CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Elsevier - Publisher Connector

ELSEVIER
Available online at www.sciencedirect.com
ScienceDirect
Virology 369 (2007) 191 – 197
VIROLOGY
www.elsevier.com/locate/yviro

The complete nucleotide sequence of a New World simian foamy virus

Leonore Thümer [a], Axel Rethwilm [a,*], Edward C. Holmes [b], Jochen Bodem [a]

[a] Universität Würzburg, Institut für Virologie und Immunbiologie, Versbacher Str. 7, 97078 Würzburg, Germany
[b] Center for Infectious Disease Dynamics, Department of Biology, The Pennsylvania State University, PA 16802, USA

Received 12 June 2007; returned to author for revision 29 June 2007; accepted 11 July 2007
Available online 31 August 2007

Abstract

We determined the complete nucleotide sequence of the New World simian foamy virus (FV) from spider monkey (SFVspm). Starting from a conserved region in the integrase (IN) domain of the *pol* gene we cloned fragments of the genome up to the 5′ end of the long terminal repeat (LTR) into plasmid vectors and elucidated their nucleotide sequence. The 3′ end of the genome was determined by direct nucleotide sequencing of PCR products. Each nucleotide of the genome was determined at least two times from both strands. All protein motifs described to be conserved among primate FVs were found in SFVspm. At both the nucleotide and protein levels SFVspm is the most divergent primate FV described to date, reflecting the long-term phylogenetic separation between Old World and New World primate host species (*Catarrhini* and *Platyrrhini*, respectively). The molecular probes developed for SFVspm will allow the investigation of trans-species transmissions of this New World foamy virus to humans by serological assays.
© 2007 Elsevier Inc. All rights reserved.

Keywords: Foamy viruses; Retrovirus; New World monkeys; Primate evolution; Spider monkey virus

Introduction

Studies into the prevalence and molecular biology of FVs have attracted increased scientific interest in recent years. First, the risk of accidentally acquiring a primate FV infection through contact to non-human primates appears to be much higher than for any other retrovirus (Heneine et al., 2003; Heneine et al., 1998; Switzer et al., 2004). Although there is, so far, no indication of a pathogenic potential of these viruses and for human-to-human transmissions (Heneine et al., 2003), the potential of a new zoonotic infection has generated considerable concern (Switzer et al., 2004). New World monkeys are often kept as pets in Northern America and the risk to attract a FV infection through contact with these has been pointed out (Brooks et al., 2002). In addition, the hunting of New World primates poses a risk of zoonotically acquiring simian virus infections (Anonymous, 2007; Wolfe et al., 2005).

Second, *Spumaretrovirinae* have adopted a replication pathway distinct from *Orthoretrovirinae* (for reviews see Linial, 2007 and Rethwilm, 2005). Although still currently under investigation, it appears that the FV replication strategy represents a bridge between those of the orthoretroviruses and hepadnaviruses (Rethwilm, 2003).

Third, in the field of gene therapy FV vectors become increasingly attractive due to certain advantages they appear to have over orthoretrovirus vectors (for a recent review, see Rethwilm, 2007). However, much has still to be learned on these viruses before it comes to the application of the vectors to humans.

Finally, FVs are the genetically most stable viruses known to replicate through an RNA intermediate and can be taken as surrogate phylogenetic markers for the evolution of their hosts (Schweizer and Neumann-Haefelin, 1995; Switzer et al., 2005; Verschoor et al., 2004). However, this property of virus–host co-divergence was only demonstrated for the Old World non-human primates, with no available data from New World primates (Switzer et al., 2005). For these reasons we determined the complete nucleotide sequence of SFVspm (previously known as SFV-8), the first from a New World primate.

* Corresponding author. Fax: +49 931 201 49553.
E-mail address: virologie@mail.uni-wuerzburg.de (A. Rethwilm).

0042-6822/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2007.07.018

Results

Strategy for cloning SFVspm

The strategy of cloning the FV isolate from the spider monkey (*Ateles* ssp.) was analogous to that used to clone other FV genomes (Bodem et al., 1996; Herchenröder et al., 1994; Holzschu et al., 1998; Tobaly-Tapiero et al., 2000). The virus was obtained from the American Type Culture Collection (ATCC) and grown on human MRC-5 fibroblasts until the pathognomonic cytopathic effect (CPE) became visible. Taking advantage of the highly conserved IN domain an approximately 410 bp fragment was initially and specifically amplified from the DNA of infected cultures with primers developed by Schweizer and Neumann-Haefelin (1995) (data not shown). After determination of the amplimer nucleotide sequence an isolate specific reverse primer was designed and a forward primer located in the protease (PR) region of the *pol* gene that we found to be relatively conserved among primate FV sequences. With these two primers a 2.50 kb fragment spanning most of the *pol* gene was then obtained. Next, we amplified a 3.50 kb fragment spanning the complete *gag* gene with a PR-specific reverse primer and a forward primer derived from the primer binding site (PBS) that is conserved among FVs (Maurer et al., 1988). To obtain a 5′ LTR fragment we digested the cellular DNA from SFVspm-infected cells with *Hin*dIII that cuts in *gag* approximately 350 bp downstream of the PBS. After re-ligation, inverse amplification was carried out with a forward and a reverse primer, which both are located in *gag* 5′ to the *Hin*dIII site. An approximately 1.55 kb amplimer was generated and cloned. The sequence analysis revealed that we had isolated the complete 5′ LTR.

However, unintegrated DNA that circularized from the *Hin*dIII restriction site in *gag* to the blunted 5′ end of the viral DNA was the source of our amplification product. We regard the illegitimate ligation of a blunted LTR DNA end with a sticky one to occur relatively frequently with viral DNA from FV-infected cells since the 3′ LTR of the prototype FV (PFV) isolate was obtained similarly after *Bam*HI digestion (Rethwilm et al., 1987). The reason may be the large amount of non-integrated viral DNA in FV-infected cells generated by "late" reverse transcription (Moebes et al., 1997; Roy et al., 2003; Yu et al., 1999).

After cloning the 5′ part of the SFVspm genome from the start of the 5′ LTR to the integrase and then determining its nucleotide sequence, we generated the 3′ part of the viral genome in a similar way. A forward primer in the IN domain was designed and a reverse primer in the 3′ LTR. The PCR reaction generated an approximately 5.98 kb fragment that we were unable to clone into a plasmid vector. Therefore we decided to determine the DNA sequence the PCR fragment directly. A genetic map of SFVspm and some conserved motifs is shown in Fig. 1.

Nucleotide sequence of SFVspm

Using among all FVs conserved 18 nt PBS for the lysine[1,2] tRNA (Maurer et al., 1988), the overall length of the LTR was determined to be 1252 bp. Based on nucleotide motifs conserved among retroviruses the approximate length of U3, R, and U5 were deduced to be 930, 165, and 158 nt, respectively. While the length of R and U5 length appear to be in the range of all other FV genomes, the U3 is considerably shorter than the U3s from primate FVs and resembles in length more the non-primate FVs (Table 1).

In addition to the start of transcription in the LTR, FVs bear an internal promoter (IP), which is located at the end of the env gene sequence (Löchelt et al., 1993). Motifs of eukaryotic promoters, such as a TATA box, that most likely belong to the IP (Fig. 1) were identified upstream of the accessory open reading frames (ORFs).

The 3′ poly purine tract (PPT) marks the border of the 3′ LTR (Fig. 1). Similar to lentiviruses FVs possess an additional PPT located in the center of the genome (cPPT), which potentially can serve as an internal start of second strand transcription (Tobaly-Tapiero et al., 1991). However, the exact function is so far ill defined. A perfect match to the downstream 9 nt of the 3′ PPT was found at the expected central position of the SFVspm genome.

As all retroviruses FVs are diploid (Erlwein et al., 1997). The dimerization sequence at the 5′ end of the PFV genome contains a perfectly conserved 10 bp palindrome (Cain et al., 2001; Erlwein et al., 1997), the integrity of which has been shown to be critical for (pre-) genome dimerization and viral replication (Cain et al., 2001). Sequence determination of ten individual molecular clones revealed an 8 bp palindrome in SFVspm in a similar position that likely serves a function in dimerizing the SFVspm (pre-) genome (Fig. 1).

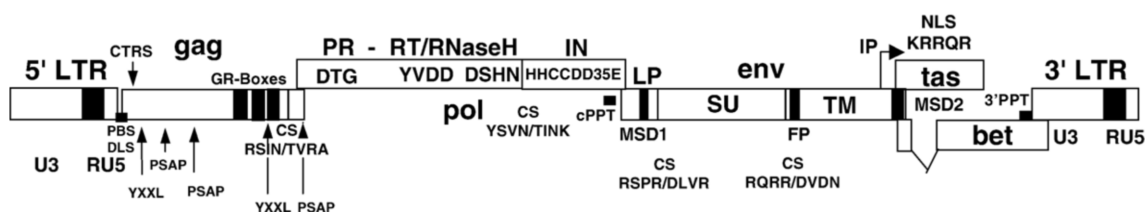

Fig. 1. Genetic map of the SFVspm provirus with conserved motifs. The primer binding site (PBS), dimer linkage site (DLS), cytoplasmic targeting and retention signal (CTRS), glycine–arginine-rich boxes (GR-boxes), proteolytic cleavage sites (CS), probable late budding and capsid assembly domains (PSAP and YXXL), residues important for enzymatic function (DTG, YVDD, DSHN, and HHCC, DD35E), the location of the central and 3′ poly purine tracts (cPPT and 3′PPT), the leader peptide (LP), surface (SU), and transmembrane (TM) proteins of Env, the internal promoter (IP), and the Tas protein-located nuclear localization signal (NLS) are indicated.

Table 1
Overview of FV LTRs and ORFs[a]

Virus/region	SFVspm	PFV	SFVcpz	SFVora	SFVmac	SFVagm	FFV	BFV	EFV
LTR	1251	1767	1760	1621	1621	1710	1353	1305	1449
U3	930	1423	1426	1286	1321	1333	1070	981	1124
R	164	190	180	177	168	215	134	173	171
U5	157	154	154	158	132	162	149	151	154
Gag	623	648	653	624	647	643	574	544	559
Pol	1148	1143	1146	1145	1149	1143	1156	1220	1153
Env	986	988	988	989	989	982	982	990	986
LP	129	126	126	126	126	126	127	123	122
SU	442	445	445	446	447	441	434	449	441
TM	415	417	417	417	416	415	419	418	423
Tas	324	300	300	278	308	298	209	249	249
Bet	374	482	490	464	487	469	387	435	440

[a] The length of the LTR regions is in nucleotides, and those of ORFs in amino acids. The length of the LP proteins was deduced from the primary sequences except for PFV and FFV where they have been determined experimentally (Duda et al., 2004, Geiselhart et al., 2004).

SFVspm amino acid sequences

While the majority of nucleic acid motifs and definitive splice sites of SFVspm need to be determined by functional characterization and transcript mapping studies, a variety of protein motifs could be unambiguously assigned to the deduced amino acid sequences based on sequence similarity to other FVs.

The Gag ORF has a length of 623 residues is more similar to the other primate than to the non-primate FVs (Fig. 1 and Table 1). The FV Gag protein is not cleaved into mature matrix, capsid, and nucleocapsid subunits (Linial and Eastman, 2003). Only a small peptide is cleaved from the Gag C-terminus. The cleavage site has been experimentally determined for PFV to be RAVN/TVTQ, which generates a 27 amino acid C-terminal peptide of approximately 3 kDa molecular weight (Flügel and Pfrepper, 2003; Zemba et al., 1998). Taking similar residues the cleavage site in SFVspm would be at aa 586 of the Gag protein and the resultant C-terminal peptide of 37 aa length and of approximately 3.7 kDa MW (Fig. 1). Three boxes rich in glycines and arginines (GR-boxes) have been identified and provisionally been characterized in the PFV isolate (Schliephake and Rethwilm, 1994; Yu et al., 1996). Three similar sequence motifs were found in SFVspm. The first is located between aa positions 454 to 483. This box I most likely harbors the nucleic acid-binding motif of SFVspm (Yu et al., 1996). Box II appears to be located between aa positions 516 to 534. Whether this element harbors a nuclear localization signal (NLS), as shown for PFV (Schliephake and Rethwilm, 1994), remains to be determined. Box III would run aa 556 to 572. Any function to the box III element has not been assigned yet. Furthermore, the presence of the GR-boxes appears to be a particular feature of primate FVs since FVs of lower animals do not bear such elements as far as this has been characterized (Bodem et al., 1998).

FV requires the presence of a late domain (L-domain) (Patton et al., 2005; Stange et al., 2005) for efficient budding. The L-domain sequence (PSAP) facilitates viral particle egress through interaction with the cellular export machinery (Freed, 2004; Stange et al., 2005). While the PSAP motif is highly conserved in primate FV Gag sequences its position within Gag is variable (Stange et al., 2005). Three perfectly conserved PSAP motifs were found in the SFVspm sequence, located at aa 223 to 226, aa 244 to 247, and aa 597 to 600 of the Gag ORF in Fig. 1. It is likely that at least one of these elements serves a similar function as the PSAP motif described for PFV (Patton et al., 2005; Stange et al., 2005). Interestingly, the downstream element was identified to be located in the C-terminal peptide that is cleaved from the main body of the Gag protein during maturation. It is unclear whether the location of this element allows for any L-domain function.

Two YXXL motifs were found at aa positions 107 and 538 in SFVspm Gag (Fig. 1). An YXXL motif has recently been shown for PFV to function unlike an L-domain and to be instead essential for correct particle assembly (Mannigel et al., 2007). The conserved R residue in Gag that is implicated in particle formation in the cytoplasma (Eastman and Linial, 2001) was located at SFVspm aa position 48, and a potential cytoplasmic targeting and retention signal (CTRS) was identified in SFVspm Gag downstream of aa position 36 (Cartellieri et al., 2005).

The Pol ORF has a length of 1148 aa (Fig. 1 and Table 1). As in all primate FVs the active center of the PR is of the sequence DT/SG. In FVs PR and the reverse transcriptase (RT) are not cleaved efficiently, if at all (Flügel and Pfrepper, 2003). The only viral PR-mediated cleavage that is known to occur in FV Pol proteins is the one between PR-RT and IN (Flügel and Pfrepper, 2003). Based on sequence similarity to other FV IN cleavage sites it is likely that the motif YSVN/TINK is used to separate the two Pol moieties in SFVspm (Fig. 1). This would result in 747 aa PR-RT and 401 aa IN proteins (Table 1). The RT active center (YVDD) is located at aa 308 to 311 in Pol and the catalytic important residues of the RNaseH domain (Boyer et al., 2004) at aa 666 to 732. The IN zinc-binding motif was identified at H^{809}-H^{813}-C^{843}-C^{846} and the IN active center (DD35E) at D^{875}-D^{932} and E^{968} (Fig. 1).

The Env ORF has 986 aa (Fig. 1 and Table 1). While the mature orthoretroviral Env is made up of two subunits, the surface (SU) and the transmembrane (TM) proteins (Swanstrom and Wills, 1997), spumaretroviral Envs are three subunit proteins, made up of a long leader peptide (LP), SU and TM

(Lindemann and Goepfert, 2003; Lindemann et al., 2001; Wilk et al., 2001). Although this has only been characterized in detail for the PFV and FFV isolates, it is likely that this topological feature applies to all FVs. The LP is a type II transmembrane protein and has the function to make specific contact to cognate viral capsids with its long cytoplasmic N-terminus (Geiselhart et al., 2003; Lindemann et al., 2001; Wilk et al., 2001). The cleavages between LP and SU and between SU and TM are mediated by cellular furin-like proteases (Duda et al., 2004; Geiselhart et al., 2004). An LP/SU-cleavage likely occurs C-terminal of R^{129} and the SU/TM-cleavage probably C-terminal of R^{571} (Fig. 1). Such cleavages would result in particle-associated 129 aa LP, 442 aa SU, and 415 aa TM proteins (Table 1).

Furthermore, some specific domains could be assigned to the individual Env proteins. The LP consists of N-terminal cytoplasmic and C-terminal extra-cellular regions, which are separated by a hydrophobic membrane-spanning region (Lindemann et al., 2001). This hinge region was located from aa 59 to 95 by hydrophobicity analysis. Following the predictions of Wang and Mulligan (68) the TM-located fusion peptide (FP) extends from residues 572 to 589, and the membrane-spanning domain (MSD) from 949 to 973. All FV Env precursors possess ER retention signals at the C-terminus of TM with the lysine in position −3 being the most important (Goepfert et al., 1995). This element was found in the expected position and consists of the residues KRK^{984}.

The Tas ORF has a length of 314 residues. FV transactivators are DNA-binding proteins that comprise an acidic activation domain at the C-terminus and a poorly characterized virus-specific DNA-interacting domain at the N-terminus which are implicated in binding IP and U3 sequences. Both domains are separated by a basic NLS (He et al., 1993; Rethwilm, 1995). These motifs could be assigned to the deduced protein sequence of the SFVspm Tas protein. The DNA-binding domain is poorly characterized for any FV and extends to aa 242, the NLS probably from aa 243 to aa 247, and the acidic activation domain is likely located downstream of the NLS with the core domain from aa 260 to aa 312 (He et al., 1993; Rethwilm, 1995).

Bet is a highly expressed and often immunodominant FV protein, which is generated by a multi-spliced mRNA containing coding exons from the first and the second accessory ORFs. The function of Bet has generated considerable controversy (Delebecque et al., 2006; Löchelt et al., 2005; Russell et al., 2005). It is possible that FV Bet proteins counteract host APOBEC cytidine deaminases by making use of an alternative strategy to the human immunodeficiency virus (HIV) Vif protein (Cullen, 2006). The proposed SFVspm Bet protein has a length of 375 residues (Fig. 1 and Table 1). Since functional domains have not yet been described for any of the FV Bet proteins, we could not assign those to SFVspm.

Evolutionary relationships of SFVspm

To determine the evolutionary relationship between SFVspm and the other simian foamy viruses we undertook a phylogenetic analysis of the translated *pol* amino acid sequences of six primate FVs (the amino acid sequences of the non-primate FVs were deemed too divergent for meaningful phylogenetic analysis). The maximum likelihood tree (Fig. 2) clearly reveals that SFVspm is the most divergent member of the simian foamy viruses identified to date, consistent with its infection of New World monkey host species (*Platyrrhini*), whereas the remaining primate FVs infect Old World monkeys and apes (*Catarrhini*).

This inference is supported by the identification in SPVspm of specific protein motifs, mainly in Gag, that are exclusive to FVs from primates. The L-domain PSAP is highly conserved among the primate viruses and is absent from the FVs of lower mammals. FFV, EFV, and BFV possess other sequences, which

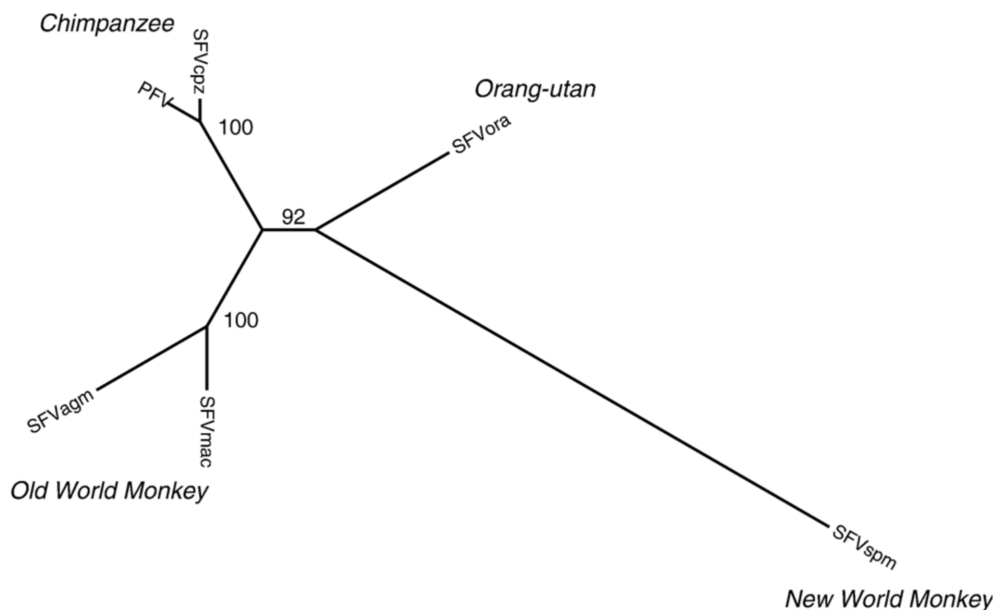

Fig. 2. Unrooted maximum likelihood (Tree-Puzzle) tree of pol amino acid sequences from six primate FVs including SFVspm. Numbers next to the specific branches depict quartet puzzling support values. All branch lengths are scaled to the number of amino acid substitutions per site.

Table 2
Percentages of amino acid similarities between SFVspm and different FVs

ORF/virus	PFV	SFVcpz	SFVora	SFVmac	SFVagm	FFV	BFV	EFV
Gag	49.2	49.8	49.3	47.7	47.8	55.9	44.9	45.4
Pol	78.8	79.8	79.1	79.9	78.7	73.8	75.3	76.2
Env (total)	71.6	71.2	68.9	69.0	67.7	62.5	62.5	63.0
LP	64.1	65.0	62.8	62.0	59.7	56.1	61.9	56.1
SU	71.1	69.8	66.7	68.8	67.0	60.5	55.1	58.0
TM	75.7	76.1	74.5	72.5	72.7	71.7	73.9	74.8

are believed to serve as L-domain motifs (Stange et al., 2005). Such elements, of the sequence YXXL, can also be found in all sequenced primate FV Gags (Patton et al., 2005; Stange et al., 2005). However, it has been shown for PFV recently that they are implicated in correct Gag assembly instead of promoting the budding process (Mannigel et al., 2007).

The GR-boxes at the C-terminus of Gag constitute the other Gag-related motifs that appear to be confined to the primate FVs. Little is known about the function of the three GR-boxes, which were first identified in PFV (Schliephake and Rethwilm, 1994). The nucleic acid-binding element in box I was found to be required for virus replication, while deletion of the NLS in box II appeared to have only modest effects on viral titers (Yu et al., 1996).

Discussion

Previous studies have shown that FV genomes from non-human primates can serve as surrogate markers for the phylogenetic relationships of their hosts (Schweizer and Neumann-Haefelin, 1995; Switzer et al., 2005). *Platyrrhini* (New World primates) diverged from a common ancestor with *Catarrhini* (Old World primates) between 43 and 35 million years ago (MYA) (Schrago and Russo, 2003; Steiper and Young, 2006). This deep divergence is also reflected in the divergent relationship of the spider monkey SFVspm from the other primate FVs (Fig. 2 and Table 2). Because of this host–virus co-divergence FV genomes are estimated to evolve at an extremely low rate of 1.7×10^{-8} per site and year, the lowest rate recorded for viruses having an RNA genome phase and approximates the pace of DNA virus evolution (Switzer et al., 2005) and which may be important when the application of FV vectors in gene therapy is considered.

The low in vivo replication rate of FVs is the most likely reason for their genomic stability (Murray and Linial, 2006; Switzer et al., 2005). This view is strengthened by the recent analysis of the fidelity of the PFV RT enzyme that was found to be in the range of HIV RT by in vitro assays (Boyer et al., 2007). When considering particular features of the FV replication strategy and reflecting the genome stability, we note that in functional terms FVs have a DNA genome and replicate through an RNA intermediate, while orthoretroviruses can be regarded as RNA viruses replicating through a DNA intermediate (Rethwilm, 2003). Furthermore, there is no evidence that the process of "late" reverse transcription involves a proof reading activity and the steps of generating viral DNA are likely to be similar between spuma- and orthoretroviruses (Rethwilm, 2003). However, the actual circumstances of reverse transcription are probably different between these two retroviral subfamilies (reviewed by Rethwilm, 2003 and 2005). Indeed, it may be that the fidelity of FV RT in its natural context is much higher than estimated by in vitro assays.

Primate FV genomes were estimated to be at least 30 million years old (Switzer et al., 2005). This calculation was based on Old World FV sequences. With the SFVspm sequence determined here we can extend this age to the separation between *Catarrhini* and *Platyrrhini*. According to our findings FVs acquired motifs in Gag, such as the L-domain PSAP and the GR-boxes, during co-divergence with their primate hosts and separating from the non-primate viruses. Since the ancestor of FVs is not known and the sequence of the prosimian FV from the Galago (SFVgal) remains to be elucidated, we cannot yet exclude the alternative possibility of non-primate FVs deleting specific Gag elements from a virus at the root of a hypothetical FV tree.

Materials and methods

Virus cultivation

Primary MRC-5 human fibroblasts were cultivated in modified Eagles medium that was supplemented with 10% fetal calf serum and antibiotics. The SFVspm isolated studied here which was originally collected by John Hooks et al. from a spider monkey brain in 1973 (Hooks et al., 1973) was obtained from ATCC (lot # 218093) and added to the cells. Virus and cells were cocultured and cells were split 1:5 after reaching confluency. After 3 days a CPE typical for FVs became visible.

Amplification and molecular cloning of SFVspm nucleic acids

DNA was extracted from SFVspm-infected cells and from uninfected MRC-5 cells on day five of culture using the DNeasy kit (Qiagen). The amplification of SFVspm was usually performed with 100–300 ng of DNA from infected cells, 15 pmol of both primers, 2.75 mM $MgCl_2$, and the Expand Long Template PCR System (Roche) for 35 cycles with 94 °C for 10 s, primer annealing at 50 °C for 30 s, and elongation at 72 °C for 2 min. From the 10th cycle onwards every polymerization step was elongated for 20 s. Amplimers were run in agarose gels, stained with crystal violet, the desired fragments were extracted with the Qiaquick Gel extraction kit (Qiagen), and inserted into the pCRII-Topo or pCR-XL-Topo cloning vectors (Invitrogen). Following screening of small-scale DNA preparations for the expected insert size by *Eco*RI digestion the viral inserts were DNA sequenced. The sequences of the primers used for amplification are available as supplementary information under http://viminfo.virologie.uni-wuerzburg.de/onlinematerial/Thuemer/pdf.

Sequence determination

Viral DNA sequences of the cloned virus were obtained using initially M13 forward or reverse primers on an ABI

PRISM 3100 Genetic Analyzer automated sequencing device (Applied Biosystems). The PCR amplimers of the 3′ region of the genome were directly sequenced using initially the amplification primers. Primer walking was then used to obtain the new sequence. All nucleotides of the SFVspm genome were determined at least two times for both strands. The sequences of the primers used for DNA sequencing and the complete SFVspm DNA sequence are available as supplementary information under the web address given above. The nucleotide sequence of SFVspm has also been deposited to GenBank under the accession number EU010385.

Evolutionary analysis

Sequence comparisons to other FVs were made using ClustalW and published sequences of PFV (prototype FV; accession number Y07725); SFVcpz (Chimpanzee FV; U04327); SFVora (Orang Utan FV; AJ544579); SFVmac (Macaque FV; X54482); SFVagm (African Green Monkey FV; M74895). The sequences from other mammalian FVs were considered too divergent for meaningful phylogenetic analysis.

A phylogenetic tree of these sequences was inferred from the translated amino acid sequences of the most highly conserved *pol* gene (alignment of 1152 residues) using the maximum likelihood (ML) method available in the Tree-Puzzle package (Strimmer and von Haeseler, 1996). The tree was inferred using the WAG+G model of amino acid substitution with the shape value (0.5) of the *G*-distribution estimated from the empirical data. Quartet puzzling was used to assess the support for specific nodes on the tree.

Finally, protein hybrophobicity was analyzed with the pepwindow software. The protein similarities among the viruses are shown in Table 2 and were calculated with the EMBOSS software at the European Molecular Biology Open Software Suite (Rice et al., 2000). Splice sites were predicted using NNSPLICE (version 0.9) (Reese et al., 1997).

Acknowledgments

We thank Benedikt Weissbrich and the institute's sequencing team for help with DNA sequencing.

This work was supported by DFG grants RE627/7 and SFB479 to A.R.

References

Anonymous, 2007. Going to the pot—the neotropical bushmeat crisis and its impact on primate populations.

Bodem, J., Löchelt, M., Winkler, I., Flower, R.P., Delius, H., Flügel, R.M., 1996. Characterization of the spliced pol transcript of feline foamy virus: the splice acceptor site of the pol transcript is located in gag of foamy viruses. J. Virol. 70, 9024–9027.

Bodem, J., Zemba, M., Flügel, R.M., 1998. Nuclear localization of the functional Bel 1 transactivator but not of the gag proteins of the feline foamy virus. Virology 251, 22–27.

Boyer, P.L., Stenbak, C.R., Clark, P.K., Linial, M.L., Hughes, S.H., 2004. Characterization of the polymerase and RNase H activities of human foamy virus reverse transcriptase. J. Virol. 78, 6112–6121.

Boyer, P.L., Stenbak, C., Hoberman, D., Linial, M., Hughes, S.H., 2007. Fidelity of the prototype primate foamy virus (PFV) RT compared to HIV-1 RT.

Brooks, J.I., Rud, E.W., Pilon, R.G., Smith, J.M., Switzer, W.M., Sandstrom, P.A., 2002. Cross-species retroviral transmission from macaques to human beings. Lancet 360, 387–388.

Cain, D., Erlwein, O., Grigg, A., Russell, R.A., McClure, M.O., 2001. Palindromic sequence plays a critical role in human foamy virus dimerization. J. Virol. 75, 3731–3739.

Cartellieri, M., Herchenroder, O., Rudolph, W., Heinkelein, M., Lindemann, D., Zentgraf, H., Rethwilm, A., 2005. N-terminal Gag domain required for foamy virus particle assembly and export. J. Virol. 79, 12464–12476.

Cullen, B.R., 2006. Role and mechanism of action of the APOBEC3 family of antiretroviral resistance factors. J. Virol. 80, 1067–1076.

Delebecque, F., Suspene, R., Calattini, S., Casartelli, N., Saib, A., Froment, A., Wain-Hobson, S., Gessain, A., Vartanian, J.P., Schwartz, O., 2006. Restriction of foamy viruses by APOBEC cytidine deaminases. J. Virol. 80, 605–614.

Duda, A., Stange, A., Luftenegger, D., Stanke, N., Westphal, D., Pietschmann, T., Eastman, S.W., Linial, M., Rethwilm, A., Lindemann, D., 2004. Prototype foamy virus envelope glycoprotein leader peptide processing is mediated by furin-like cellular protease, but cleavage is not essential for viral infectivity. J. Virol. 78, 13865–13870.

Eastman, S.W., Linial, M.L., 2001. Identification of a conserved residue of foamy virus Gag required for intracellular capsid assembly. J. Virol. 75, 6857–6864.

Erlwein, O., Cain, D., Fischer, N., Rethwilm, A., McClure, M.O., 1997. Identification of sites that act together to direct dimerization of human foamy virus RNA in vitro. Virology 229, 251–258.

Flügel, R.M., Pfrepper, K.I., 2003. Proteolytic processing of foamy virus Gag and Pol proteins. Curr. Top. Microbiol. Immunol. 277, 63–88.

Freed, E.O., 2004. HIV-1 and the host cell: an intimate association. Trends Microbiol. 12, 170–177.

Geiselhart, V., Schwantes, A., Bastone, P., Frech, M., Löchelt, M., 2003. Features of the Env leader protein and the N-terminal Gag domain of feline foamy virus important for virus morphogenesis. Virology 310, 235–244.

Geiselhart, V., Bastone, P., Kempf, T., Schnölzer, M., Löchelt, M., 2004. Furin-mediated cleavage of the feline foamy virus Env leader protein. J. Virol. 78, 13573–13581.

Goepfert, P.A., Wang, G., Mulligan, M.J., 1995. Identification of an ER retrieval signal in a retroviral glycoprotein. Cell 82, 543–544.

He, F., Sun, J.D., Garrett, E.D., Cullen, B.R., 1993. Functional organization of the Bel-1 transactivator of human foamy virus. J. Virol. 67, 1896–1904.

Heneine, W., Switzer, W.M., Sandstrom, P., Brown, J., Vedapuri, S., Schable, C.A., Khan, A.S., Lerche, N.W., Schweizer, M., Neumann-Haefelin, D., Chapman, L.E., Folks, T.M., 1998. Identification of a human population infected with simian foamy viruses. Nat. Med. 4, 403–407.

Heneine, W., Schweizer, M., Sandstrom, P., Folks, T., 2003. Human infection with foamy viruses. Curr. Top. Microbiol. Immunol. 277, 181–196.

Herchenröder, O., Renne, R., Loncar, D., Cobb, E.K., Murthy, K.K., Schneider, J., Mergia, A., Luciw, P.A., 1994. Isolation, cloning, and sequencing of simian foamy viruses from chimpanzees (SFVcpz): high homology to human foamy virus (HFV). Virology 201, 187–199.

Holzschu, D.L., Delaney, M.A., Renshaw, R.W., Casey, J.W., 1998. The nucleotide sequence and spliced pol mRNA levels of the nonprimate spumavirus bovine foamy virus. J. Virol. 72, 2177–2182.

Hooks, J.J., Gibbs Jr., C.J., Chou, S., Howk, R., Lewis, M., Gajdusek, D.C., 1973. Isolation of a new simian foamy virus from a spider monkey brain culture. Infect. Immun. 8, 804–813.

Lindemann, D., Goepfert, P.A., 2003. The foamy virus envelope glycoproteins. Curr. Top. Microbiol. Immunol. 277, 111–129.

Lindemann, D., Pietschmann, T., Picard-Maureau, M., Berg, A., Heinkelein, M., Thurow, J., Knaus, P., Zentgraf, H., Rethwilm, A., 2001. A particle-associated glycoprotein signal peptide essential for virus maturation and infectivity. J. Virol. 75, 5762–5771.

Linial, M., 2007. Foamy viruses, In: Kniepe, D.M., Howley, P.M. (Eds.), 5th Ed. Fields Virology, vol. 2. Lippincot Williams & Wilkins, Philadelphia, pp. 2245–2262.

Linial, M.L., Eastman, S.W., 2003. Particle assembly and genome packaging. Curr. Top. Microbiol. Immunol. 277, 89–110.

Löchelt, M., Muranyi, W., Flügel, R.M., 1993. Human foamy virus genome possesses an internal, Bel-1-dependent and functional promoter. Proc. Natl. Acad. Sci. U. S. A. 90, 7317–7321.

Löchelt, M., Romen, F., Bastone, P., Muckenfuss, H., Kirchner, N., Kim, Y.B., Truyen, U., Rosler, U., Battenberg, M., Saib, A., Flory, E., Cichutek, K., Munk, C., 2005. The antiretroviral activity of APOBEC3 is inhibited by the foamy virus accessory Bet protein. Proc. Natl. Acad. Sci. U. S. A. 102, 7982–7987.

Mannigel, I., Stange, A., Zentgraf, H., Lindemann, D., 2007. Correct capsid assembly mediated by a conserved YXXLGL motif in prototype foamy virus gag is essential for infectivity and reverse transcription of the viral genome. J. Virol. 81, 3317–3326.

Maurer, B., Bannert, H., Darai, G., Flügel, R.M., 1988. Analysis of the primary structure of the long terminal repeat and the gag and pol genes of the human spumaretrovirus. J. Virol. 62, 1590–1597.

Moebes, A., Enssle, J., Bieniasz, P.D., Heinkelein, M., Lindemann, D., Bock, M., McClure, M.O., Rethwilm, A., 1997. Human foamy virus reverse transcription that occurs late in the viral replication cycle. J. Virol. 71, 7305–7311.

Murray, S.M., Linial, M.L., 2006. Foamy virus infection in primates. J. Med. Primatol. 35, 225–235.

Patton, G.S., Morris, S.A., Chung, W., Bieniasz, P.D., McClure, M.O., 2005. Identification of domains in gag important for prototypic foamy virus egress. J. Virol. 79, 6392–6399.

Reese, M.G., Eeckman, F.H., Kulp, D., Haussler, D., 1997. Improved splice site detection in Genie. J. Comput. Biol. 4, 311–323.

Rethwilm, A., 1995. Regulation of foamy virus gene expression. Curr. Top. Microbiol. Immunol. 193, 1–24.

Rethwilm, A., 2003. The replication strategy of foamy viruses. Curr. Top. Microbiol. Immunol. 277, 1–26.

Rethwilm, A., 2005. Foamy viruses, In: Mahy, B.W.J., Meulen, V.t. (Eds.), 10th Ed. Topley & Wilson's Microbiology and Microbial Infections—Virology, vol. 2. Hodder Arnold, London, pp. 1304–1321.

Rethwilm, A., 2007. Foamy virus vectors: an awaited alternative to gamma retro-viral and lentiviral vectors. Curr. Gene Ther. 7, 261–271.

Rethwilm, A., Darai, G., Rosen, A., Maurer, B., Flügel, R.M., 1987. Molecular cloning of the genome of human spumaretrovirus. Gene 59, 19–28.

Rice, P., Longden, I., Bleasby, A., 2000. EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet. 16, 276–277.

Roy, J., Rudolph, W., Juretzek, T., Gärtner, K., Bock, M., Herchenröder, O., Lindemann, D., Heinkelein, M., Rethwilm, A., 2003. Feline foamy virus genome and replication strategy. J. Virol. 77, 11324–11331.

Russell, R.A., Wiegand, H.L., Moore, M.D., Schafer, A., McClure, M.O., Cullen, B.R., 2005. Foamy virus Bet proteins function as novel inhibitors of the APOBEC3 family of innate antretroviral defense factors. J. Virol. 79, 8724–8731.

Schliephake, A.W., Rethwilm, A., 1994. Nuclear localization of foamy virus Gag precursor protein. J. Virol. 68, 4946–4954.

Schrago, C.G., Russo, C.A., 2003. Timing the origin of New World monkeys. Mol. Biol. Evol. 20, 1620–1625.

Schweizer, M., Neumann-Haefelin, D., 1995. Phylogenetic analysis of primate foamy viruses by comparison of pol sequences. Virology 207, 577–582.

Stange, A., Mannigel, I., Peters, K., Heinkelein, M., Stanke, N., Cartellieri, M., Gottlinger, H., Rethwilm, A., Zentgraf, H., Lindemann, D., 2005. Characterization of prototype foamy virus gag late assembly domain motifs and their role in particle egress and infectivity. J. Virol. 79, 5466–5476.

Steiper, M.E., Young, N.M., 2006. Primate molecular divergence dates. Mol. Phylogenet. Evol. 41, 384–394.

Strimmer, K., von Haeseler, A., 1996. Quartet puzzling: a quartet maximum likelihood method for reconstructing tree topologies. Mol. Biol. Evol. 13, 964–969.

Swanstrom, R., Wills, J.W., 1997. Synthesis, assembly, and processing of viral proteins. In: Coffin, J.M., Hughes, S.H., Varmus, H.E. (Eds.), Retroviruses. Cold Spring Harbour Labratory Press, Plainview, N.Y., pp. 263–334.

Switzer, W.M., Bhullar, V., Shanmugam, V., Cong, M.E., Parekh, B., Lerche, N.W., Yee, J.L., Ely, J.J., Boneva, R., Chapman, L.E., Folks, T.M., Heneine, W., 2004. Frequent simian foamy virus infection in persons occupationally exposed to nonhuman primates. J. Virol. 78, 2780–2789.

Switzer, W.M., Salemi, M., Shanmugam, V., Gao, F., Cong, M.E., Kuiken, C., Bhullar, V., Beer, B.E., Vallet, D., Gautier-Hion, A., Tooze, Z., Villinger, F., Holmes, E.C., Heneine, W., 2005. Ancient co-speciation of simian foamy viruses and primates. Nature 434, 376–380.

Tobaly-Tapiero, J., Kupiec, J.J., Santillana-Hayat, M., Canivet, M., Peries, J., Emanoil-Ravier, R., 1991. Further characterization of the gapped DNA intermediates of human spumavirus: evidence for a dual initiation of plus-strand DNA synthesis. J. Gen. Virol. 72, 605–608.

Tobaly-Tapiero, J., Bittoun, P., Neves, M., Guillemin, M.C., Lecellier, C.H., Puvion-Dutilleul, F., Gicquel, B., Zientara, S., Giron, M.L., de The, H., Saïb, A., 2000. Isolation and characterization of an equine foamy virus. J. Virol. 74, 4064–4073.

Verschoor, E.J., Langenhuijzen, S., Bontjer, I., Fagrouch, Z., Niphuis, H., Warren, K.S., Eulenberger, K., Heeney, J., 2004. The phylogeography of orangutan foamy viruses supports the theory of ancient repopulation of Sumatra. J. Virol. 78, 12712–12716.

Wilk, T., Geiselhart, V., Frech, M., Fuller, S.D., Flügel, R.M., Löchelt, M., 2001. Specific interaction of a novel foamy virus Env leader protein with the N-terminal Gag domain. J. Virol. 75, 7995–8007.

Wolfe, N.D., Daszak, P., Kilpatrick, A.M., Burke, D.S., 2005. Bushmeat hunting, deforestation, and prediction of zoonoses emergence. Emerg. Infect. Dis. 11, 1822–1827.

Yu, S.F., Edelmann, K., Strong, R.K., Moebes, A., Rethwilm, A., Linial, M.L., 1996. The carboxyl terminus of the human foamy virus Gag protein contains separable nucleic acid binding and nuclear transport domains. J. Virol. 70, 8255–8262.

Yu, S.F., Sullivan, M.D., Linial, M.L., 1999. Evidence that the human foamy virus genome is DNA. J. Virol. 73, 1565–1572.

Zemba, M., Wilk, T., Rutten, T., Wagner, A., Flügel, R.M., Löchelt, M., 1998. The carboxy-terminal p3Gag domain of the human foamy virus Gag precursor is required for efficient virus infectivity. Virology 247, 7–13.