thus the infection in both cases was systemic.

Infectivities of the PVX strain Tula and a "reference" UK3 strain were compared by semiquantitative determination of the virus coat protein accumulation in the infected plants. From the leaves (directly infiltrated with agrobacteria and the others) of the infected plant protein preparations were isolated and analyzed by dot-blotting with antibodies to the PVX coat protein. For each strain, 23 samples were analyzed including eight preparations from the agroinfiltrated leaves. The band intensities on the digital photo were measured using the Tina program, and the signal of all samples was summarized for each of

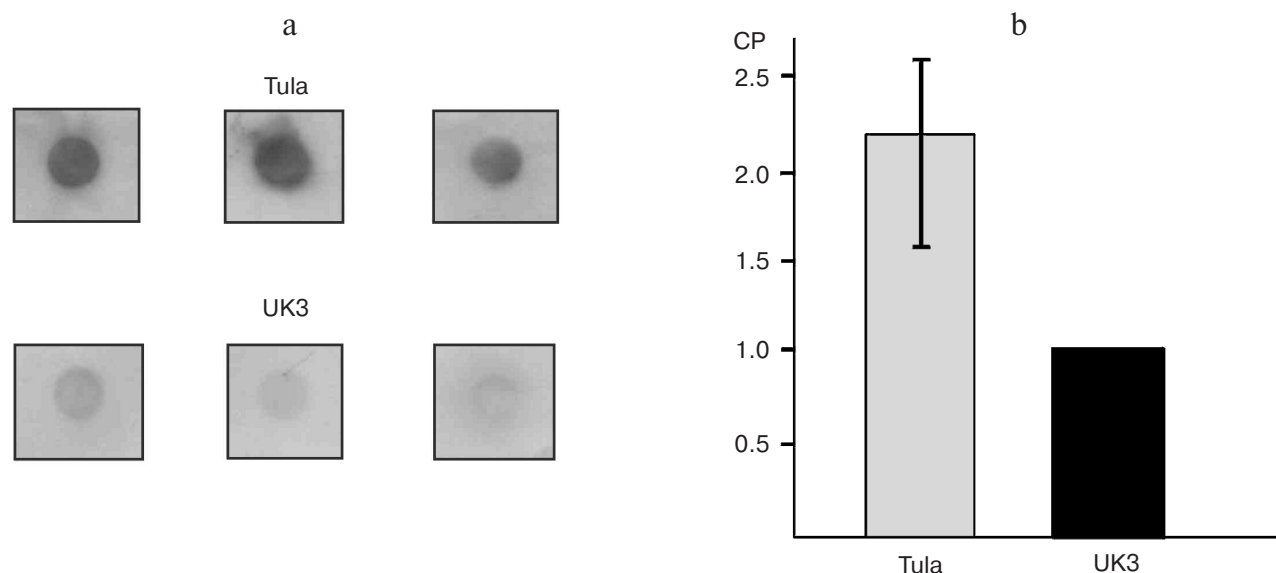


Fig. 3. Coat protein accumulation in plants infected with PVX strains Tula and UK3. a) Dot-blot analysis of protein extracts from infected plants (examples are given for three specimens of each strain). b) Quantitative evaluation of relative levels of CP synthesis based on comparison of intensities of signals obtained by dot-blot analysis of all 46 protein extracts from the infected plants. The value for the PVX strain UK3 is taken as unity.

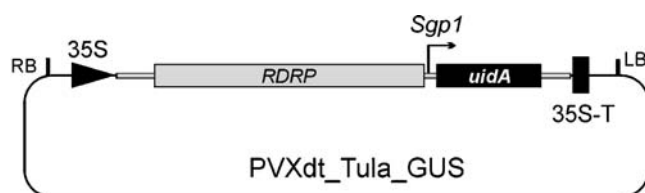


Fig. 4. Structure of the expressing vector PVXdt_Tula_GUS. The cDNA copy sequences of the PVX Tula genome are shown in gray color; the bent arrow indicates the subgenomic RNA 1; uidA presents the β-glucuronidase gene; 35S and 35S-T are promoter and terminator, respectively, of the cauliflower mosaic virus 35S RNA. RB and LB are the right and left limits of the T-DNA binary vector region.

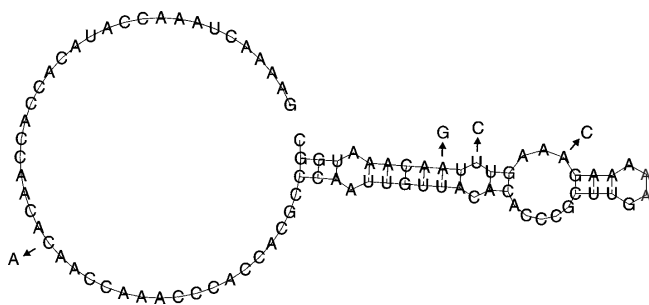


Fig. 5. Secondary structure of RNA of PVX strain UK3 5' non-translated region predicted by the RNA Fold program (<http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi>). Arrows indicate nucleotide substitutions in the Tula strain genome. The free energy of the secondary structures calculated by the RNA Fold program is -11.24 kcal/mol for the PVX strain UK3 RNA and -14.01 kcal/mol for the Tula strain RNA.

the two strains. The coat protein accumulation was about twofold higher in the plants infected with the PVX Tula strain than in those infected with the UK3 strain (Fig. 3).

Construction of expressing viral vector based on PVX Tula genome. A new vector was created on the basis of the viral vector PVXdt described in [24], which was founded on the PVX strain UK3 genome. This vector involves a 5' nontranslated region of the PVX genome, the polymerase gene, the subgenomic RNA first promoter, the last 60 nucleotides of the coat protein gene, and the 3' nontranslated region of the PVX genome. This construct is placed between the 35S promoter and 35S terminator and cloned in the binary vector pBIN19. Agroinfiltration of leaves results in infection of the majority of the leaf cells in the infected region, replication of the viral vector in individual cells, and a high level of synthesis of the product [24].

We used the recombinant vector PVXdt containing the β-glucuronidase gene of *E. coli*, uidA (PVXdt_GUS). To create an equivalent vector based on the PVX Tula genome, the fragment PVXdt_GUS located between the 35S promoter and uidA gene was substituted by an equivalent region of the PVX Tula, and as a result, the PVXdt_Tula_GUS vector was obtained (Fig. 4). Note that sequences of the 3' nontranslated regions of the PVX Tula and UK3 are not different.

Efficiencies of the newly constructed vector PVXdt_Tula_GUS and the vector PVXdt_GUS based on the standard strain UK3 genome were compared by determining the transient expression of the uidA gene. The agrobacteria containing the expressing vectors were agroinfiltrated into the *N. benthamiana* cells. The agroin-

Determination of GUS activity in plants infected with viral vectors PVXdt_Tula_GUS and PVXdt_GUS

Leaf number	GUS activity, arbitrary units		GUS relative activity (Tula/UK3)
	PVXdt_Tula_GUS	PVXdt_GUS	
1	35	25	1.4
2	25	14	1.8
3	27	19	1.4
4	110	33	3.3
5	16	29	0.6
6	49	21	2.3
Average value	—	—	1.8

filtration zones corresponding to both vectors were simultaneously present on the same leaf. Three days later samples of the plant tissues were taken from the agroinfiltrated zones, and the GUS activity was quantified. The GUS synthesis on using the vector based on the strain Tula genome was about 1.5–2-fold higher than on using the vector based on the UK3 genome (table).

Note that the higher productivity of the vector based on the PVX strain Tula genome correlated with the higher level of the protein coat accumulation in the plants infected with this strain. Because the PVXdt_Tula_GUS and PVXdt_GUS vectors are discriminated only by the 5' nontranslated regions and the RNA-dependent RNA polymerase genes, it was suggested that just PVX Tula polymerase should be responsible for the more effective replication of viral RNA in *N. benthamiana* and the higher accumulation of the virus during the infection.

However, the observed effect could also be caused by differences in the 5' nontranslated regions of the genomes of these strains, which formed stable hairpin structures essential for the reproduction of the virus [25]. The 5' nontranslated regions of the Tula and UK3 strain genomes are different in four nucleotides (Fig. 5), and one of them (C→A in position +27 from the 5'-end of RNA) is located beyond the hairpin structure limits, two other substitutions (U→C in position +78 and A→G in position +80) do not change the hairpin structure, and only one substitution (A→C in position +73) stabilizes the hairpin in mRNA at the cost of lengthening the double-stranded part.

The authors are grateful to Yu. L. Dorokhov for the presented pBIN_PVX201 and PVXdt plasmids and antibodies to the PVX coat protein, B. B. Kuznetsov and E. S. Bulygina who provided oligonucleotides for amplification of the PVX genome fragments, and T. V. Kolganova for her help in the sequencing.

This work was supported by the Federal Agency for Science and Innovation of the Russian Federation (contract No. 02.435.11.3012) and the Russian Academy of Sciences Presidium program on Molecular and Cellular Biology.

REFERENCES

- Goelet, P., Lomonosoff, G. P., Butler, P. J., Akam, M. E., Gait, M. J., and Karn, J. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 5818–5822.
- Skryabin, K. G., Morozov, S. Yu., Kraev, A. S., Rozanov, M. M., Chernov, B. K., and Atabekov, J. G. (1988) *FEBS Lett.*, **240**, 33–40.
- Skryabin, K. G., Kraev, A. S., Morozov, S. Yu., Rozanov, M. N., Chernov, B. K., Lukasheva, L. I., and Atabekov, J. G. (1988) *Nucleic Acids Res.*, **16**, 10929–10930.
- Marillonnet, S., Thoeringer, C., Kandzia, R., Klimyuk, V., and Gleba, Y. (2005) *Nat. Biotechnol.*, **23**, 718–723.
- Yusibov, V., Hooper, D., Spitsin, S., Fleysh, N., Kean, R., Mikheeva, T., Deka, D., Karasev, A., Cox, S., Randall, J., and Koprowski, H. (2002) *Vaccine*, **20**, 3155–3164.
- Santi, L., Giritch, A., Roy, C. J., Marillonnet, S., Klimyuk, V., Gleba, Y., Webb, R., Arntzen, C. J., and Mason, H. S. (2006) *Proc. Natl. Acad. Sci. USA*, **103**, 861–866.
- Wagner, B., Hufnagl, K., Radauer, C., Wagner, S., Baier, K., Scheiner, O., Wiedermann, U., and Breiteneder, H. (2004) *J. Immunol. Meth.*, **287**, 203–215.
- Giritch, A., Marillonnet, S., Engler, C., van Eldik, G., Botterman, J., Klimyuk, V., and Gleba, Y. (2006) *Proc. Natl. Acad. Sci. USA*, **103**, 14701–14706.
- Canizares, M. C., Nicholson, L., and Lomonosoff, G. P. (2005) *Immunol. Cell Biol.*, **83**, 263–270.
- Huisman, M. J., Linthorst, H. J. M., Bol, J. F., and Cornelissen, B. J. C. (1988) *J. Gen. Virol.*, **69**, 1789–1798.
- Kagiwada, S., Yamaji, Y., Nakabayashi, H., Ugaki, M., and Namba, S. (2002) *J. Gen. Plant Pathol.*, **68**, 94–98.
- Kavanagh, T., Goulden, M., Santa Cruz, S., Chapman, S., Barker, I., and Baulcombe, D. C. (1992) *Virology*, **189**, 609–617.
- Malcuit, I., de Jong, W., Baulcombe, D. C., Shields, D. C., and Kavanagh, T. A. (2002) *Virus Genes*, **20**, 165–172.
- Komatsu, K., Kagiwada, S., Takahashi, S., Yamaji, Y., Hirata, H., Ugaki, M., and Namba, S. (2005) *Virus Genes*, **31**, 99–105.
- Querci, M., Vlucht, R., Goldbach, R., and Salazar, L. F. (1993) *J. Gen. Virol.*, **74**, 2251–2255.
- Orman, B. E., Celnik, R. M., Mandel, A. M., Torres, H. N., and Mentaberry, A. N. (1990) *Virus Res.*, **16**, 293–306.
- Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (1997) *Nucleic Acids Res.*, **24**, 4876–4882.
- Saitou, N., and Nei, M. (1987) *Mol. Biol. Evol.*, **4**, 406–425.
- Van de Peer, Y., and de Wachter, R. (1994) *Comput. Appl. Biosci.*, **10**, 569–570.
- Felsenstein, J. (1985) *Evolution*, **39**, 783–791.
- Hendrix, R. (2003) *Curr. Opin. Microbiol.*, **6**, 506–511.
- Lai, M. M. C. (1992) *Microbiol. Rev.*, **56**, 61–79.
- Figlerowitz, M., Nagy, P. D., and Bujarski, J. J. (2003) *Proc. Natl. Acad. Sci. USA*, **94**, 2073–2078.
- Komarova, T. V., Skulachev, M. V., Zvereva, A. S., Shwarts, A. M., and Atabekov, J. G. (2006) *Biochemistry (Moscow)*, **71**, 846–850.
- Kwon, S. J., and Kim, K. H. (2006) *Mol. Cells*, **21**, 63–75.