Conserved sequence motifs involving the tat reading frame of Brazilian caprine lentiviruses indicate affiliations to both caprine arthritis–encephalitis virus and visna–maedi virus

Roberto Soares Castro,1, 2 † Timothy Greenland,1, 2 Rômulo Cerqueira Leite,3 Aurora Gouveia,3 Jean-François Mornex1, 2 and Geneviève Cordier1, 2

1 Laboratoire d‘Immunologie et de Biologie Pulmonaire, Hôpital Louis Pradel, Université Claude Bernard, 69394 Lyon Cedex 03, France
2 Laboratoire associé de Recherches sur les Lentivirus chez les petits Ruminants, INRA and Ecole Vétérinaire, Lyon, France
3 Universidade Federal de Minas Gerais – Escola de Veterinária – Depto. Medicina Veterinária Preventiva, C. Postal 567, CEP 30161-970 Belo Horizonte, MG, Brazil

The sequence variation in small ruminant lentiviruses from Brazilian herds of milking goats was sampled in a representative region of the pol gene and in a region including the entire tat open reading frame. Clones were amplified from cDNA derived from virus produced in vitro using primers targetting conserved sequences of the pol gene. Iterative sequencing of clones indicated that animals from two herds in the Minas Gerais area were infected by a caprine arthritis–encephalitis virus (CAEV)-like virus and that individual animals carried variant virus populations. Sequences derived from an infected goat from a herd in Pernambuco showed no nucleic acid variation and were distant from the CAEV-type sequence but marginally closer to ovine visna–maedi virus (VMV) sequences. Sequences amplified from a region including the tat gene, amplified with a common upstream primer within the vif coding region and different downstream primers because of the local divergence between CAEV- and VMV-type sequences, confirmed the affiliation of the Minas Gerais sequences to CAEV and indicated that the Pernambuco isolate was indeed related to VMV, which had not previously been reported to cause natural caprine infection. The overlap between the vif and tat open reading frames clearly distinguished between CAEV-like small ruminant lentiviruses, which shared eight common nucleotides, and the VMV group, where the overlap was reduced to a single base; the final adenine of the vif terminator (TAA) is the initial adenine of the presumptive tat initiator codon. This may be useful for epizoological tracing of the origin of outbreaks.

Introduction

Small ruminant lentiviruses (SRLVs) form a closely related group (Zanoni, 1998) and are widespread in populations of domestic sheep and goats throughout the world (Adams et al., 1984). Goats infected with caprine arthritis–encephalitis virus (CAEV) are at risk of arthritis, principally affecting the carpal joint, mastitis or encephalitis, particularly in kids, whereas sheep infected with visna–maedi virus (VMV) may develop severe pulmonary symptoms, mastitis or, more rarely, encephalitis (reviewed in Narayan et al., 1997). Many infected animals develop no overt symptoms during their whole lifetime, although infection occurs principally in the young animal through virus in milk or colostrum. Viruses from sheep with neurological or pulmonary disease from Iceland (K1514; Sonigo et al., 1985), South Africa (SA-OMVV; Quérat et al., 1990) and Scotland (EV-1; Sargan et al., 1991) have been fully sequenced and show clear similarity although, like other lentiviruses, strains differ considerably, with the greatest
Table 1. Origins and characteristics of the virus isolates

<table>
<thead>
<tr>
<th>Goat</th>
<th>Serological status*</th>
<th>Symptoms</th>
<th>Isolation procedure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mg1-01</td>
<td>+</td>
<td>Arthritis</td>
<td>Explant</td>
</tr>
<tr>
<td>Mg1-02</td>
<td>+</td>
<td>None</td>
<td>Leukocyte co-culture</td>
</tr>
<tr>
<td>Mg2-02</td>
<td>+</td>
<td>Arthritis</td>
<td>Explant</td>
</tr>
<tr>
<td>Mg2-03</td>
<td>+</td>
<td>Arthritis</td>
<td>Explant</td>
</tr>
<tr>
<td>Pe1-01</td>
<td>+</td>
<td>Respiratory distress</td>
<td>Explant</td>
</tr>
</tbody>
</table>

*Seropositivity was evaluated by immunoprecipitation in gels and by ELISA using locally-prepared antigen. The percentages of seropositive animals were 49.7% in herd Mg1, 51.2% in herd Mg2 and 14.7% in herd Pe1.

Conservation in the pol and parts of the gag genes and wide variation over parts of the env gene. The genomic sequence of a virus from an American goat with arthritis (CAEV-Co; Saltarelli et al., 1990), although clearly related, diverges from the VMV sequences in all genes, as does the recent French isolate CA680 (Valas et al., 1997). Viruses from populations of naturally infected sheep in mainland Europe (Zanoni et al., 1992; Leroux et al., 1995) and North America (Chebloune et al., 1996) seem from partial nucleotide sequences of several of their genes to resemble CAEV rather than VMV. This might indicate that an ancestral CAEV-like virus present in sheep has more recently spread into the goat population (Karr et al., 1996; Valas et al., 1997).

Lentiviruses are complex retroviruses carrying genes encoding regulatory functions involved in their replication cycle. One such gene is tat, which has been reported to mediate upregulation of gene transcription from the VMV (Carruth et al., 1994) and CAEV (Saltarelli et al., 1993) long terminal repeat (LTR) promoters, although these are not predicted to form stem–loop structures clearly analogous to the transactivator–response elements of primate or bovine lentiviruses. The basal level of transcription, particularly from the CAEV LTR, is elevated even in the absence of tat, however, and a clone of CAEV lacking functional tat has been shown to be capable of maintaining infection and causing disease in experimentally infected goats (Harmache et al., 1995). The tat open reading frame (ORF) of the VMV isolates differs clearly from that of CAEV-Co, both in sequence (Saltarelli et al., 1990), although there are significant conserved regions, and in its overlap with the preceding vif ORF. The tat region has proved useful in establishing epidemiological relationships in human immuno-deficiency virus (HIV) infections (Lorenzo et al., 1996).

Serological surveys have shown that SRLVs are present in commercial goat flocks in Brazil but uncommon in traditionally herded native animals (Saraiva Neto et al., 1995; R. S. Castro, unpublished data). To determine the overall affiliation of Brazilian caprine lentiviruses, we have recovered virus by limited culture from diseased animals and seropositive herdmates from geographically separated regions and compared sequences amplified by RT–PCR from their tat and pol genes.

Methods

- Animals and virus isolation. Animals and virus isolates are identified by two letters, Mg or Pe, signifying their geographical origin from the Minas Gerais or Pernambuco regions, respectively, followed by two numbers, indicating the herd and the individual animal. Synovial explants from symptomatic animals from two milking herds from the Minas Gerais region (Mg1-01, Mg2-02 and Mg2-03) and one from the Pernambuco region (Pe1-01) of Brazil were cultured in Eagle's minimal essential medium (Gibco BRL) supplemented with 15% foetal calf serum, penicillin (200 U/ml), streptomycin (200 µg/ml) and amphotericin B (2.5 µg/ml). Ficoll-separated blood leukocytes from a seropositive asymptomatic goat (Mg1-02) were co-cultured with synovial cells from a virus-free foetal goat in the above medium containing 10% foetal calf serum (Table 1). Virus-containing supernatants were collected after visible syncytium formation and stored at −80 °C until use.

- RT–PCR, cloning and sequencing. RNA was extracted from 500 µl samples of clarified supernatant by using acid guanidinium thiocyanate–phenol–chloroform (Chomczynski & Sacchi, 1987) precipitated in cold ethanol, vacuum dried, dissolved in 30 µl diethyl-pyrocarbonate-treated water and stored at −80 °C. cDNA obtained by reverse transcription with Moloney virus reverse transcriptase was used as template for PCR amplification of a 475 nucleotide fragment from a conserved region of the pol gene, as previously described (Leroux et al., 1995). tat sequences were amplified by using a common upstream primer (1CV, 5’ AAGTATTGCCATTGTGGAGAGC 3’) and one of two alternative downstream primers (2C, 5’ CCGTGGATTTGTTCCCA-CCC 5’, for the CAEV-like sequences; 3V, 5’ TCCCTTGGTCTTCTTCTCCTCC 3’, for the VMV-like sequences) under the same PCR conditions.

The resulting PCR products were cloned into TA-Vector (Invitrogen) according to the supplier’s instructions. Inserts were released by EcoRI (Boehringer Mannheim) digestion of plasmid DNA from the positive recombinant Escherichia coli colonies. Plasmid DNAs from five positive clones from the pol amplicon from each isolate and from a variable number of tat clones were purified by using the midi plasmid protocol (Qiagen). Released inserts were sequenced by ACT Gene-Euro Sequence Genes Services (Genopole, Evry, France) on an ABI377 sequencer by using the ABI PRISM dye-terminator cycle sequencing ready reaction kit with AmpliTaq DNA polymerase (Perkin-Elmer/Applied Biosystems).
**Sequence analysis.** Sequences were aligned and their relationship was evaluated by using GeneJockey II (Biosoft) or Geneworks 2.0 (Intelligenics) software on a Macintosh computer. For phylogenetic studies, alignments were made by using the CLUSTAL V multiple sequence alignment program (Thompson et al., 1994) and a neighbour-joining tree was constructed with 1000 bootstrap replicates to assess support for each internal node.

**Results**

Syncytia, indicating virus replication, appeared only after prolonged culture (1–3 months) of synovial membrane explants from an arthritic goat from herd Mg1 in the Minas Gerais region (Mg1-01) or in co-cultures of white blood cells from an asymptomatic herdmate (Mg1-02). Similar syncytia appeared after only 1 week in synovial membrane cultures from arthritic goats from the second Minas Gerais herd (Mg2-02 and Mg2-03) or from a goat with isolated respiratory symptoms from Pernambuco (Pe1-01).

Universal primers permitting reliable amplification by RT–PCR from all isolates of SRLVs have not yet been developed but we were able to obtain at least five independent clones of part of the pol gene of virus isolates from four individual goats from three different herds in the two geographical regions studied by the use of degenerate primers from well-conserved regions. The sequence variation and consequences for coding potential are summarized in Table 2.

It is striking that all three isolates from Minas Gerais herds showed considerable variation between replicate clones, whereas the Pernambuco isolate showed none at all (Table 2; homologous diagonal, bold print). Variation between the consensus sequences from the two animals from herd Mg1 was less than that between replicate clones from the individual animals, suggesting a common infectious pool with divergence in the affected host. The consensus sequence of the isolate from herd Mg2 was clearly more distantly related to those from Mg1, but the replicate clones again showed considerable variability. The isolate from herd Pe1 showed wide divergence from the Minas Gerais isolates, with mutations in nearly half of the codons, resulting in 13% amino acid changes. Comparisons between clones from goat Mg1-01 showed seven base changes (six transitions, one transversion) at six positions in the 443 bp excluding the degenerate primer regions, and one clone had a 2 bp deletion leading to a premature termination codon.

Clones from goat Mg1-02 had seven independent base changes (two transversions), one of which created a premature terminator, and those from goat Mg2-03 showed 15 changes at 12 sites (two transversions). All five clones obtained independently from goat Pe1-01 were identical at all positions outside the primers. These values correspond to an accumulation of 31–67 mutations in 10% bases in viruses from the Minas Gerais herds and to < 5 mutations in 10% bases for the BFIF isolates.

**Table 2. Codon variation between pol fragments from Brazilian SRLV isolates**

| Degenerate primer regions were excluded from the comparison. Results are expressed as: number of non-synonymous changes/total changed triplets for the 144 complete codons sequenced. Homologous comparisons (bold type) show variation between the five replicate clones sequenced from each isolate. Comparisons between different isolates used the majority sequence at codons where intra-isolate variation occurred.|
|---|---|---|---|
| **Mg1-01** | **Mg1-02** | **Mg2-03** | **Pe1-01** |
| Mg1-01 | 2*/7 | 0/4 | 6/39 | 18/67 |
| Mg1-02 | – | 5/7 | 7/42 | 19/70 |
| Mg2-03 | – | – | 3/12 | 19/73 |
| Pe1-01 | – | – | – | 0/0 |

*Includes one 2 bp deletion.

**Fig. 1.** Sequence alignments of the vif/tat region of Brazilian SRLVs with homologous sequences from reference SRLV isolates. (a) vif/tat from Mg1-01 (top) compared with CAEV-Co (bottom); (b) vif/tat from Pe1-01 (top) compared with K1514 (bottom). *, nt 125 was A in 3/6 clones; †, nt 226 was G in 1/6 clones; ‡, nt 386 was G in 1/6 clones; §, nt 428 was C in 1/6 clones.
Table 3. Codon variation in the vif and tat ORFs from Brazilian SRLV isolates

Degenerate primer regions were excluded from the comparison. Results are shown as number of non-synonymous codons/total number of changed triplets. The upper right sector (bold type) shows the tat ORF comparison (94 comparable complete codons for isolate Pe1-01; 87 comparable complete codons for other isolates; 86 codons aligned between Pe1-01 and each other isolate). The lower left sector shows the vif ORF comparison (41 comparable complete codons for Pe1-01 and 43 comparable complete codons for other isolates). The homologous comparisons in the table were made between replicate sequences of independently prepared clones (number given in parentheses in the left-hand column) of the different isolates. Comparisons between different isolates involved the majority codon for each position; when two equally frequent variants were present, that most like the comparison sequence was scored. –, Not applicable, as only a single clone was sequenced.

<table>
<thead>
<tr>
<th></th>
<th>Mg1-01</th>
<th>Mg1-02</th>
<th>Mg2-02</th>
<th>Mg2-03</th>
<th>Pe1-01</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mg1-01</td>
<td>–/–</td>
<td>7/15</td>
<td>22/40</td>
<td>23/36</td>
<td>38/61</td>
</tr>
<tr>
<td>Mg1-02</td>
<td>3/5</td>
<td>2/4</td>
<td>0/2</td>
<td>2/6</td>
<td>42/64</td>
</tr>
<tr>
<td>Mg2-02</td>
<td>10/17</td>
<td>10/19</td>
<td>0/1</td>
<td>1/4</td>
<td>40/61</td>
</tr>
<tr>
<td>Mg2-03</td>
<td>9/18</td>
<td>11/20</td>
<td>1/2</td>
<td>0/0</td>
<td>1/2</td>
</tr>
<tr>
<td>Pe1-01</td>
<td>14/24</td>
<td>15/23</td>
<td>20/28</td>
<td>19/28</td>
<td>1/1</td>
</tr>
</tbody>
</table>

Pernambuco isolate. In our laboratory, iterative RT–PCR from an expressed stable plasmid generates approximately 2:5 base changes for $10^4$ bases of sequence analysed (Leroux et al., 1997), suggesting that the majority of the sequence changes in our isolates reflect differences in the virus populations and not laboratory artefacts. The Minas Gerais isolates differed from the CAEV-Co sequence by 37–40 bases, causing only 2–5 amino acid substitutions, but by 57–66 bases from K1514, taken as typical of VMV sequences, causing 19 amino acid substitutions. The Pernambuco isolate was more distant from the CAEV-Co sequence by 37–40 bases, causing only 2–5 amino acid substitutions, but by 57–66 bases from K1514, taken as typical of VMV sequences, causing 19 amino acid substitutions. The Minas Gerais isolates differed from each other (57% in pol) or with the Pernambuco isolate (66% in both tat and vif) (Table 3); however, most substitutions were conservative.

In order to confirm these suggested affinities between the Brazilian isolates and CAEV and VMV prototypes, we next sequenced a region encompassing the tat ORF, which shows characteristic differences between the VMV-like viruses and CAEV-Co. An isolate from an additional animal from the Pernambuco herd was included to evaluate within-herd variation further and between one and six clones were sequenced for each isolate. The common upstream primer 1CV covers a conserved region of the vif ORF and when used with the downstream primer 2C on cDNA from CAEV-related isolates generated a 397 bp fragment ending 7 bp after the terminator of the tat ORF, which was sequenced entirely (Fig. 1a). In conjunction with the downstream primer 3V, amplification from cDNA from Pe1-01 virus produced a 557 bp fragment, of which 449 bp were sequenced to include the whole of the tat ORF (Fig. 1b). Variant positions in the multiply cloned and sequenced isolates from the Minas Gerais herds, expressed as number per $10^4$ bp, ranged from 25 in Mg2-02 to 66 in Mg1-02, similar to the variation observed for the pol fragment (see above). The Pernambuco isolates showed rather less variation, 15 sites per $10^4$ bp, which was nevertheless greater than that observed in the pol region. The ratio of non-synonymous to total replacements by codon was high in the isolates from Minas Gerais herds when compared with each other ($57\%$ in tat, $54\%$ in vif) or with the Pernambuco isolate ($66\%$ in both tat and vif) (Table 3); however, most substitutions were conservative.

An alignment of the amplified fragment, which includes the last 43 complete codons of the vif gene and the 87 codons comprising the tat ORF from virus Mg1-01 (representative of the Minas Gerais sequences), with the homologous region from CAEV-Co is shown in Fig. 1(a). A similar alignment of 41 vif codons and the 94 codons comprising tat from virus Pe1-01 and K1514 (representative of the VMV sequences) is shown in Fig. 1(b). Sequence Mg1-01 was clearly homologous to the CAEV-Co sequence both in the length of the coding regions, with the terminator for tat at positions 387–389, and the overlap between vif and tat, which contained eight nucleotides. Sequence Pe1-01 was closer to the VMV group, with a single nucleotide in the vif/tat overlap and a tat terminator at positions 410–412. The deduced amino acid sequences show variation, mostly involving conservative replacements, both in the C-terminal portion of Vif and throughout much of Tat.
alignment of the consensus sequences from the different Minas Gerais isolates (Fig. 2) shows that the nucleotide changes, both between the different isolates and within clones from each isolate (shown in bold type), were scattered throughout the sequence with no apparent concentration in any ‘hyper-variable’ region. The number of clones sequenced for each isolate is indicated in parentheses. Underlined sequences at the beginning and end of the alignments were used as PCR primers. Variant bases (found in single clones unless indicated) are shown in italic type. The sequence at positions 132–134.

Discussion

Nucleotide sequencing of selected fragments generated by RT–PCR from caprine virus isolates confirmed serological indications that Brazilian commercial milking herds are infected by SRLVs. The viruses recovered showed significant clonal variation, even in isolates from Mg1-02 despite a phase of virus amplification by co-culture, suggesting the existence of lentivirus quasispecies in these animals, as reported for naturally infected French sheep (Leroux et al., 1997). The very small amount of variation, particularly in the pol region, between independent clones derived without co-culture from a goat from a herd in Pernambuco suggests a naturally less-variable virus. This herd of animals is descended from goats imported from Great Britain some 20 years ago. Interestingly, the virus isolated from Pe1-01 showed clear affinity to the classical ovine VMV rather than to CAEV-Co, the prototype caprine strain. VMVs K1514 and K1772, which are related by serological and the present report of a VMV-related sequence in a goat appears to be the first observation of such an infection occurring naturally. Viruses isolated from seropositive animals from two herds in the Minas Gerais region were closely related to CAEV-Co. Virus from pairs of animals from the same herd showed little more variation between their sequences than did replicate clones from the same animal, suggesting that there is efficient mixing of virus populations among animals in constant close contact. More distinct differences were seen between viruses from the two herds. Even greater variations may exist in the field, because virus isolation was based on the formation of syncytia and infections detectable only by specific staining for virus antigen were not investigated.

The relationships between sequences from different goats was the same whether the pol or tat regions were considered, but the tat regions, together with the 3’ end of the vif gene, showed a higher rate of nucleotide and, especially, amino acid variation than did the pol fragments (Fig. 3; note scale difference). No recombination between the pol and tat regions, which are separated by over 2500 nucleotides, was observed in the present series, but such an event remains a possibility if animals from one herd with a distinct virus population are introduced to another infected herd. It is not known whether VMV-like and CAEV-like viruses can recombine, but the presence of CAEV-related sequences in sheep and the present report of a VMV-related sequence in a goat provide means for the experimental investigation of this possibility.

Comparison between different sequences of classical ovine VMV has suggested that the tat ORF is surprisingly tolerant of variation, despite its presumed importance for virus replication. The tat gene in HIV-1, the vital function of which is much more clearly understood, also shows much variability (Lorenzo et al.,
Fig. 3. Phylogenetic trees of the relationships between nucleotide sequences of the pol (a) and vif/tat (b) regions of Brazilian SRLVs and prototype isolates. Trees were constructed by the neighbour-joining method from the CLUSTAL V package on sequences from four (pol) or five (vif/tat) Brazilian SRLV isolates represented by following symbols: ▽, Mg2-02; ■, Mg2-03; ▲, Mg1-01; ●, Mg1-02; ○, Pe1-01. All cloned sequences available are included. Numbers at branch-points represent the percentage of 1000 bootstrap repetitions generating the illustrated configuration; only values > 85% are shown. Sequence K1514 was defined as the outgroup. Note the different scales (inferred changes per nucleotide position) in (a) and (b).

1996). Our sequences related to CAEV-Co show that the caprine virus tat region also shows high variability, with an accumulation of non-synonymous codon changes. The VMV and CAEV sequences do, however, preserve a clearly distinct difference in their relation with the preceding vif ORF. All CAEV-like viruses had an eight nucleotide overlap between the two ORFs, while the VMV group showed a minimum overlap, with the last adenine of the TAA terminator of vif being the initial adenine of the tat ATG initiator. The VMV tat genes were also regularly longer than the CAEV homologues, encoding proteins with a seven amino acid C-terminal extension. These differences might be useful for distinguishing between the two virus types. The real part played by the Tat peptide in the life history of the virus and the infected cell is still open to some speculation. A CAEV-Co-derived clone, artificially deleted in the tat gene, has been shown to be capable of infecting goats and causing disease symptoms (Harmache et al., 1995) and a partially sequenced VMV isolate from a symptomatic ewe proved to have a nine nucleotide, non-homologous stretch introducing two in-frame terminators into the middle of the tat ORF (Campbell et al., 1993). Most sequenced isolates do, however, preserve the integrity of this variable ORF and in the present series no clones showed premature termination of tat, although two clones from different goats contained in-frame terminators in the pol gene, presumably producing defective virus. Comparison of the VMV-type and CAEV-type tat sequences showed two short regions of high amino-acid conservation; a perfectly conserved octapeptide LQRWLAML near the middle of the peptide and the sequence CGCRXCNPGWG(S/T) towards the C terminus in the CAEV-type sequences. These sequences appear to have no very strong homology with other lentivirus Tat peptides or with other known gene products, but may provide a basis for further functional investigations.

This work was supported in part by the Agence Nationale de Recherches sur le SIDA and the Hospices Civils de Lyon. R.S.C. was the recipient of a fellowship from Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) do Ministério da Educação e do Desporto do Brasil. We express our thanks to Joelle Chastang for her excellent technical assistance and to Caroline Leroux for assistance in producing the phylogenetic trees and for helpful discussion.
References


Received 24 November 1998; Accepted 20 March 1999