Each of the Eight Simian Hemorrhagic Fever Virus Minor Structural Proteins is Functionally Important

Heather A. Vatter  
*Georgia State University*, hvatter@gsu.edu

Han Di  
*Georgia State University*, hdi1@student.gsu.edu

Eric F. Donaldson  
*University of North Carolina at Chapel Hill*, eric_donaldson@med.unc.edu

Ralph S. Baric  
*University of North Carolina at Chapel Hill*, rbaric@email.unc.edu

Margo A. Brinton  
*Georgia State University*, mbrinton@gsu.edu

Follow this and additional works at: [https://scholarworks.gsu.edu/biology_facpub](https://scholarworks.gsu.edu/biology_facpub)

This Article is brought to you for free and open access by the Department of Biology at ScholarWorks @ Georgia State University. It has been accepted for inclusion in Biology Faculty Publications by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact scholarworks@gsu.edu.
Each of the eight simian hemorrhagic fever virus minor structural proteins is functionally important

Heather A. Vatter a, Han Di a, Eric F. Donaldson b, Ralph S. Baric b, Margo A. Brinton a,*

a Department of Biology, Georgia State University, Atlanta, GA 30303, United States
b Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, United States

ARTICLE INFO
Available online 16 July 2014

Keywords:
Simian hemorrhagic fever virus
Subgenomic mRNAs
Minor structural proteins
Infectious clone

ABSTRACT
The simian hemorrhagic fever virus (SHFV) genome differs from those of other members of the family Arteriviridae in encoding two adjacent sets of four minor structural protein open reading frames (ORFs). A stable, full-length, infectious SHFV-LVR cDNA clone was constructed. Virus produced from this clone had replication characteristics similar to those of the parental virus. A subgenomic mRNA was identified for the SHFV ORF previously identified as 2b. As an initial means of analyzing the functional relevance of each of the SHFV minor structural proteins, a set of mutant infectious clones was generated, each with the start codon of one minor structural protein ORF mutated. Different phenotypes were observed for each ortholog of the pairs of minor glycoproteins and all of the eight minor structural proteins were required for the production of infectious extracellular virus indicating that the duplicated sets of SHFV minor structural proteins are not functionally redundant.

© 2014 Elsevier Inc. All rights reserved.

Introduction
Simian hemorrhagic fever virus (SHFV) was first identified in the 1960s when it was found to be the causative agent of fatal hemorrhagic fever outbreaks in captive rhesus macaque colonies in the United States, Russia and Europe (Tauraso et al., 1968). Experimental SHFV infections in several macaque species produced clinical symptoms including fever, anorexia, adipsia, cyanosis, skin petechial and nose bleeds and ultimately, led to death by 7 to 13 days after infection (Allen et al., 1968; London, 1977; Palmer et al., 1968). These symptoms closely resembled those induced by other types of hemorrhagic fever viruses, such as Ebola Zaire and Marburg viruses in rhesus macaques (Bray, 2005). Various species of African monkeys are the natural hosts of SHFV, and in baboons, vervets, patas monkeys and African green monkeys, SHFV infections are typically asymptomatic and often persistent (Gravell et al., 1986). Previous SHFV outbreaks in macaque colonies are thought to have been caused by inadvertent mechanical transfer of SHFV present in the blood of a persistently infected African monkey to a macaque followed by efficient transmission of the virus between macaques (Palmer et al., 1968). Humans exposed to SHFV-infected macaques did not develop disease symptoms or seroconvert (Dalgard et al., 1992; Palmer et al., 1968).

SHFV is a member of the family Arteriviridae that also includes equine arteritis virus (EAV), porcine reproductive and respiratory syndrome virus (PRRSV), lactate dehydrogenase elevating virus (LDV) and wobbly possum disease virus (Dunowska et al., 2012; Snijder and Meulenberg, 1998). The Arteriviridae, Coronaviridae and Roniviridae families are classified within the Order Nidovirales based on similarities in genome organization and replication strategy (Snijder and Kikkert, 2013). The SHFV genome is a 5′ capped and 3′ polyadenylated, positive-sense, single-stranded RNA of approximately 15.7 kb. The 5′ two-thirds of the genome encodes the ORF1a and ORF1ab polyproteins (Fig. 1A). The nonstructural proteins proteolytically cleaved from these polyproteins are required for the replication and transcription of the viral genomic and subgenomic (sg) RNAs (Beerens et al., 2007; Snijder and Meulenberg, 1998).

The viral structural proteins are encoded at the 3′ end of the genome and are expressed from a 3′ and 5′ coterminal, nested set of sg mRNAs. The 3′ most gene (ORF7) encodes the 15 kDa nucleocapsid (N) protein which forms disulfide-linked homodimers that interact to form the nucleocapsid (Dea et al., 2000). The interaction of the basic N-terminal domain of the N protein with genomic RNA is thought to facilitate packaging of the viral RNA into the nucleocapsid (Dea et al., 2000). The two other major structural proteins are the 19 kDa non-glycosylated membrane (M) protein and the 26 kDa major glycoprotein (GP5). These two virion envelope-associated proteins interact to form disulfide-linked heterodimers
that function in virion attachment (Delputte et al., 2002, 2007). The M protein is the most conserved arterivirus structural protein and is essential for progeny virion assembly (Delputte et al., 2002; Snijder and Meuleenberg, 1998). The macrophage-restricted protein sialoadhesin (CD169) serves as the arterivirus receptor and mediates virion internalization (Van Breedam et al., 2010; Welch and Calvert, 2010), while the macrophage-specific CD163 antigen is required for virion entry/uncoating (Calvert et al., 2007; Van Gorp et al., 2008). Recently, an additional ORF (GP5a) that starts at an alternative AUG within the major glycoprotein GP5 ORF was identified in the genomes of each of the known arteriviruses (Firth et al., 2011). An EAV reverse genetic study indicated that GP5a was not an essential protein but its absence reduced virus yields and GP5a was detected in purified PRRSV particles (Johnson et al., 2011; Sun et al., 2013).

The EAV, PRRSV and LDV genomes also encode four minor structural proteins: E, GP2, GP3 and GP4. The sequences of these proteins are less conserved between different arteriviruses than those of the major structural proteins. The minor structural GPs form complexes located on the surface of virions that are postulated to be involved in receptor binding and virion uncoating (Snijder and Kikkert, 2013). The E protein is translated from bicistronic sg mRNA 2 that also expresses GP2 and it has been proposed that E oligomers function as ion-channels during virion entry (Lee and Yoo, 2006; Snijder et al., 1999). The SHFV genome encodes two adjacent sets of four minor structural protein ORFs (Godeny et al., 1998).

A stable, full-length cDNA infectious clone for SHFV-LVR (SHFVic) was constructed. The replication characteristics of the virus produced by MA104 cells transfected with in vitro-transcribed SHFVic RNA were similar to those of the parental SHFV-LVR virus. A separate sg mRNA was identified as 2b. A set of mutant infectious clones, each constructed using a previously described strategy (Fang et al., 2007; Snijder and Meuleenberg, 1998). The macrophage-restricted protein sialoadhesin (CD169) serves as the arterivirus receptor and mediates virion internalization (Van Breedam et al., 2010; Welch and Calvert, 2010), while the macrophage-specific CD163 antigen is required for virion entry/uncoating (Calvert et al., 2007; Van Gorp et al., 2008). Recently, an additional ORF (GP5a) that starts at an alternative AUG within the major glycoprotein GP5 ORF was identified in the genomes of each of the known arteriviruses (Firth et al., 2011). An EAV reverse genetic study indicated that GP5a was not an essential protein but its absence reduced virus yields and GP5a was detected in purified PRRSV particles (Johnson et al., 2011; Sun et al., 2013).

The EAV, PRRSV and LDV genomes also encode four minor structural proteins: E, GP2, GP3 and GP4. The sequences of these proteins are less conserved between different arteriviruses than those of the major structural proteins. The minor structural GPs form complexes located on the surface of virions that are postulated to be involved in receptor binding and virion uncoating (Snijder and Kikkert, 2013). The E protein is translated from bicistronic sg mRNA 2 that also expresses GP2 and it has been proposed that E oligomers function as ion-channels during virion entry (Lee and Yoo, 2006; Snijder et al., 1999). The SHFV genome encodes two adjacent sets of four minor structural protein ORFs (Godeny et al., 1998).

A stable, full-length cDNA infectious clone for SHFV-LVR (SHFVic) was constructed. The replication characteristics of the virus produced by MA104 cells transfected with in vitro-transcribed SHFVic RNA were similar to those of the parental SHFV-LVR virus. A separate sg mRNA was identified for the SHFV ORF previously identified as 2b. A set of mutant infectious clones, each lacking the expression of one of the duplicated 3' ORFs, was made and their ability to produce infectious virions was analyzed. The data indicate that all eight of the minor structural proteins are required for the efficient production of infectious virions and suggest that SHFV particles are more complex than those of the other known arteriviruses.

Results

Construction of a full-length cDNA infectious clone of SHFV-LVR

A full-length cDNA clone of the SHFV, strain LVR, genome was constructed using a previously described strategy (Fang et al., 2006b; Nielson et al., 2003; Yount et al., 2000, 2002). Five overlapping cDNA fragments covering the genome were amplified from SHFV RNA by RT-PCR as described in the Materials and Methods section (Fig. 1B and C). A few nts in each fragment differed from the Genbank SHFV consensus sequence (Accession number AF180391.1) and initially each of these nts was corrected by site-directed mutagenesis of the appropriate fragment clone. The corrected individual fragment clones were then cut with PflMI, simultaneously ligated and cloned into pACYC184. The complete cDNAs in three clones were sequenced and no unexpected mutations were found. MA104 cells transfected with either 100 or 500 ng of in vitro transcribed SHFV RNA were observed daily for the development of cytopathic effect (CPE). At 120 h after transfection, cells were lysed and lysates analyzed for SHFV nucleocapsid and nsp1β proteins by Western blotting. The transfected cells did not show any obvious CPE and viral proteins were not detected in cell lysates through 120 h (data not shown). The harvested culture fluid was serially passaged four times in an attempt to recover virus but neither CPE nor intracellular viral protein was detected after any of these passages (data not shown).

Because the Genbank consensus sequence (Genbank Assesion number AF180391.1) of SHFV LVR was obtained by older sequencing methods from shotgun clones and it was likely that this sequence contained deleterious errors. Viral RNA extracted from parental SHFV LVR was subjected to 454 sequencing as described in the Materials and Methods section. In addition, each of the fragment clones was reamplified from SHFV LVR RNA by RT-PCR and sequenced. The same 18 nt differences compared with the previous Genbank consensus sequence were detected by both 454 genomic sequencing and individual fragment sequencing (Table 1). The Genbank sequence was updated based on 454 sequencing data (Genbank Assesion number AF180391.2). The reamplified fragment clones were then used to construct new full-length clones (SHFVic). The sequences of three independently constructed full-length clones were confirmed. Viral RNA was in vitro transcribed from linearized plasmid RNA and analyzed on an RNase-free denaturing agarose gel (Fig. 1D). Six RNA bands with sizes of about 15, 13, 12, 7, 4 and 3 kb were consistently detected. The 15 kb band was appropriately sized to represent the full-length viral genome RNA. The shorter products were likely the result of premature termination or aberrant transcription. MA104 cells transfected with either 100 or 500 ng of in vitro transcribed SHFVic RNA showed CPE by 72 h after transfection and Western blot analysis detected nsp1β in lysates from cells transfected with either concentration of viral RNA at 72 h after transfection (Fig. 2A). The higher amount of viral protein detected after transfection of 100 ng of RNA suggested the possibility that increased amounts of the shorter RNAs negatively affected the efficiency of viral replication. To confirm that extracellular infectious virus was produced, 100 μl of undiluted culture fluid from the transfected cells was passaged onto fresh MA104 cells. These cells showed CPE starting at 24 h after infection and Western blot analysis detected intracellular nsp1β protein in cell lysates harvested at 72 h (Fig. 2A).

<table>
<thead>
<tr>
<th>Position</th>
<th>GeneBank</th>
<th>New sequencing</th>
<th>Amino acid change</th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>511</td>
<td>C</td>
<td>T</td>
<td>P101L</td>
<td>nsp1α</td>
</tr>
<tr>
<td>849</td>
<td>G</td>
<td>T</td>
<td>V214F</td>
<td>nsp1β</td>
</tr>
<tr>
<td>1658</td>
<td>G</td>
<td>A</td>
<td>silent</td>
<td>nsp1y</td>
</tr>
<tr>
<td>2503</td>
<td>C</td>
<td>G</td>
<td>G655A</td>
<td>nsp2</td>
</tr>
<tr>
<td>3726</td>
<td>C</td>
<td>–</td>
<td>10 aa frame shift</td>
<td>nsp2</td>
</tr>
<tr>
<td>3727</td>
<td>–</td>
<td>A</td>
<td>10 aa frame shift</td>
<td>nsp2</td>
</tr>
<tr>
<td>3785</td>
<td>A</td>
<td>G</td>
<td>silent</td>
<td>nsp2</td>
</tr>
<tr>
<td>4277</td>
<td>T</td>
<td>G</td>
<td>silent</td>
<td>nsp3</td>
</tr>
<tr>
<td>4895</td>
<td>A</td>
<td>G</td>
<td>silent</td>
<td>nsp4</td>
</tr>
<tr>
<td>5575</td>
<td>G</td>
<td>T</td>
<td>G1889V</td>
<td>nsp5</td>
</tr>
<tr>
<td>5695</td>
<td>G</td>
<td>A</td>
<td>R2018H</td>
<td>nsp5</td>
</tr>
<tr>
<td>5704</td>
<td>A</td>
<td>T</td>
<td>K20211</td>
<td>nsp6</td>
</tr>
<tr>
<td>5707</td>
<td>C</td>
<td>G</td>
<td>P2022R</td>
<td>nsp6</td>
</tr>
<tr>
<td>5727</td>
<td>A</td>
<td>T</td>
<td>I202IF</td>
<td>nsp6</td>
</tr>
<tr>
<td>6088</td>
<td>G</td>
<td>C</td>
<td>G1960A</td>
<td>nsp7</td>
</tr>
<tr>
<td>8276</td>
<td>T</td>
<td>G</td>
<td>S2879A</td>
<td>nsp9</td>
</tr>
<tr>
<td>8340</td>
<td>G</td>
<td>T</td>
<td>W2900L</td>
<td>nsp9</td>
</tr>
<tr>
<td>10359</td>
<td>T</td>
<td>C</td>
<td>L3737S</td>
<td>nsp11</td>
</tr>
</tbody>
</table>
replicate cultures. Similar kinetics of production and levels of intracellular nsp1β and nucleocapsid were detected by Western blotting in parental and SHFVic virus infected cell lysates (Fig. 2C). The data indicate that the replication efficiency of virus derived from the infectious clone is comparable to that of the parental LVR virus.

Analysis of SHFV sg mRNA production in infected cells

EAV produces six sg mRNAs in infected MA104 cells (Snijder and Meulenberg, 1998; Van Berlo et al., 1982) and sg mRNA 2 was subsequently shown to be bicistronic expressing both GP2 and E (Snijder et al., 1999). Because the sequence obtained for the 3′ end of the SHFV genome indicated the presence of a duplicated set of minor structural ORFs (Godeny et al., 1998), 9 sg SHFV mRNAs were predicted. However, evidence was obtained for only 8 due to the lack of detection of a PCR band containing a sg mRNA junction sequence for the second minor GP ORF (Godeny et al., 1998).

The first and second SHFV minor GP ORFs were therefore named ORF2a and ORF2b because it was assumed that these GPs were expressed from a bicistronic sg mRNA (Godeny et al., 1998). The existence of an ORF for the E’ protein overlapping the first minor GP ORF was not known at that time.

The sg mRNAs produced by MA104 cells infected with SHFVic virus were analyzed by Northern blotting using a DIG-labeled 5′ leader probe. The 15.7 kb genome and nine sg mRNA bands with sizes of 5.0, 4.4, 4.0, 3.5, 2.8, 2.6, 1.9, 1.2 and 0.6 kb were detected starting at 8 h after infection (Fig. 3A). The detection of 5.0 and 4.4 kb sg mRNAs suggested that separate sg mRNAs are expressed for the previously named ORF2a and ORF2b. To confirm that separate sg mRNAs were expressed for these two ORFs, a set of RT-PCR primers (Table 4) was designed that were predicted to amplify the junction regions of both ORF2a (now ORF2’ and ORF2b (now ORF3’) sg mRNAs (Fig. 3B, top). The RT-PCR products generated from total cell RNA isolated from MA104 cells infected with SHFVic virus for 24 h were separated on an agarose gel and...
visualized by ethidium bromide staining. A strong band of ~1000 bp and a weaker PCR band of ~500 bp were detected (Fig. 3B, bottom). These bands were excised and the PCR products present in each band were TA cloned and sequenced. The junction sequence identified in the predominant ~1000 kb band was identical to the one previously reported for sg mRNA 2α (data not shown) (Godeny et al., 1998). The junction sequence found in the weaker ~500 kb band is shown in Fig. 3C. Based on the detection of the 4.4 kb sg mRNA and current knowledge that the sg mRNA of the former ORF2α is bicistronic because it contains an E protein ORF, the SHFV ORF formerly named ORF2b was renamed ORF3α (Fig. 3A). To facilitate comparison of the structural ORFs among the arteriviruses, the duplicated sets of SHFV minor structural ORFs and proteins were renamed ORF2a’ (GP2α), ORF2b’ (E), ORF3α (GP3), ORF4’ (GP4) and ORF2a (GP2), ORF3 (GP3), ORF4 (GP4) (Fig. 1A) (Snijder and Kikkert, 2013).

The relative intensities of the individual sg mRNA bands differed with sg mRNAs 2’, 2, 5, 6, and 7 being the most abundant. Previous Northern blot analyses of the SHFV sg mRNAs done before it was known that the SHFV genome encoded additional 3’ ORFs detected only these five strong sg mRNA bands (Godeny et al., 1995; Zeng et al., 1995). Interestingly, the relative abundances of the sg mRNAs for the orthologs in the duplicated sets of SHFV 3’ ORFs was similar with sg mRNAs 2’ and 2 being abundantly expressed while sg mRNAs 3’, 4’ and 4 were expressed at low levels.

**Analysis of the functional importance of the individual minor structural proteins**

The SHFV genome is predicted to encode two sets of minor structural proteins while the LDV, EAV, and PRRSV genomes encode a single set (Godeny et al., 1998). Recent data indicate that all four of the minor structural proteins of EAV and PRRSV are required for viral attachment/entry (Snijder and Kikkert, 2013; Tian et al., 2012). To analyze the functions of the individual SHFV minor structural proteins, the translation start codon for each of the 8 minor structural proteins was individually mutated in the appropriate genomic fragment clone without altering the amino acid coded in an overlapping ORF as described in the Materials and Methods section. Three full-length clones were then assembled for each of the minor structural protein knock-out mutants. In vitro transcribed viral RNA (100 ng) for each of the 3 clones was used as a positive control. Transfected cells were observed daily for CPE through 120 h after transfection. Cells transfected with SHFVic RNA showed no CPE by 120 h after transfection with any of these double mutant RNAs (Fig. 4C–E). A sustained low level of extracellular viral RNA would be expected if only a single cycle infection occurred and virus particles were produced from only the few cells transfected with a replication competent, full length copy of in vitro transcribed mutant viral RNA. Interestingly, between 24 and 72 h, the relative amounts of extracellular viral RNA were higher in culture fluids from cells transfected with ΔGP4’, ΔGP2 or ΔGP3 mutant RNA than in culture fluids from cultures transfected with wildtype RNA. This observation is consistent with the accumulation of non-infectious virus particles in the mutant RNA transfected culture fluids. Infectious wildtype virus would not be expected to accumulate initially due its ability to efficiently attach to and enter cells. Cells transfected with ΔGP4 mutant RNA produced increased levels of extracellular viral RNA by 24 h after transfection than cells transfected with the wildtype RNA suggesting that the particles produced by this mutant RNA were also not infectious (Fig. 4F). However, the level of ΔGP4 extracellular viral RNA decreased by 48 h suggesting a possible role for GP4 in virion stability. For each of the mutants, intracellular nsp1β levels in cell lysates collected at 120 h after transfection and at 120 h after a single blind passage were analyzed by Western blotting. Nsp1β was not detected in cells transfected with any of the mutant RNAs or in any of the passage 1 cells (Fig. 4J). The lack of detection of intracellular viral protein is consistent with a lack of virus spread and the levels of viral protein in the few productively transfected cells being too low to detect by Western blotting. Mutant infectious clones were next made that had the AUGs of both GP2’ and GP2, both GP3’ and GP3, or both GP4’ and GP4 mutated. No CPE was observed by 120 h after transfection with any of these double mutant RNAs or during four subsequent serial blind passages of the culture fluids. No extracellular viral RNA was detected in culture fluids harvested from cells transfected with any of the double mutant RNAs (Fig. 4G–I) and no detectable intracellular nsp1β protein was detected after transfection or passage (Fig. 4K).

No CPE was observed in cells transfected with ΔE’ or ΔE mutant viral RNA or during five subsequent serial blind passages of culture fluid. No significant increase in extracellular viral RNA above background levels was detected in culture fluids from cells transfected with either ΔE’ or ΔE mutant RNA (Fig. 5A and B). No intracellular nsp1β protein was detected at 120 h after transfection or at 120 h after passage 1 (Fig. 5C). The results suggest that both SHFV E’ and E proteins play an essential role in virus particle production.

**Discussion**

The EAV, PRRSV and LDV genomes encode four minor structural proteins, GP2, E, GP3 and GP4. The SHFV genome is unique in having a second set of the minor structural protein ORFs that are proposed to have been acquired by copy choice recombination...
between two SHFV genomes (Smith et al., 1997). It had also been suggested that the minor structural protein gene sequence duplication might be a laboratory artifact unique to the SHFV LVR strain (Tauraso et al., 1968). However, the additional set of SHFV 3' ORFs was reported to be present in SHFV isolates recently obtained from naturally infected Red colobus and Red-tailed guenon (Lauck et al., 2011, 2013) and also from persistently infected baboons (Vatter and Brentn, unpublished data).

Infectious clones for multiple strains of EAV and PRRSV have been constructed (Balasuriya et al., 2007; Fang et al., 2006a; Lv et al., 2008; Meulenbergh et al., 1998; Nielson et al., 2003; Van Dinten et al., 1997) and these infectious clones have been used to investigate viral replication as well as the induction and counter-action of host immune responses. SHFVic is the first SHFV infectious clone and the longest arterivirus RNA genome cDNA to date to be stably maintained within a single plasmid. The successful generation of SHFVic provides a unique tool for analyzing the functions of individual virus proteins and for identifying SHFV virulence determinants for macaques. In the present study, SHFVic was used for a reverse genetics analysis of the effects of separately knocking out the expression of each of the eight SHFV minor structural proteins in the duplicated sets.

Data previously obtained with EAV and PRRSV indicate that virus particles containing genome RNA are still produced when only the major structural proteins, N, M and GP5 are expressed (Wieringa et al., 2004a; Wissink et al., 2005; Zevenhoven-Dobbe et al., 2004). EAV GP2 and GP4 form heterodimers through intermolecular cysteine bridges (Wieringa et al., 2003a). After virion release, a post-assembly maturation occurs resulting in EAV GP3 becoming disulfide bonded to GP2-GP4 heterodimers forming GP2-GP3-GP4 heterotrimers (Wieringa et al., 2003b, 2004b). GP3 of Type 2 PRRSV strains is required for formation of the GP2/GP4 heterodimer (Dea et al., 2000). The PRRSV GP4 is required for the interaction of GP2-GP3-GP4 heterotrimers with the major glycoprotein GP5 (Das et al., 2010). The C-terminal domain of PRRSV GP4 functions as a GPI anchor and expressed GP4 colocalizes with CD163 in lipid rafts on the plasma membrane (Du et al., 2012). The E protein appears to be associated with GP2-GP3-GP4 heterotrimer on the surface of virions (Wieringa et al., 2004a) and to function as an ion channel during virus entry (Lee and Yoo, 2006). Knockout of the expression of GP2, GP3, GP4 or E blocked the incorporation of the minor GP heterotrimer into virion particles (Wieringa et al., 2003b, 2004a; Wissink et al., 2005; Zevenhoven-Dobbe et al., 2004).

Based on sequence homology between SHFV minor structural protein orthologs and the respective homologs of other arteriviruses, it was previously suggested that the two sets of SHFV minor structural glycoproteins might be functionally redundant (Godeny et al., 1998). However, in the present study different phenotypes were detected for the ΔGP2 and ΔGP2 mutant RNAs (Fig. 4A and D), for the ΔGP3 and ΔGP3 mutant RNAs (Fig. 4B and E) and for the ΔGP4 and ΔGP4 RNAs (Fig. 4C and F). No extracellular viral RNA was detected when either GP2 or GP3 expression was knocked out indicating that these two proteins are required for virion assembly, release and/or stability. However, the possibility that one or both of these proteins may also be involved in virion infectivity cannot be ruled out. Sustained low levels of extracellular viral RNA were produced when GP2, GP3 or GP4 expression was knocked out but no amplification of the viral RNA was detected at later times indicating that non-infectious virus particles were produced. It is possible that these three minor SHFV GP proteins form a complex that functions similar to the GP2-GP3-GP4 heterotrimer of other arteriviruses. Although both ΔGP4 and ΔGP4 mutant RNAs produced non-infectious virus particles, the production of these particles was not sustained when GP4 was not expressed suggesting that GP4 may play a role in virion particle stability. The finding that each of the SHFV minor GPs was required for the production of infectious virions as well as the observation of different phenotypes for the orthologs of the two sets of SHFV minor GPs indicate that the duplicated sets of SHFV minor structural proteins are not functionally redundant.

The LDV, EAV and PRRSV E proteins are each translated from the bicistronic sg mRNA 2. However, the location of the start codon of the E protein ORF relative to that of the GP2 protein ORF in sg mRNA 2 differs among these genomes. GP2 translation is initiated from the 5' AUG (ORF2a) and E translation is initiated at a downstream AUG (ORF2b) in the PRRSV sg mRNA 2, whereas in the sg mRNA 2 of LDV and EAV, the E protein is initiated from the 5' most AUG (ORF2a) and GP2 is initiated from a downstream AUG (ORF2b). In the SHFV sg mRNA 2', ORF2a' (GP2') is upstream of ORF2b' (E) similar to the order in PRRSV sg mRNA 2. In contrast, in the SHFV sg mRNA 2, ORF2b (E) is upstream of ORF2a (GP2) similar to the EAV and LDV gene order (see Fig. 1A). The E protein sequences of EAV, PRRSV and LDV contain a conserved N-terminal myristoylation site (Thaa et al., 2009). Inhibition of myristoylation during EAV replication decreased virus infectivity and plaque size but not the amount of extracellular virus suggesting that myristoylation of E is not required for virion budding or release but facilitates infectivity (Thaa et al., 2009). The SHFV E protein sequence contains the conserved N-terminal myristoylation site but the SHFV E' protein sequence does not. Extracellular viral RNA was not produced by cells transfected with either ΔE or ΔE' mutant RNA. These data suggest that both SHFV E and E' also play a role in virion production/stability. Although previous studies with EAV showed that knocking out GP2, GP3, GP4 or E expression blocked the incorporation of the minor GP heterotrimer into virion particles, a role for E in virion assembly or stability has not been reported (Wieringa et al., 2003b, 2004a; Wissink et al., 2005; Zevenhoven-Dobbe et al., 2004). Additional studies are required to determine whether SHFV E and/or E' proteins have functions during virus attachment and/or entry. The data obtained indicate that each of the proteins in the duplicated sets of SHFV minor structural proteins are required for the production of infectious virions. Among RNA viruses, SHFV particles appear to contain the largest number of envelop proteins.

Materials and methods

Cells and virus

The MA104 cell line was a gift from O. Nianan, Centers for Disease Control and Prevention (Atlanta, GA). These cells were cultured in Minimal Essential Medium (MEM) supplemented with 10% heat-inactivated fetal bovine serum and 10 μg/ml gentamicin at 37°C in 5% CO₂ atmosphere.

SHFV, strain LVR 42-0/6941 (American Type Culture Collection) was sequentially plaque-purified three times and then amplified once on MA104 cells (MOI of 0.01) to produce a stock pool. Virus stocks used for experiments were made by infecting confluent MA104 monolayers with the stock virus at an MOI of 0.2 and harvesting culture fluid at 32 h after infection. Clarified culture fluid (~ 10⁵ PFU/ml) was aliquoted and stored at −80°C.

Generation of overlapping genomic fragments

Genomic RNA was extracted from SHFV-LVR with TRI Reagent (Molecular Research Center, Cincinnati, OH) and 5 overlapping cDNA fragments were synthesized using SuperScript II reverse transcriptase (Invitrogen) and the forward primers listed in Table 2. The fragment junctions were selected based on the locations of PRI1 recognition sequences (CCANNNN/NTGG), Each
fragment cDNA was then amplified by PCR using an appropriate set of primers (Table 2). Cycling conditions were 48 °C for 30 min, 94 °C for 2 min, then 37 cycles of 94 °C for 30 s, 55 °C for 30 s and 65–72 °C for 1–5 min followed by a final cycle at 72 °C for 7 min to facilitate subsequent TA cloning. A unique SpeI cutting site (for insertion into the pACYC184 vector) and the SP6 promoter were added to the 5' end of fragment 1 in the forward primer and a 47 nt poly A tail, a unique PvuI site (for linearization) and a unique NotI site (for insertion into the pACYC184 vector) were added to the 3' end of fragment 5 in the reverse primer (Table 2). Each of the PCR fragments was gel purified and cloned into the pCR-XL-TOPO vector. These plasmids were maintained in TOP10 or INV110 Escherichia coli cells (Invitrogen) grown in LB media containing kanamycin (50 mg/ml). Multiple clones of each fragment were generated and sequenced.

**Assembly of full-length cDNA clones**

Plasmid DNA of each fragment was digested with PstI alone (fragments 2 and 4) or also with SpeI (fragment 1), Rsrl (fragment 5) or NotI and RsrlII (fragment 5). The pACYC184 vector was digested with XbaI and EagI to generate cohesive ends that were compatible with SpeI and NotI, respectively, and then gel purified. The digested SHFV cDNA fragments and vector DNA were simultaneously ligated using a T4 Rapid Ligation kit (Thermo Scientific) and the resulting DNA was used to transform XL10-Gold KanR cells (Invitrogen) grown in LB media containing chloramphenicol. The full-length cDNA clones were then used as template for random amplification of cDNA using a mMessage mMachine kit (Ambion) according to the manufacturer’s protocol. Reactions were incubated at 37 °C for 2 h then with DNase for 15 min, followed by LiCl precipitation and an ethanol wash of the pellet. The final RNA pellet was resuspended in nucleic-acid-free deionized water. An aliquot of the RNA was denatured with a formamide-based RNA sample buffer (Ambion) and analyzed for size and quality on an RNase-free denaturing gel.

MA104 cells were seeded in a six-well plate and grown overnight to <50% confluency. The cells were washed once with serum-free Opti-MEM (Gibco). The in vitro transcribed infectious clone RNA (100 or 500 ng) was mixed with 10 μl DMIRE-C (1, 2-dimyristoylxyloproplyl-3-dimethyl-hydroxy ethyl ammonium bromide and cholesterol) (Invitrogen) in serum-free Opti-MEM (Gibco) and added to the cells. DMIRE-C without RNA was used as a negative control and DMIRE-C with RNA extracted with TRI Reagent from parental SHFV-LVR virions was used as a positive control. The transfection media was removed after 4 h and 2 ml of fresh media were added to each well. Culture fluid (100 μl) harvested at 120 h was used to infect a fresh MA104 cell monolayer. This procedure was repeated for a total of three or four serial passages. In some cases, the harvested culture fluid was analyzed for viral RNA by qRT-PCR and cell lysates were analyzed for the presence of SHFV nucleocapsid and nsp1β proteins by Western blotting.

### Table 2

Primer used to generate overlapping cDNA fragments from SHFV-LVR RNA.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5'-3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fragment 1</td>
<td>act agt ATT TAG GTG ACA CTA TAG ATT AAA ATA AAA GTG TGA AGC</td>
</tr>
<tr>
<td>R-2343</td>
<td>CAA CCA CCC AAT GGT CAG AGT CAA GG</td>
</tr>
<tr>
<td>Fragment 2</td>
<td>TAC GTT GTG CGG CTT GAC TCT GAC</td>
</tr>
<tr>
<td>R-7981</td>
<td>GCT TGC CAG ACA GAA ATT TGA GAC TG</td>
</tr>
<tr>
<td>Fragment 3</td>
<td>TGG TCT CTC CTC AGG TCA TG</td>
</tr>
<tr>
<td>R-12306</td>
<td>AGT CAT GTT GCC TGT AAT TGT CTC</td>
</tr>
<tr>
<td>Fragment 4</td>
<td>TAT GTC TAT CTT CCA CTA CTT GTC</td>
</tr>
<tr>
<td>R-12190</td>
<td>AGG TAT TTA GAA AGT CAC GTA ACG</td>
</tr>
<tr>
<td>Fragment 5</td>
<td>TCT GCT GTG TGG TAA AAT GCT CTC ACT CAC</td>
</tr>
<tr>
<td>R-12652</td>
<td>ggg gcc gcc gat cgt (T)42 AAT TAT GGT ATA</td>
</tr>
</tbody>
</table>

a SP6 promoter consensus sequence is indicated by underlining.
b Recognition sites for restriction enzymes SpeI in F-SP6-1 and NotI and PvuI in R-15717 are in lowercase.

**454 sequencing of viral RNA**

Viral RNA was isolated from SHFV-LVR and processed as previously described (Donaldson et al., 2010). Briefly, culture fluid was first passed through a 0.45 μm filter to remove large contaminants and then centrifuged at 50,000 × g for 2 h. The pellets were resuspended in PBS overnight at 4 °C and then treated with DNase Turbo, Benzonase, and RNase One for 2 h at 37 °C to remove any RNA and DNA that was not protected due to being inside a virion. Virion RNA was then extracted using a Qiagen Viral RNA kit (Qiagen, Valencia, CA) according to manufacturer’s instructions. Extracted RNA was reverse transcribed, purified, and amplified as described previously (Donaldson et al., 2010). Briefly, the RNA was reverse transcribed with Superscript III (Life Technologies, Grand Island, NY) according to manufacturer’s instructions using random hexamers with a unique barcode. A Klenow reaction was performed to generate dsDNA from the cDNA. The DNA was then incubated with phosphatase and exonuclease to remove phosphates and unincorporated primers or single stranded bases, respectively. A total of 10 μl from the Klenow reaction was used as template for random amplification using the unique barcode, without the random hexamer portion, as a primer. The SISPA reactions were gel purified using a QiaQuick Gel Extraction Kit (Qiagen, Valencia, CA). The purified DNA samples were submitted to the Research Development Sequencing facility for sequencing using the Roche 454 Life Science FLX Titanium chemistry (454 Life Sciences of Roche, Branford, CT) for the first sequence run. The second sequence run used the Roche
Sequences derived from Roche 454 sequencing were binned based on the barcode sequence added during the SISPA preparation, and the barcode sequences were trimmed from the reads in each unique bin, representing individual samples. Each sample bin was mapped to the SHFV reference genome (accession no. AF180391.1) using the CLC Genomics Workbench version 5.1 (CLCBio, Aarhus, Denmark; http://www.clcbio.com) using default settings. Consensus sequences were determined for each contig using the 50% majority rule. The 454 sequence bins for each sample were also assembled by *de novo* assembly to determine if the two assembly approaches produced similar results and to identify sequences of unknown identity. The consensus sequences derived by *de novo* assembly were used to query the non-redundant nucleotide database using BLAST within the CLC Genomics Workbench interface. Differences between the reference genome and the genome derived by sequencing were annotated using a multiple sequence alignment generated using the CLC Genomics Workbench.

**Construction of minor structural protein knock out mutants**

To generate individual mutant infectious clones that did not express one of the eight minor structural proteins and primers were designed to mutate the start codon of one ORF without changing the amino acid sequence of the protein expressed from an overlapping ORF (Table 3). Mutations were introduced
Fig. 4. Effects of knocking out the expression of individual minor structural glycoprotein ORFs on extracellular viral RNA levels. (A–I) Analysis of extracellular viral RNA levels. Culture fluids were harvested at the indicated times after transfection with 100 ng of wildtype (gray bars), replication defective SAA mutant (white bars) or (A) ΔGP2, (B) ΔGP3, (C) ΔGP4, (D) ΔGP2, (E) ΔGP3, (F) ΔGP4, (G) ΔGP2Δ/2, (H) ΔGP3Δ/3 or (I) ΔGP4Δ/4 mutant (black bars) in vitro transcribed viral RNA. Extracellular viral RNA was extracted from 200 ul of harvested culture fluid and then quantified by qRT-PCR using nucleocapsid region primers and probe and a standard curve generated with a known amount of in vitro transcribed viral RNA. Each sample was assayed in triplicate. Error bars represent SD. Values shown are representative of two independent experiments. The complete range of values for the extracellular wildtype viral RNA levels produced is shown only in panel A. (J and K) Western blots of MA104 cell lysates collected 120 h after a blind passage infection. Viral proteins were detected with anti-nsp1β antibody. The blots shown are representative of data from three independent experiments. Actin was used as a loading control.
using the appropriate primer set (Table 3), infectious clone fragment [Fragment 3 for ORFs 2a (ΔGP2), 2b (ΔE) and 3 (ΔGP3)]. Fragment 4 for ORF4 (ΔGP4) and Fragment 5 for ORF2b (ΔE), 2a (ΔGP2), 3 (ΔGP3) and 4 (ΔGP4)] and a Quick-Change Lightning Site-Directed Mutagenesis kit (Agilent Technologies) according to manufacturer's protocol. Briefly, wild type fragment constructs were subjected to PCR using the mutagenic primers, the PCR products were then digested with DpnI to cleave the parental methylated and hemimethylated templates, and the mutant fragment DNAs were transformed into XL10-Gold ultra-competent cells. The mutations and fragment sequences were confirmed by sequencing. One mutant fragment and four wildtype fragments were then digested with the appropriate restriction enzymes, simultaneously

Table 3

<table>
<thead>
<tr>
<th>ORF (gene)</th>
<th>Primer sequence (5'-3')</th>
</tr>
</thead>
</table>
| ORF2a (GP2) | F CACTCTCTCCTCTGcTGAGTTTCTGTCCAGG  
            | R CCTGGACAGAAACTCAgCAGAGGAGAGGG |
| ORF2b (ΔE) | F CACTCTCTCCTCCTCAAGTACTAcGATCACACCCACATTCATACTG  
            | R CAGTATGAATGTGGGTGTGATCgTAGTACTTGAAGGAGGAGAAGTG |
| ORF3 (ΔGP3) | F GAAGGAGAGTTCTGTACCcTGGTTTCAACTTGGTTTC  
              | R GAAACCAGATTGAAACCAgGGTACAGAACTCTCCTTC |
| ORF4 (ΔGP4) | F CCTAACAATCAAGCAATCcTGGAACGAGACAATGATG  
            | R CATCATTGTCTCGTTCCAgGATTGCTTGATTGTTAGG |
| ORF2b (ΔE) | F CGATCTTTAGAGCATTCTATTCTTAgTGGGTTCTATACTCACCCACATC  
            | R GATGTGGGTGAGTATGACCCAcTAAGAATAGAATGCTCTAAAGATCG |
| ORF2a (ΔGP2) | F GCCTTCCACCAcGCTGTCCACGAGCTCC  
            | R GGAGCTCGTGGACAGCgTGGTGGAAGGC |
| ORF3 (ΔGP3) | F CAGTTATCTCACTAACCCcTGGATGTCCGTGG  
            | R CCACGGACATCCAgGGGTTAGTGAGATAACTG |
| ORF4 (ΔGP4) | R CCAGTTTGATAACGAGACATTcTGCTACATTGCAAACCTTAC  
            | F GTAAGGTTTGCAATGTAGCAgAATGTCTCGTTATCAAACTGG |

Effects of the mutations on an overlapping ORF

CTCTCCTCTGcTGAGTTTCTGTCCAGG (ORF2a): ATG to CTG (knock out start codon)
TACTCTGCTCA (ORF2b): ATG to ACC (knock out start codon)
TACTCTGCTCA (ORF2a): TAT to TAC (Tyrosine to Tyrosine)
GACCCTCGGT (ORF3): ATG to CTG (knock out start codon)
GACCCTCGGT (ORF3): TAA to TAG (STOP codon to STOP codon)
CACCCACCTCACTAACCCcTGGATGTCCGTGG (ORF4): ATG to CTG (knock out start codon)
CACCCACCTCACTAACCCcTGGATGTCCGTGG (ORF4): TAA to TAG (STOP codon to STOP codon)
CACCCACCTCACTAACCCcTGGATGTCCGTGG (ORF4): CCA to CCC (Proline to Proline)
CACCCACCTCACTAACCCcTGGATGTCCGTGG (ORF4): CAT to CAC (Histidine to Histidine)
AACCCTCGCGATATCGCCCAgAATGTCTCGTTATCAAATCCACATC (ORF4): ATG to CTG (knock out start codon)
AACCCTCGCGATATCGCCCAgAATGTCTCGTTATCAAATCCACATC (ORF4): TTA to TTC (Leucine to Leucine)

Fig. 5. Effects of knocking out the expression of the SHFV E' or E protein ORF on extracellular viral RNA levels. (A and B) Analysis of extracellular viral RNA levels. Culture fluids were harvested at the indicated times after transfection with 100 ng of wildtype (gray bars), replication defective SAA mutant (white bars) or (A) ΔE' or (B) ΔE mutant (black bars) in vitro transcribed viral RNA. Extracellular viral RNA was extracted from 200 ul of harvested culture fluid and then quantified by qRT-PCR using nucleocapsid region primers and probe and a standard curve generated with a known amount of in vitro transcribed viral RNA. Each sample was assayed in triplicate. Error bars represent SD. Values shown are representative of two independent experiments. The complete range of values for the extracellular wildtype viral RNA levels produced is shown only in panel A. (C) Western blots of MA104 cell lysates collected 120 after transfection with the indicated in vitro transcribed viral RNAs or 120 h after a blind passage infection. Viral proteins were detected with anti-nsp1β antibody. The blots shown are representative of data from three independent experiments. Actin was used as a loading control.
assembled, and cloned into the pACYC184 vector as described above. The resulting plasmid was used to in vitro transcribe mutant genome RNA for transfection into MA104 cells.

**Plaque assay**

Plaque assays were performed on confluent monolayers of MA104 cells in six-well plates. Harvested culture fluids were clarified by centrifugation at 1000 rpm for 5 min at 4 °C, serially diluted 10-fold in growth media and 100 μl of each dilution was added per well. Each sample was assayed in duplicate. After adsorption for 1 h at room temperature, the inoculum was removed, 2 mls of overlay media [1% SeaKem ME agarose (Bio-Whittaker Molecular Applications) mixed 1:1 with 2 × MEM containing 5% FCS] were added per well and plates were incubated at 37 °C for 72 h. After removal of the agarose plugs, cells were stained with 0.05% crystal violet in 10% ethanol.

**Western blot**

Culture media was removed from infected MA104 monolayers in 6 well plates and RIPA buffer (1 × phosphate-buffered saline, 1% Nonidet P-40, 0.5% sodium deoxycholate, and 0.1% SDS) containing Halt protease inhibitor cocktail (Thermo Scientific) was added to lyse the cells. Lysate proteins were separated by SDS-PAGE and then transferred to a nitrocellulose membrane by electrophoresis. Membranes were blocked with 1 × Tris-buffered saline (TBS), pH 8.0, containing 5% non-fat dry milk and 0.1% Tween 20 before incubation with a polyclonal primary antibody specific for an SHFV nucleocapsid peptide (NETHYVFAEPGDLRC), an SHFV nsp1 peptide (FAQKVITAF-PEGVLC) or actin (C-11; Santa Cruz Biotechnology, Santa Cruz, CA) in the presence of blocking buffer. Blots were washed with 1 × TBS and incubated with secondary antibody (horseradish peroxidase-conjugated anti-rabbit or anti-mouse; Santa Cruz Biotechnology) and then processed for chemiluminescence using a Super-Signal West Pico detection kit (Pierce, Rockford, IL) according to manufacturer’s protocol.

**Digoxigenin (DIG)-labeled SHFV 5’ leader probe**

PCR primers were designed to amplify a region in the 5’-leader (genome nts 7–200) for use as a template for making a 5’-leader probe (Table 4). The reverse primer included a T7 promoter. The PCR product generated was validated by sequencing and then in vitro transcribed with T7 RNA polymerase in the presence of DIG-labeled UTP using a DIG Northern Starter Kit (Roche) according to manufacturer’s protocol. The concentration of the DIG-labeled probe was determined by a dot-blot assay using dilutions of a DIG-labeled human actin RNA probe of known concentration as a reference (Roche). Briefly, 1 μl of each dilution was spotted onto an Amersham Hybond-N+ membrane (GE Healthcare) and then UV-crosslinked. The membrane was blocked with DIG blocking solution (Roche), incubated with anti-DIG antibody (1:10,000 dilution, Roche) and then with CDP-Star (Roche). The signal was detected with an LAS4000 mini Luminescent Image Analyzer (GE Healthcare). Spot intensities were determined using Multi Gauge V2.3 software.

**Northern blot hybridization**

MA-104 cell monolayers in 6 well plates were either mock infected or infected with SHFV infectious clone virus at an MOI of 1. At different times after infection, total cell RNA was extracted with TRI reagent (Molecular Research Center, Inc.). One μg of RNA was mixed with NorthernMax formaldehyde loading dye (Ambion), denatured at 80 °C for 10 min and then separated on 1% formaldehyde agarose gel for 2.5 h at 100 V. RNA markers (Millennium Markers-Formamide, Ambion) were run on one lane of the gel. RNA was transferred overnight by capillary action onto an Amersham Hybond-N+ membrane (GE Healthcare) and the RNA was then UV-crosslinked to the membrane. The lane containing RNA standards was cut from the membrane and stained with methylene blue. The rest of the membrane was pre-hybridized with DIG Easy Hyb buffer (Roche) at 68 °C for 30 min and then hybridized with 100 ng/mL of a DIG-labeled, denatured RNA probe at 68 °C overnight. The hybridized membrane was washed first with low stringency buffer (2 × SSC + 0.1% SDS) at room temperature, then with high stringency buffer (0.1 × SSC + 0.1% SDS) at 68 °C followed by blocking with DIG blocking solution (Roche). To detect the RNA bands, membranes were incubated with anti-DIG antibody (1:1,000 dilution) (Roche), developed with CDP-Star (Roche) and imaged with an LAS4000 mini Luminescent Image Analyzer (GE Healthcare).

**Analysis of extracellular viral RNA by quantitative real-time RT-PCR**

RNA in 200 μl of culture fluid harvested from infectious clone transfected MA104 monolayers was extracted with TRI Reagent. Reaction mixtures contained extracted RNA, the primer pair (1 μM) and probe (0.2 μM) in a final volume of 10 μl. Primer/probe sequences targeted either the nucleocapsid (forward primer 5′- tccacactcacacacatca-3′, TaqMan probe 5′-6FAMaaactcctgtcatg caggtMGBNFQ-3′ and reverse primer 5′-ccgcctctgtgtagt-3′) or helicase (ns5p) (forward primer 5′-ctgacacggctctgc-3′, TaqMan probe 5′-6FAMttgacgtttctacagagMGBNFQ-3′ and reverse primer 5′-ccgcaagtgcctcataa-3′) regions. Quantitative real time RT-PCR was performed and the data were analyzed as previously described (Scherbik et al., 2006). Extracellular SHFV genomic RNA was quantified using a standard curve generated with serial dilutions of a known concentration of SHFV RNA that had been in vitro transcribed with a MAXiScript SP6 transcription kit (Ambion) from either the nucleocapsid or ns5 gene region of the SHFVic cDNA, digested with DNase, extracted with phenol-chloroform, precipitated with ethanol and then quantified by UV spectrophotometry.

**Table 4**

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5′-3′)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5′-leader-probe-F</td>
<td>AATAAAAAGTGTGAAGCTTCCCTGCTTTCGTCCAGG</td>
</tr>
<tr>
<td>5′-leader-probe-R</td>
<td>CCGTAGCTATAGACGTCTACAGGCTGTCGCAATTCCCAAGGCAC</td>
</tr>
<tr>
<td>5′-Leader-RT-PCR-F</td>
<td>TAGCCCGATTGGATAAGC</td>
</tr>
<tr>
<td>3′-RT-PCR-R</td>
<td>CGTACGGTAGAATGATGG</td>
</tr>
<tr>
<td>sg mRNAs2 and 3′-RT-PCR-R</td>
<td>CCGTACGGTAGAATGATGG</td>
</tr>
</tbody>
</table>


