A common RNA motif in the 3′ end of the genomes of astroviruses, avian infectious bronchitis virus and an equine rhinovirus

Christine Monceyron Jonassen, Tom Ø. Jonassen and Bjørn Grinde

Department of Virology, National Institute of Public Health, PO Box 4404 Torshov, N-0403 Oslo, Norway

In the 3′ non-coding region of the genomes of infectious bronchitis virus, an avian coronavirus and the picornavirus equine rhinovirus serotype 2, there is a motif with remarkable similarity, both in sequence and folding, to the second RNA stem–loop from the 3′ end of the genomes of human astroviruses. This motif was also found in astroviruses of sheep, pig and turkey, suggesting that it is a common feature of all astroviruses. The conserved nature of the motif indicates that there has been strong selection for its preservation. There is significant homology between the regions flanking this motif in infectious bronchitis virus and a continuous RNA sequence at the same distance from the 3′ poly(A) tail in some related mammalian coronaviruses. These observations suggest that the presence of the motif in these three viral families is the result of at least two separate RNA recombination events.

Astroviruses are non-enveloped particles with a plus-strand RNA genome with a 3′ poly(A) tail. They can infect the gastrointestinal tract of humans and various animals (Madeley & Cosgrove, 1975; Saif et al., 1990; Shimizu et al., 1990; Snodgrass & Gray, 1977). About 110 nucleotides (nt) adjacent to the poly(A) tail of the genomes of human astroviruses (HAstV) are highly conserved both in nucleotide sequence and in the folding into two RNA stem–loop motifs (Monceyron et al., 1997; Monroe et al., 1993). Nucleotide database searches for sequences resembling this part of the HAstV genome have yielded an RNA sequence in the conserved part of the 3′ non-coding region (NCR) of infectious bronchitis virus (IBV) with 90% similarity to the second stem–loop from the 3′ end of the HAstV genome. This HAstV stem–loop II consists of a basal stem of 6 base pairs and a 31 nt loop region with more uncertain interaction between the nucleotides. The biological importance of the basal stem is substantiated by nucleotide covariations among HAstV serotypes, i.e. sequence differences are compensated by other differences to maintain base pairing (Monceyron et al., 1997).

IBV is a coronavirus with a 27·6 kb plus-strand RNA genome containing a 0·3–0·5 kb 3′ NCR and a poly(A) tail (Williams et al., 1993). Different strains have different tissue tropisms, infecting respiratory, reproductive and gastrointestinal organs as well as the kidneys of chicken (Siddell et al., 1983). In IBV this stem–loop II-like motif (s2m) can be folded exactly like the HAstV stem–loop II, and we have suggested its presence in these very different viruses to be the result of a natural RNA recombination between an astrovirus and IBV (Monceyron et al., 1997).

Recombination among distantly related groups of viruses is probably important in the evolution of RNA viruses. In contrast to recombination within certain viral species, however, transfer of RNA from one virus family to another seems very rare. Because of the high rates of nucleotide substitution in RNA viruses, traces of such recombinations are not easily detected. Examples of recombination among distant groups of viruses are reviewed in Koonin & Dolja (1993) and Lai (1992).

In this paper we describe the presence of s2m in previously unpublished RNA sequences of three animal astroviruses and in the 3′ NCR of the recently published equine rhinovirus serotype 2 (ERV-2) sequence. This unclassified picornavirus causes a mild respiratory infection in horses, and has an 8·8 kb plus-strand RNA genome with a 167 nt 3′ NCR and a poly(A) tail. Both the total genome and its 3′ NCR are somewhat longer than in other vertebrate picornaviruses (Wutz et al., 1996).

To investigate the origin and distribution of s2m, we have sequenced the 3′ region of the genomes of astroviruses from sheep (Snodgrass & Gray, 1977), pig (Shimizu et al., 1990) and turkey (Saif et al., 1990), extracted from faecal samples containing particles morphologically similar to HAstV. We used a plus-strand primer from the HAstV s2m and (T)₉₋₁₀GC as the minus-strand primer for PCR amplification of cDNA and for subsequent sequencing. These sequences were expanded...
Fig. 1. Alignment of the 3′ end of the genomes of HAstV (serotype 1 shown), pig astrovirus (PAstV), sheep astrovirus (SAstV), turkey astrovirus (TAstV), an IBV consensus and ERV-2. Identity in at least five of the sequences is shown in boxes. Gaps are shown as dashes. The beginning and end of s2m are marked by asterisks above the sequences.

Fig. 2. The s2m of ERV-2. Proposed canonical interactions between nucleotides are indicated by horizontal lines. Differences in other viruses containing s2m are described to the right of each position. HAstV serotype number is indicated where appropriate.

Computer predictions of possible foldings of the 3′ 500 nt including 20 nt of the poly (A) tail were performed using the MFOLD program in the Wisconsin sequence analysis package version 8.0 (Genetics Computer Group, Wisconsin, USA). The folding of s2m presented in Fig. 2 is suggested for IBV, ERV-2 and the astroviruses. In the loop region of s2m, there are alternative foldings with similar free energies. However, covariations suggest that the two base pairs 11:34 and 18:24 were conserved in the animal astroviruses, but in the turkey virus an upstream stop codon had made it obsolete. In IBV and ERV-2 the entire s2m is within the 3′ NCR.

Proposed canonical interactions between nucleotides are indicated by horizontal lines. Differences in other viruses containing s2m are described to the right of each position. HAstV serotype number is indicated where appropriate.

by amplification of cDNA using the 5′ RACE kit (Gibco BRL) with minus-strand primers based on the initially obtained sequences, and subsequent primer walking in the 5′ direction.

The results showed that s2m was also present in the 3′ end of the genomes of these three non-human astroviruses, and that, except between the 3′ NCRs of astroviruses from pig and humans, there was little sequence similarity in the regions flanking s2m (Fig. 1). Its presence in all sequenced astroviruses makes it likely that s2m is a universal feature of astroviruses. The HAstV capsid protein precursor gene stop codon is located within s2m (positions 22–24 in Fig. 2). The stop codon is G in turkey astrovirus, A in the HAstV-1 and pig astrovirus; U in other astroviruses and most IBV.

Fig. 2. The s2m of ERV-2. Proposed canonical interactions between nucleotides are indicated by horizontal lines. Differences in other viruses containing s2m are described to the right of each position. HAstV serotype number is indicated where appropriate.
implies that a replication-dependent recombination must have localization of s2m in the 3'-20 nt of viral RNA genomes may serve various functions by interacting with host-cell structures, viral proteins, other parts of the viral genome or as a structural feature of the RNA per se. The presence of s2m in these otherwise very different genomes makes it unlikely that the selective advantage of s2m depends on its interaction with viral structures.

While recombination within certain types of viruses appears to be common in various viral families (Lai, 1992), s2m might be one of the strongest traces of natural recombination between non-related viruses.

We thank Dr D. R. Snodgrass (Moredun Research Institute, Edinburgh, UK), Dr Y. M. Saif (The Ohio State University, Ohio, USA) and Dr H. Ushijima (Institute of Public Health, Tokyo, Japan) for providing the animal virus samples.

References


Received 5 September 1997; Accepted 10 December 1997