

Cloning Of A Transmissible Gastroenteritis Coronavirus Full-Length cDNA

JOSE M. GONZALEZ, FERNANDO ALMAZAN, ZOLTAN PENZES,
ENRIQUE CALVO, AND LUIS ENJUANES

*Department of Molecular and Cell Biology, Centro Nacional de Biotecnología, CSIC, Campus
Universidad Autónoma. Cantoblanco, 28049 Madrid, Spain*

1. INTRODUCTION

To understand gene function and expression in coronaviruses it would be of interest to obtain a cDNA encoding a full-length infectious RNA. This has not yet been possible due to the large size of the coronavirus genome and the instability of plasmids carrying coronavirus replicase sequences. In this report we describe the construction of a transmissible gastroenteritis virus (TGEV) full-length cDNA. For this purpose, we started from a cDNA encoding the defective RNA DI-C, that was stably and efficiently replicated by the helper virus (Izeta *et al.*, 1999). During the completion of the cDNA an ORF 1a fragment that was toxic to the bacteria was identified. Advantage of this finding was taken by reintroducing the toxic fragment into the viral cDNA in the last cloning step. To enhance the stability of the viral sequence, the cDNA was cloned as a bacterial artificial chromosome (BAC), a system that has been useful to stably clone large DNA from a variety of complex genomic sources into bacteria. The cytomegalovirus (CMV) immediate-early promoter was placed upstream of the cDNA to make use of a two-step amplification system that couples RNA pol II-driven transcription in the nucleus with the amplification by the viral replicase in the cytoplasm.

2. MATERIALS AND METHODS

2.1 Cells, Viruses, Plasmids and Bacteria Strains

The TGEV strains PUR46-MAD and PUR46-C11 were grown and titered as described (Sánchez *et al.*, 1999). Plasmid pACNR1180 was kindly provided by J.D. Tratschin (Institute of Virology and Immunoprophylaxis, Mittelhäusern, Switzerland). Plasmid pBeloBAC11 was kindly provided by H. Shizuya and M. Simon (California Institute of Technology, Pasadena, CA). Plasmid pRL-CMV was obtained from Promega. *E. coli* DH10B strain was obtained from GIBCO/BRL.

2.2 Plasmid DNA Preparation, Sequence Analysis and Plasmid Stability

The preparation of plasmid DNA, the sequence analysis and the study of pBeloBAC11-based plasmids stability were performed as described (Almazán *et al.*, 2000).

3. RESULTS AND DISCUSSION

3.1 Cloning and Assembly of a TGEV Genome cDNA

As a backbone for the assembly of a TGEV full-length cDNA, the plasmid pDI-C, encoding a TGEV-derived defective RNA was used. The DI-C cDNA was flanked by the CMV immediate early promoter at the 5' end and a synthetic poly-A tail of 24 bp followed by the hepatitis delta virus ribozyme and the bovine growth hormone termination and polyadenylation sequences at the 3' end. DI-C RNA has three deletions: $\Delta 1$ of 10 kb, $\Delta 2$ of 1.1 kb and $\Delta 3$ of 7.7 kb. A set of cDNAs encoding the missing sequences was generated from the respiratory PUR46-MAD strain RNA by standard RT-PCR techniques and assembled as described (Fig. 1). To make a fully functional virus, regions corresponding to the S gene were obtained from the virulent PUR46-C11 strain, which replicated to high titers in both the enteric and the respiratory tracts.

D1-5' fragment (positions 9349-11173) could not be cloned stably in *E. coli* cells, as DNA underwent sequence rearrangements (González *et al.*, 2000). Interestingly, D1-5' sequence was split into two overlapping clones spanning positions 8978-9758 and 9581-11173, respectively, that

were fully stable. These results indicated that none of the individual fragments, but their joint sequences, were the source of toxicity. In the last step, the Clal^{4417} - Clal^{9615} fragment was inserted into the plasmid encoding the rest of the TGEV genome ($\text{pAC-TGEV}^{\Delta\text{Clal}}$), previously transferred to the low-copy number vector pACNR1180 (10-12 copies/cell) to minimize instability. However, *E. coli* cells did not tolerate the full-length cDNA.

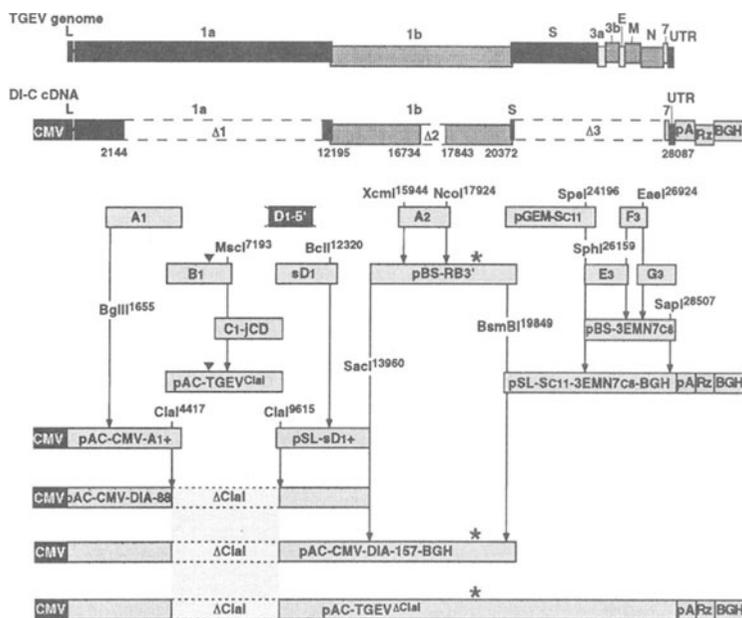


Figure 1. Strategy for the assembly of a TGEV full-length cDNA starting from the DI-C cDNA. CMV, cytomegalovirus immediate-early promoter; L, leader sequence; UTR, untranslated region; pA, tail of 24 A residues; Rz, hepatitis delta virus ribozyme; BGH, bovine growth hormone termination and polyadenylation sequences; ▼, marker at position 6752; *, marker at position 18997.

To improve definitively the stability of TGEV sequences in *E. coli* cells, the last step of the assembly was performed in the very-low-copy number vector pBeloBAC11 (1-2 copies/cell), allowing the construction of a full-length TGEV cDNA in *E. coli* DH10B cells.

3.2 Stability of the TGEV Full-Length cDNA as a BAC

The stability of the BAC-based plasmids was checked by restriction endonuclease analysis (González *et al.*, 2000). The plasmid encoding the full-length TGEV cDNA was stable in DH10B cells for at least 80 generations. Afterwards, it began to undergo rearrangements, mostly insertions from 0.7 to 1.6 kb in length throughout the ORF 1a (data not shown). The sequence of a ~0.7 kb insertion located at position 9480 was

identical to that of an *E. coli* insertion sequence IS1, a transposon of 768 bp that is randomly inserted, generating 9-bp direct repeats in the target DNA.

To disrupt the toxic sequence and increase the stability of the TGEV full-length cDNA, a 133-nt intron from the plasmid pRL-CMV was inserted between the positions 9595 and 9596 to generate 5' and 3' intron splice sites matching the consensus sequence of mammalian introns. The intron-containing TGEV full-length cDNA was completely stable after 120 generations. *In vivo* transcription of the intron-containing full-length clones from the CMV promoter should assure splicing of the intron in the nucleus of the transfected cells..

4. CONCLUSION

During the assembly of a TGEV full-length cDNA a ORF 1a region was determined as the cause of instability. The TGEV full-length cDNA was cloned as a bacterial artificial chromosome (BAC) in *E. coli* DH10B cells, where it was stable for, at least, 80 generations. The introduction of an intron in the ORF 1a completely stabilized the full-length cDNA.

ACKNOWLEDGMENTS

This research was supported by grants from the Comisión Interministerial de Ciencia y Tecnología (Spain), the Community of Madrid, the European Community (Biotechnology, FAIR and Control of Infectious Diseases Programs), and Fort Dodge Veterinaria (Spain).

REFERENCES

- Almazán, F., González, J. M., Péñzes, Z., Izeta, A., Calvo, E., Plana-Durán, J. and Enjuanes, L. (2000). Engineering the largest RNA virus genome as an infectious bacterial artificial chromosome. *Proc. Natl. Acad. Sci. USA* **97**, 5516-5521.
- González, J. M., Almazán, F., Péñzes, Z., Calvo, E. and Enjuanes, L. (2000). Construction of a stable transmissible gastroenteritis coronavirus full-length cDNA. Submitted.
- Izeta, A., Smerdou, C., Alonso, S., Péñzes, Z., Méndez, A., Plana-Durán, J. y Enjuanes, L. (1999). Replication and packaging of transmissible gastroenteritis coronavirus-derived synthetic minigenomes. *J. Virol.* **73**, 1535-1545.
- Sánchez, C. M., Izeta, A., Sánchez-Morgado, J. M., Alonso, S., Sola, I., Balasch, M., Plana-Durán, J. y Enjuanes, L. (1999). Targeted recombination demonstrates that the spike gene of transmissible gastroenteritis coronavirus is a determinant of its enteric tropism and virulence. *J. Virol.* **73**, 7607-7618.