A Highly Divergent Picornavirus in a Marine Mammal

A. Kapoor,1,2 J. Victoria,1,2 P. Simmonds,3 C. Wang,4 R. W. Shafer,4 R. Nims,5 O. Nielsen,9 and E. Delwart1,2

Blood Systems Research Institute, San Francisco, California 94118; University of California, San Francisco, San Francisco, California2; Centre for Infectious Diseases, University of Edinburgh, Summerhall, Edinburgh EH9 1QH, United Kingdom1; Division of Infection Diseases, Stanford University, Stanford, California4; BioReliance Corporation, 14920 Broschart Road, Rockville, Maryland 20850; and Department of Fisheries and Oceans, Central and Arctic Region, 501 University Crescent, Winnipeg, Manitoba R3T 2N6, Canada6

Received 6 June 2007/Accepted 6 October 2007

Nucleic acids from an unidentified virus from ringed seals (Phoca hispida) were amplified using sequence-independent PCR, subcloned, and then sequenced. The full genome of a novel RNA virus was derived, identifying the first sequence-confirmed picornavirus in a marine mammal. The phylogenetic position of the tentatively named seal picornavirus 1 (SePV-1) as an outlier to the grouping of parechooviruses was found consistently in alignable regions of the genome. A mean protein sequence identity of only 19.3 to 30.0% was found between the 3D polymerase gene sequence of SePV-1 and those of other picornaviruses. The predicted secondary structure of the short 506-base 5'-untranslated region showed some attributes of a type IVB internal ribosome entry site, and the polyprotein lacked an apparent L peptide, both properties associated with the Parechovirus genus. The presence of two SePV-1 2A genes and of the canonical sequence required for cotranslational cleavage resembled the genetic organization of Ljungan virus. Minor genetic variants were detected in culture supernatants derived from 8 of 108 (7.4%) seals collected in 2000 to 2002, indicating a high prevalence of SePV-1 in this hunted seal population. The high level of genetic divergence of SePV-1 compared to other picornaviruses and its mix of characteristics relative to its closest relatives support the provisional classification of SePV-1 as the prototype for a new genus in the family Picornaviridae.

Materials and Methods

Extraction of viral nucleic acids. A nasal swab from a seal hunted in 2002 was used to inoculate Vero cells, which developed CPE. CPE was transferable to a fresh cell culture. The supernatant from this cell culture was used as the input to nonspecifically amplify viral nucleic acids. For enrichment of viral particles, 2 ml of culture supernatant was clarified (5,000 × g for 10 min) and filtered through a 0.45-μm pore-size sterile filter. The filtrate was then centrifuged at 35,000 × g for 3 h at 10°C, and the resulting pellet was resuspended in 100 μl of water containing 1× Turbo DNase buffer. To remove non-viral-particle-protected DNA from the cultured cells, 20 U of Turbo DNase (Ambion) was added, followed by incubation at 37°C for 90 min. Particle-protected nucleic acids were then extracted using a viral RNA extraction kit (Qiagen).

Random amplification, subcloning, and sequencing. Viral RNA was mixed with 50 pmol of primer RA01 (GCGCAGGCTCAGATCATCAGN) and reverse transcription reaction mix was incubated at 25°C for 10 min and then at 42°C for 50 min. After a denaturation step at 94°C for 3 min and chilling on ice (to reanneal primer RA01 to cDNA), 2.5 units of Turbo DNase, 1.5 ml of filtrate was then centrifuged at 35,000 × g for 3 h at 10°C, and the resulting pellet was resuspended in 100 μl of water containing 1× Turbo DNase buffer. To remove non-viral-particle-protected DNA from the cultured cells, 20 U of Turbo DNase (Ambion) was added, followed by incubation at 37°C for 90 min. Particle-protected nucleic acids were then extracted using a viral RNA extraction kit (Qiagen).

Random amplification, subcloning, and sequencing. Viral RNA was mixed with 50 pmol of primer RA01 (GCGCAGGCTCAGATCATCAGN) and reverse transcription reaction mix was incubated at 25°C for 10 min and then at 42°C for 50 min. After a denaturation step at 94°C for 3 min and chilling on ice (to reanneal primer RA01 to cDNA), 2.5 units (0.5 μl) of 3'-5' Exo-Klenow DNA polymerase (New England Biolabs) was added to extend RA01 and incubated at 37°C for

* Corresponding author. Mailing address: Blood Systems Research Institute, 270 Masonic Ave., San Francisco, CA 94118. Phone: (415) 923-5763. Fax: (415) 276-2311. E-mail: delwarte@medicine.ucsf.edu.

† Published ahead of print on 17 October 2007.
50 picomoles of a random octamer oligonucleotide was added to 10 mM Tris · HCl [pH 8.3], 300 mM KCl (Applied Biosystems), 3 mM MgCl₂, each dNTP at 0.3 mM, 50 pmol of primer RA02 (GCCGGAGGCTTCGAGATATC), and 2.5 units of AmpliTaq Gold DNA polymerase LD (Applied Biosystems). An initial denaturation step for 5 min at 95°C was followed by 40 cycles of PCR (94°C for 1 min, 55°C for 1 min, and 72°C for 2 min). Random PCR products were then separated on a 1.5% agarose gel, and DNA smears ranging in size from approximately 400 to 1,500 bp were excised and extracted using the QIAquick gel extraction kit (Qiagen). Five microliters of the eluted, purified PCR product was ligated into the pGEMT-Easy vector (Promega Inc.) and introduced into chemically competent Escherichia coli TOP-10 cells (Topo One Shot; Invitrogen). Bacteria were plated onto LB agar plates containing ampicillin, X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) and IPTG (isopropyl-β-D-thiogalactopyranoside). Ninety-six white colony inserts were sequenced using the T-7 forward primer.

**Sequence analysis.** Sequence data for all clones were imported into Sequence 4.1 (Genecode) and trimmed of vector and primer (RA02) sequences. The remaining sequences were then assembled into contigs using an assembly parameter of a minimum 90% base identity with at least a 30-nucleotide overlap. A sequence similarity search was performed using tblastx (http://www.ncbi.nlm.nih.gov/BLAST).

**SePV-1 genome sequencing.** Sequence contigs of sequences were assembled using Sequencer software. Contigs showing significant tblastx hits to picornaviruses (E value of <0.001) were then linked using PCR overlap. To acquire the 3' end of viral genome, 10 μl of extracted RNA was mixed with 10 pmol of primer DT-01 (ATTCTAGAGGCCGAGGCGCCGACATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT!n
Table 1. Pairwise amino acid sequence identities between SePV-1 and other picornavirus genera in the P1 and 3D regionsa

<table>
<thead>
<tr>
<th>Virus</th>
<th>SePV-1</th>
<th>Parechovirus</th>
<th>DHV</th>
<th>Aphthovirus</th>
<th>Cardiovirus</th>
<th>Erbovirus</th>
<th>Teschovirus</th>
<th>SVV</th>
<th>Enterovirus</th>
<th>Sapelovirus</th>
<th>Hepatovirus</th>
<th>Kobuvirus</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1 region</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parechovirus</td>
<td>25.6</td>
<td>61.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DHV</td>
<td>24.3</td>
<td>28.9</td>
<td>78.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aphthovirus</td>
<td>11.6</td>
<td>12.1</td>
<td>12.4</td>
<td>61.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cardiovirus</td>
<td>9.3</td>
<td>12.2</td>
<td>12.5</td>
<td>25.9</td>
<td>70.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Erbovirus</td>
<td>10.3</td>
<td>12.3</td>
<td>12.6</td>
<td>30.2</td>
<td>31.5</td>
<td>73.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teschovirus</td>
<td>9.4</td>
<td>11.5</td>
<td>11.8</td>
<td>20.0</td>
<td>23.4</td>
<td>21.6</td>
<td>79.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SVV</td>
<td>10.6</td>
<td>11.8</td>
<td>10.7</td>
<td>25.2</td>
<td>31.9</td>
<td>29.4</td>
<td>20.2</td>
<td>N4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enterovirus</td>
<td>11.0</td>
<td>10.0</td>
<td>12.3</td>
<td>18.9</td>
<td>20.1</td>
<td>21.0</td>
<td>17.6</td>
<td>19.8</td>
<td>54.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sapelovirus</td>
<td>10.5</td>
<td>11.5</td>
<td>12.4</td>
<td>19.8</td>
<td>21.2</td>
<td>22.3</td>
<td>18.7</td>
<td>21.1</td>
<td>34.8</td>
<td>49.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kobuvirus</td>
<td>12.4</td>
<td>12.2</td>
<td>11.0</td>
<td>13.9</td>
<td>14.2</td>
<td>14.1</td>
<td>13.0</td>
<td>13.4</td>
<td>15.5</td>
<td>69.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3D region</td>
<td>12.0</td>
<td>12.7</td>
<td>12.5</td>
<td>15.5</td>
<td>18.0</td>
<td>18.4</td>
<td>14.8</td>
<td>17.6</td>
<td>16.9</td>
<td>16.1</td>
<td>12.8</td>
<td>63.7</td>
</tr>
</tbody>
</table>

Note: Boldface type, averages of pairwise amino acid sequence identities between SePV-1 and members of different picornavirus genera. Italics, averages of pairwise amino acid sequence identities between members of the same picornavirus genus. NA, not applicable.

Results

Sequencing of new virus. Virions in CPE-causing cell culture supernatant were first purified from larger particles by filtration, and contaminating cellular DNAs were digested by DNase, while viral nucleic acids remained protected from digestion within viral capsids (1, 15). Total nucleic acids were then extracted and purified, and viral RNA was amplified following reverse transcription with an oligonucleotide primer containing a randomized 3′-end sequence followed by second-strand synthesis using the same primer with Klenow DNA polymerase. The fixed portion of this oligonucleotide was then used to PCR amplify the DNA fragments with complementary ends (see Materials and Methods). The random PCR product was purified and subcloned, and the inserts of 96 plasmids were sequenced. Sequence data were analyzed using tBLASTx (i.e., translated in all six frames and aligned against the entire similarly translated GenBank database).

The viral RNA-derived sequences revealed the presence of four contigs showing sequence similarities to picornavirus proteins with amino acid identities ranging from between 22 and 41%. Except for the extremities of the contigs, each base of the contigs was sequenced at least twice from different overlapping subclones. The three gaps between the four contigs were amplified from the randomly primed cDNA using primers located within contigs followed by direct PCR product sequencing. The 5′ and 3′ extremities of the viral genome were acquired using rapid amplification of cDNA ends (see Materials and Methods). Relationship to other picornaviruses. The SePV-1 genome was 6,693 nucleotides long, excluding the poly(A) tail, and encoded a 2,027-amino-acid-long polyprotein. The genome showed a relatively low G+C content of 44% [excluding the poly(A) tail] compared to other picornaviruses. Its composition was most similar to those of HPeV and LVs in the Parechovirus genus (40% and 42%, respectively), DHV (44%), teschoviruses (44% to 45%), and equine rhinitis A virus equine rhinitis A virus (46 to 48%); greater than that of HAV (38%); and generally lower than those of other picornaviruses (e.g., FMDV [52 to 55%], cardioviruses [48 to 49%], enteroviruses [39 to 49%], and kobuviruses [55% to 60%]). To investigate its sequence relationship with other picornaviruses, alignments of the SePV-1 polyprotein were made in the P1 and 2C-3D regions of the genome with representative members of each picornavirus genus (listed in Materials and Methods). Regions of amino acid sequence similarities were most apparent in the nonstructural region of the SePV-1 genome, such as the region surrounding highly conserved motifs in the active site of the RNA-dependent RNA polymerase. Alignment of the 5′ region was more problematic, with several regions with no detectable homology between genera and a frequent necessity to introduce long gaps to align homologous sites. Although it is possible that both the 3D and P1 alignments could be optimized further, this would have only a small effect on sequence similarity values (Table 1 and Fig. 1A) or phylogenetic relations (Fig. 1B).
Both the P1 and 3D regions of SePV-1 showed the greatest similarity to DHV (24% and 30%, respectively) and to the parechoviruses LV and HPeV (26% and 29%) (Table 1). Similarity was less than 13% (P1) and 24% (3D) with all other picornavirus genera (Table 1). The level of sequence identity between SePV-1 and other picornaviruses was within the range seen between other genera (average of 17.6% and range of 10% to 35% for P1 and average of 23% and range of 19 to 56% for 3D) (Table 1). Many intervening regions between conserved motifs in the P1 and 2C-3D region displayed no detectable sequence similarities between genera. This is demonstrated by the sliding window of amino acid sequence divergence between different genera of picornaviruses (Fig. 1A).

In the 2C-3D structural gene region, pronounced dips in amino acid divergence between SePV-1 and other genera corresponded to conserved motifs within 2C, the end of 3C, and four regions in 3D (Fig. 1A). SePV-1 showed no close sequence relatedness to any of the picornavirus genera or the currently unclassified SVV or DHV sequences, with the possible exception of marginally greater similarity to parechoviruses and DHV (6, 17, 38, 39). In the P1 region, SePV-1 showed a more apparent greater degree of sequence similarity with parechoviruses and DHV (Fig. 1A, black and gray lines, respectively) than other genera.

For analysis of sequence relationships between SePV-1 and different picornavirus genera in the P1 and nonstructural protein regions A to H (positions indicated by solid bars in A). The same color code is used in both analyses.
and 6167 (3D region G), and 6263 to 6521 (3D region H) (Fig. 1A). Phylogenetic trees constructed from picornavirus sequences from each region demonstrate consistent branching patterns with generally good differentiation of the nine currently classified genera, although the proposed “Sapelovirus” genus showed a substantially greater similarity to the Enterovirus genus than to other picornaviruses in both structural and nonstructural regions (Fig. 1B). An exception to the otherwise consistent branching patterns of the trees was SVV, which generally grouped within or as an outlier to the cardioviruses but fell within the Aphthovirus genus in region E. In this region, members of the proposed “Sapelovirus” genus also became interspersed with enteroviruses (Fig. 1B).

The phylogenetic relationship between SePV-1 and other picornavirus genera was similarly highly consistent where all regions except E grouped closest to the parechovirus genus and DHV (Fig. 1B). In regions A, B, F, G, and H, these sequences grouped in bootstrap-supported clades. Given the anomalous position of SVV in region E, the lack of grouping of SePV-1 with parechoviruses in this region may be the result of a lack of phylogenetic resolution in this relatively short fragment. Indeed, analysis of regions D and E combined restored the SePV-1/parechovirus/DHV grouping (data not shown).

The relationships between SePV-1 with DHV and parechoviruses differed in different genomic regions. In the P1 region, the three virus groups formed a much more distinct, well-defined clade than in nonstructural regions of the genome (trees C to H), a difference also apparent from the similarity scan (Fig. 1A) (see above). Within the SePV-1/parechovirus/DHV grouping, DHV split earlier in the lineage in P1, while in the 2C-3D region, SePV-1 was consistently ancestral to the other viruses.

Despite the frequent bootstrap-supported grouping of SePV-1 with parechoviruses and DHV in the nonstructural 2C-3D region and the existence of identifiable homology elsewhere in the capsid-encoding region, the newly discovered virus was nevertheless highly divergent from each of these virus groups and indeed similar to sequence divergences that exist between other picornavirus genera, supporting the possibility that a new picornavirus genus was identified.

**Polyprotein.** A methionine codon starting at nucleotide position 507 was found in a standard Kozak context (RNNA UGG) and used to deduce the start of the polyprotein (22). Picornavirus structural and nonstructural proteins are typically generated by cleavage with virus-encoded proteinases. The hypothetical cleavage map of the SePV-1 polyprotein was derived from an alignment with other picornaviruses whose experimentally determined or hypothetical protease cleavage sites have been reported (Fig. 2). The presence of cleavage sites at interdomain junctions was sought based on the preference of picornaviruses for Q and E at the P1 position and a small amino acid residue (e.g., G, S, R, M, A, and N) at the P1/H11032 position (2, 14). Prediction of the cleavage sites in Fig. 2 were therefore determined by scanning for pairs of amino acids fitting this pattern around the interdomain regions inferred based on the amino acid alignment. The predicted cleavage sites result in a typical picornavirus gene order of VP0–VP3–VP1-2A1-2A2-2B–2C–3A–3B–3C–3D (Fig. 1A and 2).

The absence of a predicted maturational cleavage site between 1A and 1B (VP0) was consistent with HPeV, DHV (6), and LV (14) in which VP0 is not cleaved.

An analysis of the N terminus of the 2A protein of SePV-1 revealed a sequence corresponding to the canonical cotranslational cleavage site DXXN (Fig. 2). Cleavage at this site is cotranslationally mediated in cis by 2A and is present in cardioviruses, erbovirus, teschovirus, and aphthoviruses as well as in LV and DHV but not in HPeV, hepatoviruses, and the kobuvirus Aichi virus (7, 12, 18, 44, 45). Proteolytic cleavage at this site would therefore release a small 2A1 protein (Fig. 2). Based on the hypothesized cleavage sites, SePV-1 therefore appears to contain two structurally unrelated 2A proteins in a manner analogous to that of LV (13, 14, 17) but unlike HPeV, which contains a single 2A protein (11). Additionally, DHV
may encode a third 2A protein absent in both LV and SePV-1 (6, 17, 38).

**Untranslated terminal regions.** The length of the SePV-1 5’UTR until the polyprotein initiation AUG was unusually short at 506 nucleotides and included two terminal uracils required to covalently link the RNA to the VPg (3B) protein, a characteristic common to all picornaviruses (3). The secondary structure of the 5’UTR RNA was predicted using a combination of a thermodynamic folding energy minimization algorithm (MFOLD) and a stochastic context-free grammar method (PFOLD), independent algorithms that produced generally concordant results for the main structural features (Fig. 3). The analysis carried out is necessarily limited by the availability of only a single SePV-1 sequence from the 5’UTR; the

![RNA secondary structure of the SePV-1 5’UTR predicted by MFOLD and PFOLD. The positions of the a polypyrimidine tract (PPT), a 22-nucleotide region with sequence identity to DHV, and the polyprotein start codon are indicated by shaded boxes. The type IVB IRES has been annotated as previously proposed (9).](image-url)
The first stem-loop was remarkably large, with a duplex region of 28 base pairs, including an uninterrupted pairing of 15 consecutive bases. This stem-loop was much larger than the 5'-H11032-terminus loops of most other picornaviruses. The downstream stem-loop was similarly highly stable (duplex length of 9 nucleotides), with a terminal loop of only three unpaired bases. This configuration would likely prevent the formation of the pseudoknot (Cpk) found in a homologous position in HPeV (27) (Fig. 3). Of uncertain functional significance was a polypyrimidine tract (14 bases between positions 90 and 102) in the predicted unstructured region downstream of the second stem-loop (Fig. 3, gray box).

Although neither MFOLD nor PFOLD can predict tertiary RNA structure interactions, sequences at the 3'-end of the 5'-UTR can be convincingly modeled onto the previously proposed type IV IRES (9). Specifically, the predicted structure was most closely similar to IRES structures of members of the Sapelovirus and Teschovirus genera and to DHV, which was classified as type IVB IRES elements. SePV-1 thus contains the highly conserved IIIe stem-loop (with the unpaired GAYA sequence), a CpG dinucleotide pairing (IIIf), a longer-range interaction to form stem 1, and, finally, a pseudoknot pairing between positions 491 and 495 and upstream positions 467 to 471). As well as structural conservation, there were also several regions of sequence identity between SePV-1 and other viruses with type IV IRESs (Fig. 4), including both paired and unpaired bases, and a large number of covariant sites between

The proposed IRES structure for SePV-1 lacked a tetraloop containing the GNRA sequence, a common feature of most other type I, type II, and type IV IRES sequences. In type IV IRESs, the tetraloop is normally found as the four unpaired bases in stem-loop IIId, which was GGGG in SePV-1. However, a GGGG pentaloop was found in a homologous position in the simian picornavirus type IVB, indicating that either it is not absolutely required or the sequence requirement for this element is slightly less stringent than originally proposed (perhaps GNRR). The 3'-UTR was also the shortest one reported for a picornavirus, with a length of 34 nucleotides, the next biggest being from rhinoviruses, at 40 nucleotides. No folded RNA structures were detected.

**Myristylation and leader peptide.** Myristylation plays a crucial role in virion morphogenesis of most picornaviruses and involves the covalent linkage of myristic acid to an N-terminal glycine residue in a canonical Gxxxt/S motif. SePV-1 contains a putative myristylation sequence starting at position 16 relative to putative methionine amino termini. This location for myristylation most resembles that of HPeV at position 13, while DHV has a canonical sequence at position 31 and LV has one at position 3. The non-near-terminal position of the putative myristylation site of SePV-1 indicates that, as proposed for DHV and as shown for HPeV, myristylation may not occur and that alternative modifications at the amino terminus of the polyprotein may direct it to a lipid environment (14, 17, 36).

Leader peptides of variable length are found in the Cardiovirus, Aphthovirus, Teschovirus, Kobuvirus, and Erbovirus genera, for which various roles have been proposed (18, 32, 33, 44, 45). HPeV and LV are not thought to encode a leader peptide,
while DHV may encode a short leader peptide (6). SePV-1, like HPeV and LV, is not expected to encode a leader peptide.

**GORS.** Previous analyses of RNA structure formation among the different picornavirus genera indicated that members of the parechoviruses were unstructured, with no significant differences in MFEs between native and sequence order-scrambled sequences (34). The lack of genome-scale ordered RNA structure (GORS) in the Parechovirus genus contrasted with the high levels of sequence order-dependent RNA structure in the genera Aphthovirus, Kobavirus, and Teschovirus. Since the possession of GORS varies between genera, we conducted a more detailed comparison of RNA structure between human parechoviruses and LV (Parechovirus genus), DHV, and SePV-1 (Fig. 5). Each of the four parechovirus-like virus groups showed low MFEDs averaged over the length of the genome, similar to values previously determined for the “unstructured” enteroviruses and hepatoviruses and distinct from those genera with the Picornaviridae previously shown to possess GORS. Analysis of subgenomic regions revealed that the sequence order-dependent RNA structure within SePV-1 and other parechovirus-like viruses was in the 5′ and 3′ noncoding regions (data not shown). Recomputation of MFEDs for sequences confined to the coding region produced mean values of 0.06% for SePV-1 and similarly lower values for HPeV, DHV, and LV (~0.57% to 0.95%).

**Prevalence and diversity of SePV-1.** Cell culture supernatants derived from 18 out of 108 seals hunted in 2000 to 2002 showed CPE and were tested for SePV-1 RNA using RT-PCR of the complete VP1, 2A, and 2B regions of SePV-1. Eight supernatants were positive (7.4% of animals). Specificity of the PCR was confirmed by sequencing 1,428 bases of the amplification products, which also revealed the presence of minor variations between isolates, ranging from 0.2 to 3.7% (average, 1.5%) nucleotide differences.

**DISCUSSION**

We have genetically characterized the genome of a new picornavirus isolated from arctic ringed seals. Although the hunted animals from which SePV-1 was recovered appeared healthy, with normal body mass indexes, and did not exhibit any specific pathologies after cursory examination, it remains unknown whether this virus may, on occasion, be pathogenic. An extensive regulatory system of wildlife management precludes the inoculation of SePV-1 in this species to test for associated symptoms. Its high prevalence in 2000 to 2002 may represent recent spread or a more stable endemic relationship within this seal population. The absence of a GORS whose presence has been correlated with viral persistence in infected hosts (34) supports the possibility that this virus results in a transient infection, with its high prevalence therefore reflecting a recent epidemic spread. Whether this virus is transmissible to humans by hunting and consuming these seals, sometimes raw, is unknown and will require the testing of exposed populations for antibodies to SePV-1.

Picornaviruses have traditionally been identified and classified based on biophysical/antigenic properties. Viral genome sequences have also been used for taxonomy purposes. The family Picornaviridae is currently divided into nine genera (Aphthovirus, Cardiovirus, Enterovirus, Hepatovirus, Parechovirus, Rhinovirus, Erbovirus, Kobavirus, and Teschovirus) (18, 35). Rhinoviruses have recently been reclassified as a new species within the Enterovirus genus, while sapeloviruses, SVV, and DHV may become new Picornaviridae genera. The nearest genetic neighbors of SePV-1 are the parechoviruses (HPeV and LV) and the currently unclassified DHV, which has been recently proposed as the prototype member of a new picornavirus genus (6, 17, 38). Relative to the parechoviruses, SePV-1 has a more basal phylogenetic root than DHV1 in the nonstructural 2C and 3D regions, while DHV1 was more basal in the P1 region. SePV-1 is therefore divergent enough from other picornaviruses in both regions to qualify as the prototype of another picornavirus genus.

SePV-1 contains some but not all the characteristics of HPeV, LV, or DHV1. All four viruses show similar G+C contents (40 to 44%), but unlike the parechoviruses HPeV and LV, which have type II IRESs, SePV-1 and DHV1 contain 5′ UTR structures similar to those of sapeloviruses and are classified as type IVB (9). Alternative relationships between these viruses existed in other attributes. For example, like HPeV and DHV but unlike LV, its canonical myristylation site appears to be too far from the C terminus of VP0 to be functional. The two 2A proteins of SePV-1 resemble the genetic organization of LV 2A1 and 2A2 proteins and are distinct from HPeV’s single 2A or DHV’s three 2A proteins. Like DHV1 and LV but unlike HPeV, its 2A1/2A2 boundary contains the canonical site required for cotranslational cleavage. Like HPeV and LV but unlike DHV1, the SePV-1 genome appears to be missing a leader protein. SePV-1 therefore also exhibits a unique mixture of picornavirus genetic characteristics.

Microarrays consisting of highly conserved viral sequences...
have been used for the identification of both known and novel viral species (20, 40–42). The level of sequence similarity between SePV-1 and the most closely related oligonucleotides on the latest version of the microarray was lower than that observed when the severe acute respiratory syndrome coronavirus was identified by cross-hybridization with preexisting sequences (42) (data not shown). It is therefore unclear if such divergent picornaviruses as SePV-1 would be detected using microarrays based on preexisting viral sequences, but the inclusion of SePV-1 sequences on future microarrays will allow further searches for new viruses in this region of viral sequence space.

Picornaviruses can replicate in numerous mammals and birds, and picornavirus-like viruses have been found in insects (18). A recent study (4) using degenerate PCR primers targeting conserved amino acids in the highly conserved RNA-dependent RNA polymerase gene (26) also found highly diverse picornavirus-like viral sequences (42) (data not shown). It is therefore unclear if such picornavirus-like sequences (20, 40–42) are related to each other and to the H-rev107 family of picornaviruses as SePV-1 would be detected using microarrays based on preexisting viral sequences, but the inclusion of SePV-1 sequences on future microarrays will allow further searches for new viruses in this region of viral sequence space.

Picornaviruses can replicate in numerous mammals and birds, and picornavirus-like viruses have been found in insects (18). A recent study (4) using degenerate PCR primers targeting conserved amino acids in the highly conserved RNA-dependent RNA polymerase gene (26) also found highly diverse picornavirus-like viral sequences (42) (data not shown). It is therefore unclear if such picornavirus-like sequences (20, 40–42) are related to each other and to the H-rev107 family of picornaviruses as SePV-1 would be detected using microarrays based on preexisting viral sequences, but the inclusion of SePV-1 sequences on future microarrays will allow further searches for new viruses in this region of viral sequence space.

Picornaviruses can replicate in numerous mammals and birds, and picornavirus-like viruses have been found in insects (18). A recent study (4) using degenerate PCR primers targeting conserved amino acids in the highly conserved RNA-dependent RNA polymerase gene (26) also found highly diverse picornavirus-like viral sequences (42) (data not shown). It is therefore unclear if such picornavirus-like sequences (20, 40–42) are related to each other and to the H-rev107 family of picornaviruses as SePV-1 would be detected using microarrays based on preexisting viral sequences, but the inclusion of SePV-1 sequences on future microarrays will allow further searches for new viruses in this region of viral sequence space.

Picornaviruses can replicate in numerous mammals and birds, and picornavirus-like viruses have been found in insects (18). A recent study (4) using degenerate PCR primers targeting conserved amino acids in the highly conserved RNA-dependent RNA polymerase gene (26) also found highly diverse picornavirus-like viral sequences (42) (data not shown). It is therefore unclear if such picornavirus-like sequences (20, 40–42) are related to each other and to the H-rev107 family of picornaviruses as SePV-1 would be detected using microarrays based on preexisting viral sequences, but the inclusion of SePV-1 sequences on future microarrays will allow further searches for new viruses in this region of viral sequence space.

References


9. Hellen, C. U. T., and S. de Breney. 2007. A distinct group of hepacivirus/picovirus-like internal ribosomal entry sites in members of diverse Picornaviridae family. To our knowledge, it is also the first sequenced picornavirus shown to infect a marine mammal.


