JOURNAL OF VIROLOGY, June 1996, p. 3346–3354 Vol. 70, No. 6
0022-538X/96/$04.00+0

Genomic Quasispecies Associated with the Initiation of Infection and Disease in Ponies Experimentally Infected with Equine Infectious Anemia Virus

DREW L. LICHTENSTEIN,[1] CHARLES J. ISSEL,[2] AND RONALD C. MONTELARO[1]*

Department of Molecular Genetics and Biochemistry, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania 15261,[1] and Department of Veterinary Science, Gluck Equine Research Center, University of Kentucky, Lexington, Kentucky 40546[2]

Received 8 September 1995/Accepted 20 February 1996

Equine infectious anemia virus (EIAV) provides a uniquely dynamic system in which to study the mechanism and role of genomic variation in lentiviral persistence and pathogenesis. We have used a Shetland pony model of infection to investigate the association of specific long terminal repeat (LTR) and *env* gene genomic sequences with the initiation of infection and the onset of disease. We analyzed viral RNA isolated from a pathogenic stock of virus ($EIAV_{PV}$) and from plasma taken during the first disease episode from two ponies infected with $EIAV_{PV}$. Overall sequence variation within gp90 was low in $EIAV_{PV}$ and only slightly higher in plasma virus samples isolated from ponies during the first disease episode. However, a high proportion of mutations were localized to the principal neutralizing domain in $EIAV_{PV}$ and to the principal neutralizing domain and the gp90 hypervariable region in the two pony-derived samples. The rate of fixation of mutations was analyzed and determined to be approximately 4×10^{-2} mutations per site per year. Sequence diversity within the U3 region of the LTR was extremely low, which suggested that the previously reported hypervariability of this region may be a consequence of selection for replication of EIAV in different host cells. The predominant $EIAV_{PV}$ *env* and LTR sequences were used to construct chimeric viruses so that the contribution of these sequences to viral pathogenicity could be examined. The chimeras replicated in cultured equine monocytes to the same extent as the parental nonpathogenic virus and did not cause disease in Shetland ponies by 120 days postinfection, suggesting that the EIAV genomic determinants of pathogenesis are complex.

RNA viruses in general, and lentiviruses in particular, undergo extensive genetic variation as a result of error-prone replication and recombination such that they are considered to exist as quasispecies (populations of closely related genomes). Infection of horses with equine infectious anemia virus (EIAV), a member of the lentivirus subfamily of retroviruses, has served as a model for immune-mediated antigenic variation. It has been demonstrated that serum from an EIAV-infected animal can neutralize virus isolates recovered from prior but not subsequent disease episodes and that each disease episode is associated with a novel predominant antigenic variant (16, 25). Biochemical (25, 33, 36, 38) and immunological (14, 32) studies of the predominant virus isolate recovered during disease episodes indicate that the two glycoproteins, gp90 (SU protein) and gp45 (TM protein), vary extensively. In addition, the predominant viruses isolated from sequential febrile episodes in a single pony were shown to be different at the genetic level by RNase mapping experiments (29, 33) and by sequence analysis of the envelope (*env*) gene (30). Sequence analysis showed that much of the genetic variation in gp90 was clustered in a single area termed the variable region (30). Ball et al. (3) localized a principal neutralizing domain (PND) to the gp90 variable region, which accounts for the variable reactivity of different virus isolates to PND-specific, neutralizing monoclonal antibodies (14). Additional studies on EIAV sequence diversity have yielded similar results and have identified a second highly variable segment of the genome within the U3 region of the long terminal repeat (LTR) (1, 6, 30, 31, 34).

Rapid genetic variation during infection with EIAV is supported by several lines of evidence. Experiments in which an attenuated strain of EIAV was serially passaged in Shetland ponies resulted in a rapid evolution in the pathogenic properties of the virus (27). Furthermore, when antigenic variation in gp90 and gp45 was probed by immunoreactivity to monoclonal antibodies (14) or by peptide mapping (33, 38), it took as few as 14 to 28 days for variation to be seen in sequential disease episodes from a single animal. Finally, sequence analysis early after infection of a horse with the Wyoming strain of EIAV showed that portions of the *env* gene exhibited significant heterogeneity during the first disease episode (1).

While lentivirus infections usually result in a gradual development of disease over the course of months to years, the rapid onset and clear demarcation of disease in EIAV-infected animals represent a unique opportunity to examine the development of lentivirus disease and persistence. We have used a Shetland pony model of infection in which experimental infection of animals with a pathogenic stock of virus ($EIAV_{PV}$) reproducibly results in the three classical stages of disease: acute, chronic, and inapparent (24). Following infection of ponies with $EIAV_{PV}$, the acute stage of disease (defined as the first clinical episode) typically occurs by 3 weeks postinfection. The chronic stage is characterized by six to eight irregularly spaced, recurrent cycles of disease that are associated with a high level of plasma viremia and symptoms that include fever, anemia, thrombocytopenia, depression, and anorexia (24). Approximately a year after infection, the ponies enter the inapparent stage, in which they are clinically quiescent, but remain infected for life.

* Corresponding author. Mailing address: Department of Molecular Genetics and Biochemistry, University of Pittsburgh School of Medicine, Pittsburgh, PA 15261. Phone: (412) 648-8869. Fax: (412) 383-8859.

$EIAV_{PV}$ was derived from the tissue culture-adapted strain of EIAV (19) by three serial back-passages in Shetland ponies followed by biological cloning of neutralization escape mutants that were selected in the presence of autologous neutralizing serum (27, 37). Although the $EIAV_{PV}$ stock has been widely used in experimental infections, the genetic heterogeneity of this biological clone has never been determined. Moreover, there is no information on the $EIAV_{PV}$ quasispecies that establish the infection in ponies and cause the initial acute episode of disease. Because of the rapid variation that occurs during EIAV infection, it was anticipated that genetic characterization of $EIAV_{PV}$ during the initial stages of infection and disease could direct the production of a genetically defined pathogenic molecular clone of EIAV, a goal that has thus far eluded researchers who have described only avirulent infectious molecular clones. To examine the diversity present in the $EIAV_{PV}$ stock and to investigate which viral sequences are associated with the first disease episode in Shetland ponies infected with $EIAV_{PV}$, the variable region of gp90 and the U3 region of the LTR (U3-LTR) were sequenced. In addition, the predominant LTR and gp90 variable region sequences were used to create chimeric recombinant viruses in an attempt to define segments of the EIAV genome that may be involved in pathogenesis.

MATERIALS AND METHODS

Cells. Equine mononuclear cells were enriched from whole blood as described previously (18) and were seeded into 24-well plates at 4×10^6 cells per well in modified Eagle's medium alpha medium (GibcoBRL, Grand Island, N.Y.) supplemented with 10% heat-inactivated horse serum (Sigma, St. Louis, Mo.), penicillin G (100 U/ml), streptomycin (100 μg/ml), and glutamine (2 mM). Monocytes were adhered to plastic for 4 h at 37°C and washed twice with Hanks' balanced salt solution (GibcoBRL) to remove nonadherent cells. The medium containing the nonadherent cells was clarified by centrifugation at $3,500 \times g$ and then placed back on the monocytes. Fetal equine kidney (FEK) cells were grown at 37°C in minimum essential medium (GibcoBRL) supplemented with 10% heat-inactivated fetal bovine serum (Atlanta Biologicals, Atlanta, Ga.), nonessential amino acids (GibcoBRL), and penicillin, streptomycin, and glutamine as described above.

Molecular cloning. The infectious, nonpathogenic molecular clone $pEIAV_{19\text{-}2}$ (31; GenBank accession number U01866) was used as a backbone for the generation of chimeric proviral clones by restriction fragment exchanges. Since the predominant LTR sequence of $EIAV_{PV}$ was identical to the $EIAV_{65}$ LTR (31), the latter was used for LTR restriction fragment exchanges. The 3′ end of the gp45 gene and the entire 3′ LTR of $EIAV_{19\text{-}2}$ (contained on a *Bst*XI-*Eco*RI restriction fragment) was replaced with the corresponding fragment of $pEIAV_{65}$ to generate the proviral clone $pEIAV_{PV100}$. A second chimeric proviral clone that contained both the predominant gp90 and LTR sequences of $EIAV_{PV}$ was created. A 736-bp *Hin*dIII-*Hin*dIII restriction fragment from clone $EIAV_{PV}$-A2, which was identical to the $EIAV_{PV}$ consensus sequence and encoded all but the first 24 amino acids of the region that we sequenced, was used to replace the corresponding region of $pEIAV_{PV100}$, such that the proviral clone $pEIAV_{PV101}$ was created.

Viruses and infections. Virus stocks were prepared by harvesting the medium from calcium phosphate-transfected FEK cells (39). The stocks were assayed by using a micro-reverse transcriptase assay (18) and titered for infectivity with an infectious center assay based on an enzyme-linked immunosorbent assay (11). Virus was absorbed to equine monocytes for 4 h at 37°C, after which the virus inoculum was washed off and fresh medium was added to the cells. Medium was replaced daily, and the clarified supernatants were assayed for reverse transcriptase activity as described previously (18). Seronegative Shetland ponies were inoculated intravenously with $10^{7.0}$ infectious center doses of $EIAV_{PV100}$ or $EIAV_{PV101}$. Rectal temperatures were monitored daily, and serum and plasma samples were taken periodically.

Derivation of $EIAV_{PV}$ and plasma samples. Serial back-passage of the cell-adapted strain of Wyoming in Shetland ponies was described previously (27). Plasma from a second back-passage pony (pony 82) was passaged a third time in pony F135 (37). To generate a neutralization escape mutant, virus isolated from pony F135 at 16 days postinfection (during the first disease episode) was serially passaged 13 times in FEK cells in the presence of autologous neutralizing serum isolated 203 days postinfection (37). The resulting neutralization-resistant virus was biologically cloned twice in FEK cells by endpoint dilution and amplified one time in FEK cells to generate the virus stock termed $EIAV_{PV}$. The titer of $EIAV_{PV}$ is $10^{6.5}$ horse infectious doses. Two seronegative Shetland ponies were intravenously infected with 300 horse infectious doses of $EIAV_{PV}$. Plasma taken from each pony during the first disease episode (day 19 for pony 06 and day 33 for pony 11) was used in this study. The titer of pony 06 plasma taken on day 19 is $10^{5.0}$ 50% tissue culture infectious doses. The titer of pony 11 plasma taken on day 33 is unavailable. However, a quantitative reverse transcription (RT)-PCR assay (18) showed that the concentration of viral RNA in the plasma sample was approximately $3 \times 10^{7.0}$ molecules per ml of plasma.

RNA purification and RT-PCR. Viral RNA was purified and RT-PCR was performed as described previously (18). Briefly, concentrated virus, pelleted from plasma or tissue culture supernatant at $120,000 \times g$, was extracted with the Trizol reagent (GibcoBRL) to purify the viral RNA. Reverse transcription of 2 μl of viral RNA (equivalent to 200 μl of $EIAV_{PV}$ stock or pony plasma) was performed with the Superscript Preamplification System (GibcoBRL) as specified by the manufacturer except that first-strand synthesis was initiated with an EIAV-specific primer (0.4 μM). cDNA synthesis for amplification of the gp90 variable region or the U3-LTR was performed with the Var2 primer (5′ GAG CAGTTATATTGGTTAAAGCTTTGG 3′) or the LTR6 primer (5′ AGGC CTTTTCAGCCCAGCAGA 3′), respectively. PCRs were carried out by using a hot-start procedure with a mixture that contained 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 0.001% gelatin, 0.05 mM each deoxynucleoside triphosphate, 0.3 μM each primer, 1 U of AmpliTaq DNA polymerase (Perkin-Elmer, Norwalk, Conn.), and 2 μl of cDNA in a final volume of 25 or 100 μl. Primers used for amplification of the variable region of gp90 were Var1 (5′ GTTCCTTCCCGGGGTGTAGACC 3′) and Var2. Primers used for amplification of the U3-LTR were LTR5 (5′ CCCCTCATAAAAACCCCACA 3′) and LTR6. PCR was carried out under the following cycling conditions: 1 min at 94°C, 1 min at 55°C, and 1 min at 72°C for 35 cycles; 10 min at 72°C for one cycle; and hold at 4°C.

Cloning of RT-PCR products. Except as noted below, a single RT reaction provided cDNA for multiple independent PCRs. PCR products were pooled prior to purification (Wizard PCR Preps DNA Purification System; Promega, Madison, Wis.) and then T-A cloned into pGEM5Zf(+) (Promega). The $EIAV_{PV}$ *env* clones were generated in two separate experiments, each of which consisted of 12 independent PCRs from a single RT reaction. Similar results were obtained after analysis of 8 or 12 clones from the first or second set of reactions, respectively. Five independent PCRs were used to generate the $EIAV_{PV}$ LTR clones. The pony 06 *env* clones were generated from 10 independent PCRs. The pony 11 *env* clones, the pony 06 LTR clones, and 12 of the pony 11 LTR clones were derived from a single RT-PCR. An additional 8 pony 11 LTR clones were generated from 8 independent PCRs (that were not pooled) from a single RT reaction. Analysis of the latter pony 11 clones confirmed the results obtained with the former group of pony 11 clones. White colonies were screened for the proper size insert by restriction enzyme analysis without prior screening or following hybridization of the appropriate ^{32}P-labeled probe to a colony blot (39).

Sequencing of RT-PCR clones. For sequencing, double-stranded plasmid DNA and 1.0 μl of the appropriate primer (0.5 pmol/μl) were denatured together in a total volume of 10.0 μl in the presence of dimethyl sulfoxide (10%) and NaOH (0.1 M) for 10 min at 68°C. Samples were neutralized with 0.4 volume of neutralizing buffer [0.28 M *N*-tris(hydroxymethyl)methyl-2-aminoethanesulfonic acid (TES), 0.08 M $MgCl_2$, 0.2 M NaCl (pH 1.1)] and then kept at room temperature for 10 min. Samples were sequenced by the chain termination method (40, 41) with a Sequenase version 2.0 sequencing kit (United States Biochemical, Cleveland, Ohio) and [α-*thio*-^{35}S]dATP (1,000 to 1,500 Ci/mmol; DuPont NEN, Boston, Mass.). Both strands of each gp90 insert were sequenced with the following primers: M13 (−40) forward primer (United States Biochemical), M13 (−24) reverse primer (New England Biolabs, Beverly, Mass.), Var3 (5′ GTCTTCTTGCACAGATAG 3′), Var4 (5′ GTGTGCCTGTCCTATAAC 3′), Var5 (5′ GGGATACATCCAATCAGGC 3′), Var6 (5′ ATCTTCTAAAAC CCCAAG 3′), Var7 (5′ CACTAACATAACTTCCTGC 3′), and Var8 (5′ TTCAGGTACTAATATAGT 3′). For the LTR clones, the M13 forward and reverse primers were used such that both strands of the insert were sequenced. All PCR and sequencing primers were designed according to the published sequence (35). The Genetics Computer Group (Madison, Wis.) package of sequence analysis software was used for sequence analysis (10).

To estimate the *Taq* DNA polymerase error rate, 10 fg of a previously sequenced gp90 variable region clone was reamplified by PCR, and the products were cloned and sequenced as described above. Six clones were sequenced, and the error rate was determined to be 0.06% (three substitutions per 4,842 bp sequenced).

Nucleotide sequence accession numbers. The sequences analyzed were submitted to GenBank and have been assigned accession numbers U35178 through U35224 and U43948 through U43950.

RESULTS

To examine EIAV genomic quasispecies that are associated with the initiation of infection and disease, we compared viral RNA sequences present in a pathogenic stock of virus termed $EIAV_{PV}$ with the sequences circulating in the plasma during the initial disease episode in two ponies infected with $EIAV_{PV}$.

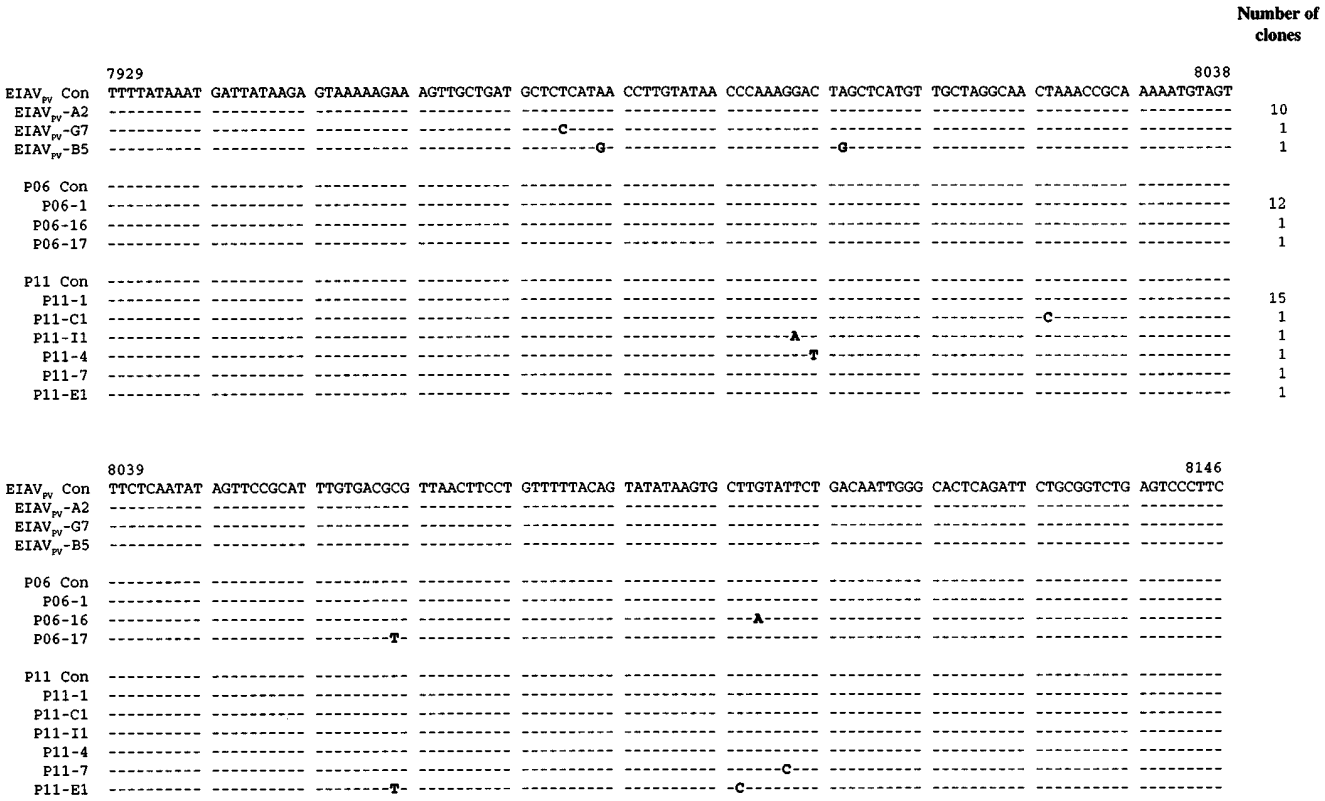


FIG. 1. U3-LTR sequences of EIAV$_{PV}$, pony 06, and pony 11. Sequences were aligned with the program Pileup from the Genetics Computer Group sequence analysis package (10). Dashes indicate a nucleotide identical to the EIAV$_{PV}$ consensus sequence (EIAV$_{PV}$ Con). For each sample, only unique clones are shown, and the number of clones with a given sequence is shown to the right. The nucleotide numbering refers to the position of each sequence with respect to the EIAV$_{Pr}$ proviral sequence.

Two portions of the viral genome reported to be variable were sequenced: (i) a 219-bp segment of the U3-LTR (6) and (ii) an 807-bp segment of the *env* gene encoding the variable region of the SU glycoprotein gp90 (30). The portion of the LTR that was sequenced contained nearly the entire U3 region, including the hypervariable region (6). The segment of the *env* gene encoding gp90 that was sequenced contained the variable region flanked by portions of two conserved regions. Located within the gp90 variable region are the hypervariable region and the PND, the latter of which is composed of two adjacent neutralizing epitopes (D_{NT} and E_{NT}) that were defined by monoclonal antibodies (see Fig. 3 and references 3 and 14).

The EIAV$_{PV}$ stock was derived as described in Materials and Methods. Pony 06 and pony 11 were each infected with EIAV$_{PV}$, and viral RNA was isolated from plasma during the first clinical episode in each pony (day 19 for pony 06 and day 33 for pony 11). We sequenced viral genomic RNA isolated directly from plasma, as opposed to sequencing proviral DNA, since previous results with EIAV (38) and recent results with human immunodeficiency virus (HIV) (13, 44) suggest that lentivirus infections are quite dynamic. Thus, examining the virus present in plasma is the best way to capture the currently replicating viral species. Genomic viral RNA was subjected to RT-PCR, then the amplified products were cloned, and multiple independent clones were sequenced.

Sequence analysis of the U3 region of the LTR. The sequences of the LTR clones from the EIAV$_{PV}$ stock and the pony-derived samples (12 to 20 clones from each sample) were nearly all identical (Fig. 1). The consensus sequences from all three samples were identical, and clones which varied from the consensus sequence did so by only one or two nucleotides. The level of variation observed for the LTR sequences (Table 1) was only slightly greater than the calculated *Taq* DNA polymerase error rate (see Materials and Methods).

Sequence comparison of the previously identified hypervariable region of the U3-LTR was used to determine the relationship of the EIAV$_{PV}$, pony 06, and pony 11 LTRs to those of pathogenic and nonpathogenic strains of EIAV (Fig. 2). The hypervariable region of the EIAV$_{PV}$ LTR and that of the pony-derived LTR had significant homology with two virulent strains of EIAV (the consensus Wyoming LTR [Fig. 2] and the Th-1 LTR [data not shown]) and were more divergent from the LTRs of the avirulent isolates EIAV$_{1369}$ and EIAV$_{19\text{-}2}$. Inter-

TABLE 1. Mean percentage variation in gp90 and the U3-LTR

Sample	Mean % variation		
	gp90		LTR, nucleotide
	Nucleotide	Amino acid	
PV stock	0.20 (33/16,140[a])	0.48 (26/5,380[b])	0.11 (3/2,628[a])
Pony 06	0.40 (36/9,684)	0.80 (27/3,228)	0.08 (2/2,628)
Pony 11	0.30 (29/9,684)	0.77 (25/3,228)	0.14 (6/4,380)

[a] Total number of nucleotide substitutions, with respect to each sample's consensus sequence, divided by the total number of nucleotides sequenced.

[b] Total number of deduced amino acid substitutions, with respect to each sample's consensus sequence, divided by the total number of amino acids.

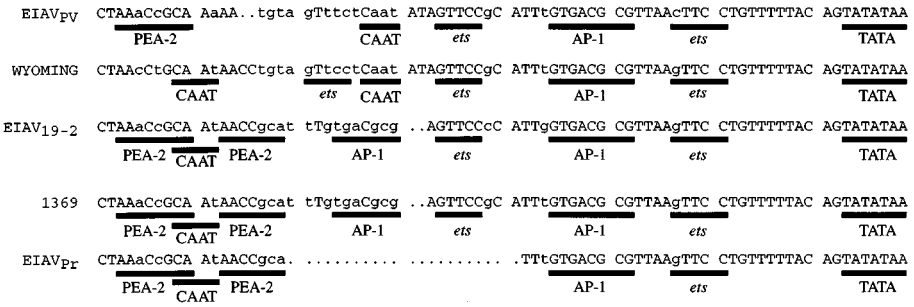


FIG. 2. Comparison of the U3-LTR hypervariable region of various virus isolates. A portion of the $EIAV_{PV}$ LTR consensus sequence was aligned with the corresponding region of other known LTR sequences, using the program Pileup (10). The pony-derived consensus sequences are not shown since they were identical to that of $EIAV_{PV}$. Potential *cis*-acting sequences (PEA-2, CAAT, *ets*, AP-1, and TATA) are underlined. A lowercase letter indicates a nucleotide position for which at least one of the sequences differs from the others. Gaps (indicated by dots) were introduced to align the sequences.

estingly, the $EIAV_{PV}$ and pony-derived LTRs lacked a second CAAT box motif and a third core *ets* binding motif that are present in the Wyoming and Th-1 LTRs (6, 21, 34). It has been suggested that two CAAT boxes and core *ets* binding motifs may be important in viral pathogenesis and replication in macrophages, respectively (6, 21).

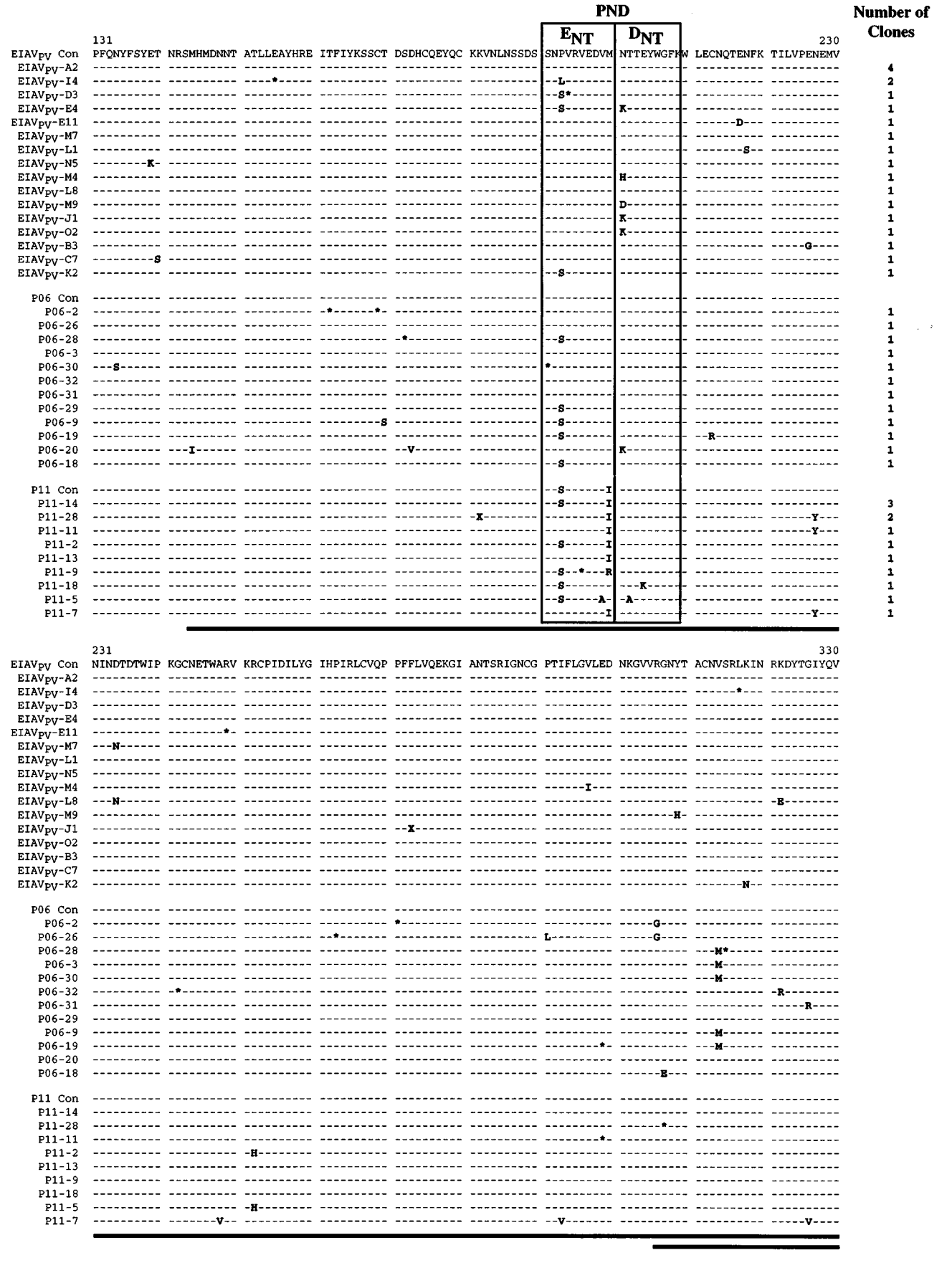
Sequence analysis of the SU protein gp90. Limited sequence variation was observed within the portion of the *env* gene that was amplified from the $EIAV_{PV}$ stock (20 clones) and from the plasma virus samples isolated from pony 06 (12 clones) and pony 11 (12 clones) (Fig. 3). Sequence variation among the latter two samples was slightly higher than among the $EIAV_{PV}$ clones, indicating that the $EIAV_{PV}$ stock underwent variation in vivo (Table 1). The level of variation observed for all three samples was significantly higher (between three- and sevenfold) than the calculated *Taq* error rate (see Materials and Methods), indicating that the majority of nucleotide changes were not the result of the *Taq* DNA polymerase. Among all three samples, the large majority of sequence changes were nucleotide substitutions. Of the 44 *env* clones examined, only three contained a 1-bp deletion resulting in a defective *env* gene (one clone from $EIAV_{PV}$ and two clones from pony 11). In addition, one clone from pony 06 contained an in-frame 3-bp deletion. There was no evidence for G-to-A hypermutation (data not shown) as described previously for EIAV and other lentiviruses (5, 9, 15, 34, 42, 43).

Nonsynonymous changes within gp90 were disproportionately located in the PND and/or the hypervariable region (HVR). Although the PND comprised 7% of the total length of the region sequenced, this segment accounted for 38% (10 of 26), 22% (6 of 27), and 48% (12 of 25) of the nonsynonymous changes that were found in $EIAV_{PV}$, pony 06, and pony 11, respectively (Fig. 3). In addition, 41% (11 of 27) of the nonsynonymous changes found in pony 06 were located in the HVR, a segment that constituted only 13% of the total length of the region sequenced (Fig. 3). These data suggested that the pattern of mutations was not random. The position and type of *Taq*-induced errors (one nonsynonymous change in each of the two conserved regions and one synonymous change in the PND) in the control experiment (see Materials and Methods) make it unlikely that the observed pattern of mutations in our samples was due to the presence of a hot spot for *Taq*-induced misincorporations.

The rate of fixation of mutations was determined from the formula $R = D/2T$, where R is the number of nucleotide substitutions per site per year, D is the mean pairwise nucleotide distance between $EIAV_{PV}$ and the pony-derived samples (data not shown), and T is the length of time after infection that the samples were isolated (pony 06 = 0.05 year and pony 11 = 0.09 year). The mutation rates were calculated to be 5.6×10^{-2} and 3.2×10^{-2} mutations per site per year for pony 06 and pony 11, respectively.

The $EIAV_{PV}$ and pony 06 consensus sequences were identical to each other, while that of pony 11 differed from them by only two amino acids, both of which were located within the PND (Fig. 4). Comparison of all three consensus sequences with other known gp90 amino acid sequences revealed significant differences, particularly within the PND and the HVR (Fig. 4). All sequences examined were only distantly related to the Wyoming sequence, diverging significantly in the PND and HVR. Our gp90 sequences showed greater homology to the prototype ($EIAV_{Pr}$) strain of virus, from which $EIAV_{PV}$ was originally derived (see Materials and Methods), although sequence divergence was still significant in the PND and HVR. The $EIAV_{PV}$ and pony-derived consensus sequences differed from the $EIAV_{Pr}$ and $EIAV_{19\text{-}2}$ sequences by seven or eight amino acids (approximately 3%). In addition, all 44 *env* clones contained two silent nucleotide changes with respect to the $EIAV_{Pr}$ and $EIAV_{19\text{-}2}$ sequences (data not shown). Five of the amino acid changes, four within the HVR and one within the PND, resulted in additional potential N-glycosylation sites within our consensus sequences compared with the $EIAV_{Pr}$ and $EIAV_{19\text{-}2}$ sequences (Fig. 4; note that 6 of the 44 clones contained an additional mutation which resulted in loss of the additional potential N-glycosylation site within the PND [see Fig. 3]). These results are not surprising since differences in the tryptic glycopeptide maps among various EIAV isolates have been noted previously (33, 38), and sequence analysis of EIAV and other lentiviruses has yielded similar changes in potential N-glycosylation sites (4, 5, 22, 28, 30). The addition of a potential N-glycosylation site within the PND of $EIAV_{PV}$ could account, in part, for its altered neutralization properties as described previously (37).

Infection with chimeric viruses. Chimeric viruses were constructed to examine the pathogenic potential of the predominant LTR and gp90 variable region sequences present in the $EIAV_{PV}$ stock. The predominant sequences present in the pathogenic $EIAV_{PV}$ stock might be associated with the generation of disease, since these sequences were very similar to those present during the first disease episode in two $EIAV_{PV}$-infected ponies and differed significantly from the corresponding sequences of the avirulent strains $EIAV_{Pr}$ and $EIAV_{19\text{-}2}$. Two chimeric proviral clones were constructed such that the only known difference between the parental and recombinant viruses was the U3-LTR (clone $pEIAV_{PV100}$) or this region plus the gp90 variable region (clone $pEIAV_{PV101}$). The aviru-


PND
E_{NT}
D_{NT}
Number of Clones
131
230
EIAVPV Con PFQNYFSYET NRSMHMDNNT ATLLEAYHRE ITFIYKSSCT DSDHCQEYQC KKVNLNSSDS SNPVRVEDVM NTTEYWGFKW LECNQTENFK TILVPENEMV
EIAVPV-A2 4
EIAVPV-I4 2
EIAVPV-D3 1
EIAVPV-E4 1
EIAVPV-E11 1
EIAVPV-M7 1
EIAVPV-L1 1
EIAVPV-N5 1
EIAVPV-M4 1
EIAVPV-L8 1
EIAVPV-M9 1
EIAVPV-J1 1
EIAVPV-O2 1
EIAVPV-B3 1
EIAVPV-C7 1
EIAVPV-K2 1
P06 Con
P06-2 1
P06-26 1
P06-28 1
P06-3 1
P06-30 1
P06-32 1
P06-31 1
P06-29 1
P06-9 1
P06-19 1
P06-20 1
P06-18 1
P11 Con
P11-14 3
P11-28 2
P11-11 1
P11-2 1
P11-13 1
P11-9 1
P11-18 1
P11-5 1
P11-7 1
231
330
EIAVPV Con NINDTDTWIP KGCNETWARV KRCPIDILYG IHPIRLCVQP PFFLVQEKGI ANTSRIGNCG PTIFLGVLED NKGVVRGNYT ACNVSRLKIN RKDYTGIYQV
EIAVPV-A2
EIAVPV-I4
EIAVPV-D3
EIAVPV-E4
EIAVPV-E11
EIAVPV-M7
EIAVPV-L1
EIAVPV-N5
EIAVPV-M4
EIAVPV-L8
EIAVPV-M9
EIAVPV-J1
EIAVPV-O2
EIAVPV-B3
EIAVPV-C7
EIAVPV-K2
P06 Con
P06-2
P06-26
P06-28
P06-3
P06-30
P06-32
P06-31
P06-29
P06-9
P06-19
P06-20
P06-18
P11 Con
P11-14
P11-28
P11-11
P11-2
P11-13
P11-9
P11-18
P11-5
P11-7

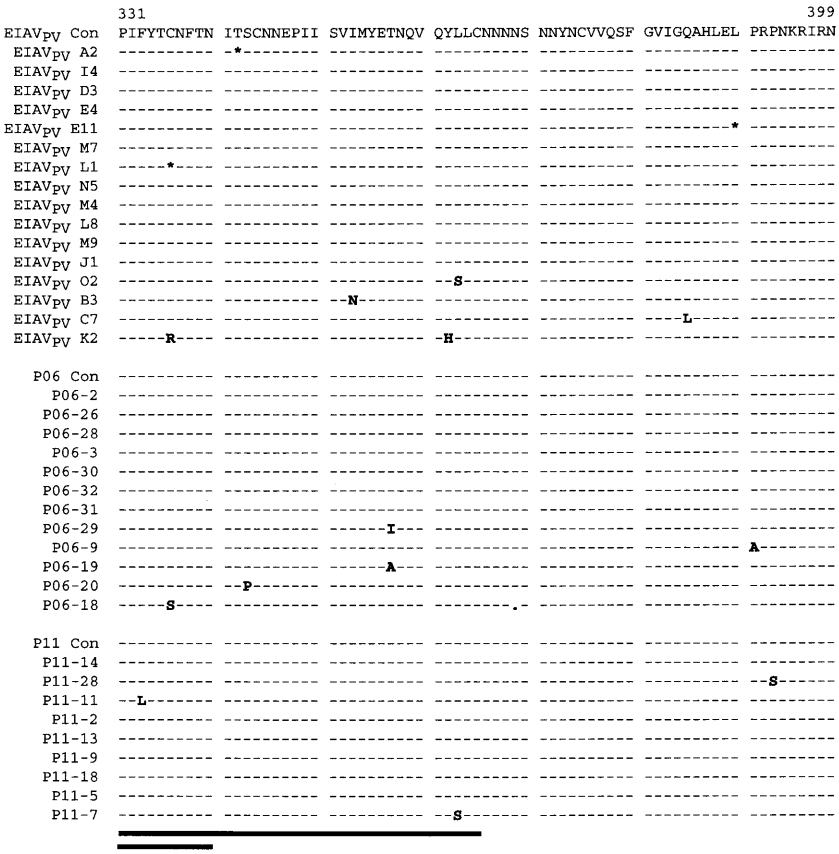


FIG. 3. Deduced amino acid sequences of the EIAV gp90 variable and flanking regions of $EIAV_{PV}$, pony 06, and pony 11. Sequences were aligned as described in the legend to Fig. 1. The single-letter amino acid code shows the consensus sequence above each group of clones. Dashes indicate an amino acid identical to the $EIAV_{PV}$ consensus sequence ($EIAV_{PV}$ Con). An asterisk indicates a silent nucleotide substitution. An X indicates the position of a single nucleotide deletion that shifts the sequence out of the proper reading frame. A gap (indicated by a dot) was introduced into clone P06-18 to align the sequences. The variable and hypervariable regions are indicated by a single and double underlines, respectively. The boxed area shows the PND, which contains two adjacent neutralizing epitopes, E_{NT} and D_{NT}. The numbering refers to the position of each sequence with respect to the first amino acid of gp90 (35). Only unique clones are shown, and the number of clones with a given amino acid sequence is shown to the right. The silent mutations shown at positions 150 and 317 of clone $EIAV_{PV}$-I4 were present only in one of the two clones with that amino acid sequence. The silent mutation shown at position 342 of clone $EIAV_{PV}$-A2 was present only in one of the four clones with that amino acid sequence. Two nucleotide substitutions were responsible for the arginine change at position 200 of clone P11-9.

lent, infectious molecular clone $EIAV_{19\text{-}2}$ was used as the backbone for restriction fragment exchanges (see Fig. 2 and 4 for the $EIAV_{19\text{-}2}$ LTR and gp90 sequences, respectively). Sequences that were identical to the consensus $EIAV_{PV}$ LTR and gp90 variable region sequences were used to generate the chimeric viruses (see Fig. 1 and 3 for the $EIAV_{PV}$ LTR and gp90 consensus sequences, respectively). Virus stocks were generated by transfection of FEK cells with the infectious molecular proviral clones. Cycle sequencing of viral RNA confirmed the presence of the expected LTR and gp90 variable region sequences in the recombinant viruses (data not shown). Comparison of the kinetics of virus replication in equine monocyte cultures showed that the parental and recombinant viruses replicated to the same level and with the same kinetics (Fig. 5). Shetland ponies (one for each chimera) were infected with $EIAV_{PV100}$ or $EIAV_{PV101}$. Rectal temperature and platelet counts were monitored as indicators of disease. Although both ponies seroconverted by 21 days postinfection, which indicated that they were infected, no clinical symptoms have been noted up to 120 days postinfection. These data suggest that the determinants of pathogenesis may not be associated with the predominant variable region of the *env* gene and the LTR of $EIAV_{PV}$.

DISCUSSION

Analysis of EIAV sequence diversity present in a pathogenic stock of virus ($EIAV_{PV}$) and during the initial disease episode in two ponies infected with $EIAV_{PV}$ was carried out to investigate the correlation between disease and the presence of particular genomic sequences. Sequence diversity in the *env* gene of the $EIAV_{PV}$ stock was low and was only slightly higher in the ponies infected with $EIAV_{PV}$. While the former result is likely due to the fact that $EIAV_{PV}$ is a biological clone (37), the latter finding was somewhat surprising since previous studies have demonstrated rapid variations in the antigenic (14, 33, 38) and pathogenic (27) properties of EIAV during persistent infections in Shetland ponies. In fact, the predominant virus isolates from successive disease episodes separated by as few as 14 days have been shown to contain different gp90 and gp45 peptide maps (27, 33). The level of EIAV *env* variation that we determined was similar to that seen early after experimental infection of monkeys with molecular clones of simian immunodeficiency virus (2, 28). Our estimate of the rate of mutation (3.1×10^{-2} to 5.4×10^{-2} mutations per site per year) was in accordance with the rates previously estimated for EIAV, HIV, and simian immunodeficiency virus (5, 12, 15, 30). Thus,

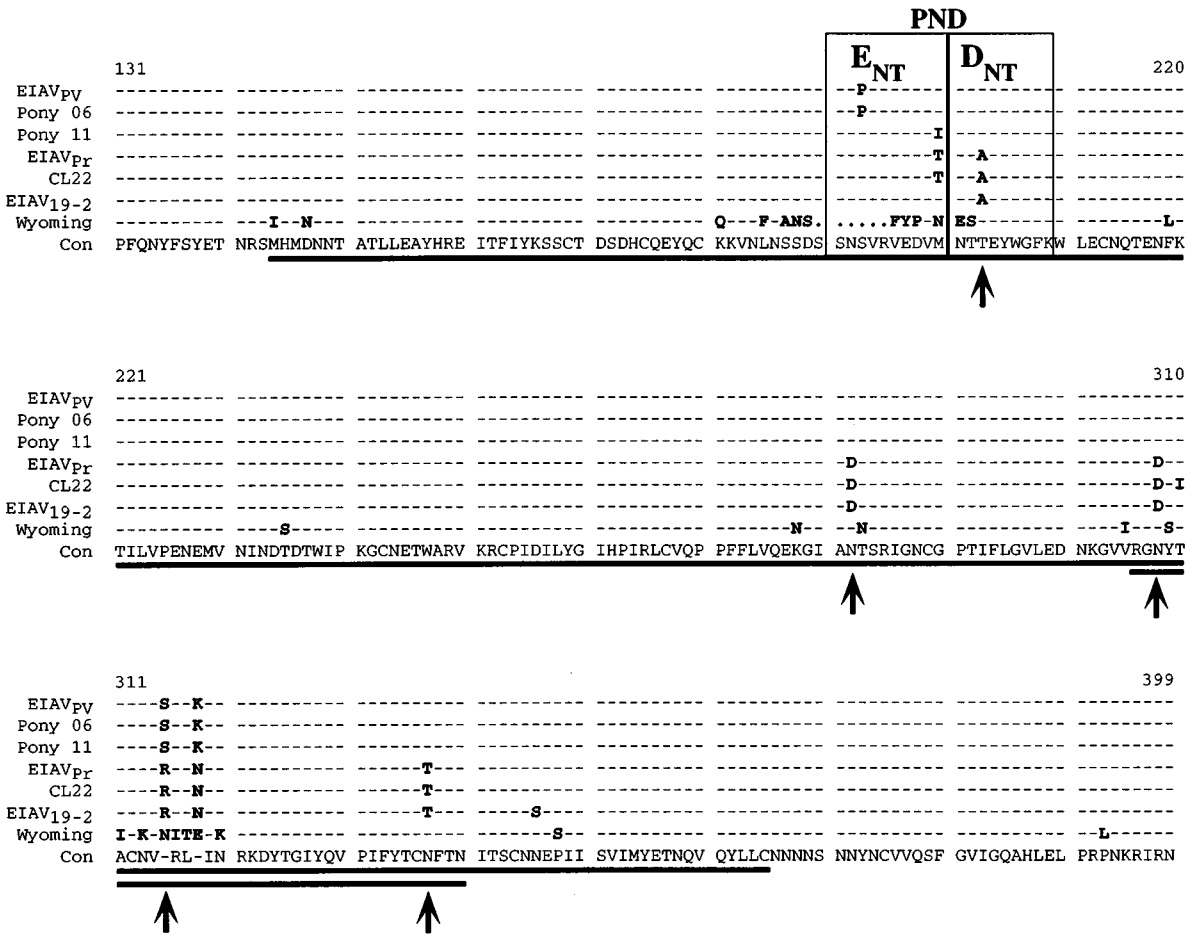


FIG. 4. Sequence alignment of the EIAV gp90 variable and flanking regions of various isolates. The consensus sequences (Con) determined in Fig. 3 were aligned with other known gp90 sequences, using the program Pileup (10). See the legend to Fig. 3 for an explanation of the sequence numbering, symbols, and features. Gaps (indicated by dots) were introduced to align the sequences. Arrows indicate the addition of potential N-glycosylation sites to the $EIAV_{PV}$ and the pony-derived sequences relative to the $EIAV_{Pr}$ and $EIAV_{19-2}$ sequences. The following sequences were obtained from GenBank and have the accession numbers shown in parentheses: $EIAV_{Pr}$ (M16575), CL22 (M87581), $EIAV_{19-2}$ (U01866), and Wyoming (this is a consensus sequence derived from the sequences assigned accession numbers M87576, M87577, M87579, M87580, M87582, M87583, and M87586 to M87589).

the low level of sequence variation detected in samples from the first disease episode may simply reflect the rate of mutation and the fact that we analyzed samples taken shortly after infection (19 or 33 days postinfection).

In contrast to our results, Alexandersen and Carpenter (1) showed that significant sequence diversity in portions of gp90 was present during the first disease episode in a sample isolated from a horse infected with a field isolate of EIAV. These workers sequenced proviral DNA isolated from macrophages after one passage of the virus in culture, as opposed to examining viral RNA isolated directly from plasma. Passage of the virus in culture may have altered the population of quasispecies as was reported previously for HIV (23). In addition, the uncharacterized, uncloned field isolate probably contained a more complex quasispecies than the biologically cloned $EIAV_{PV}$ stock, which could account for the discrepancy in the levels of variation reported by the two studies.

The observed sequence changes in the replicating EIAV quasispecies isolated from infected ponies were not randomly distributed; they were localized to either the PND (pony 11) or the PND and HVR (pony 06) (Fig. 3). This localized sequence variation was detected concomitantly with (pony 06) or shortly after (pony 11) seroconversion (data not shown), which sug-

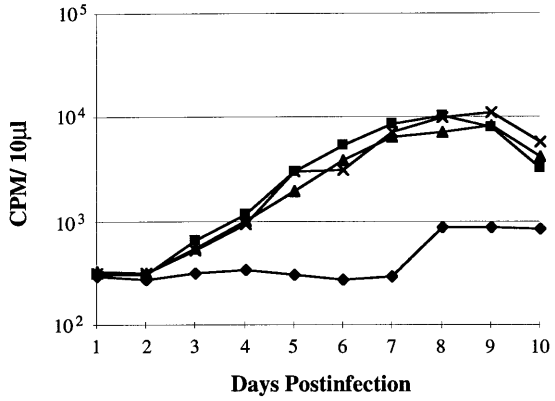


FIG. 5. Infection of equine monocytes with parental and chimeric EIAV. Equine monocytes were mock infected (◆) or infected with parental ($EIAV_{19-2}$) (■), $EIAV_{PV100}$ (▲), or $EIAV_{PV101}$ (x) at a multiplicity of 0.01 infectious center dose per cell. After 4 h at 37°C, the virus inoculum was removed and replaced with fresh medium. The medium was replaced daily. The harvested medium was clarified at 800 × *g*, and the supernatants were assayed for reverse transcriptase activity.

gested that selection by the humoral immune response could account for the genetic changes. It is important to note, however, that neutralizing antibodies were not detected at the time the plasma samples were taken. Although anti-EIAV T-cell responses have not been studied in detail, these responses may also contribute to genetic variation if the situation for EIAV is analogous to that of HIV, in which case T-cell responses are seen prior to seroconversion in HIV-infected individuals (8, 17).

Rwambo et al. (37) reported changes in the antigenic and biochemical properties of $EIAV_{PV}$ gp90, which supported the proposal that antigenic variation is responsible for escape from neutralization. The consensus $EIAV_{PV}$ gp90 sequence differed substantially from the $EIAV_{Pr}$ sequence, particularly within the PND, where there were three amino acid changes and the addition of a potential N-glycosylation site (Fig. 4). This may account for the fact that two neutralizing monoclonal antibodies directed against the $EIAV_{Pr}$ PND did not react with $EIAV_{PV}$ gp90 (37). In addition, the $EIAV_{PV}$ consensus sequence differed from the $EIAV_{Pr}$ sequence at four amino acid positions within the hypervariable region (three of which introduced potential N-glycosylation sites), suggesting that changes within this region may also be important for escape from immune surveillance. Together, these data strongly suggest that escape from neutralization was correlated, in part, to sequence changes within the PND and possibly the hypervariable region as well.

All three samples had identical consensus LTR sequences (Fig. 2), and the variation within samples was very low (Table 1). This finding is in contrast to those of previous reports which showed that this region of the LTR is hypervariable (6, 30, 31). Since the $EIAV_{PV}$ stock was biologically cloned twice, the lack of diversity for this sample was not surprising. In regard to the uniformity exhibited among the pony-derived LTRs, it is important to note that much less variability is seen when field isolates (or pony isolates) are compared with each other than when they are compared with tissue culture-adapted strains (6, 31). Our sequencing results support the idea elaborated previously (21) that selection for replication in different host cells (i.e., macrophages in vivo versus tissue culture cells in vitro) is responsible for the hypervariability seen among the LTRs of different virus stocks.

The $EIAV_{PV}$ and pony-derived LTR sequences were most similar to the Wyoming LTR sequence (Fig. 4) (34). However, analysis of potential transcription factor binding motifs showed that the latter contained a second CAAT box motif and a third potential core *ets* motif that were absent in the former. These two motifs have been suggested to be important in virus pathogenesis and replication. Two CAAT boxes are present in the LTRs of several viruses associated with disease (e.g., Wyoming, Th-1, and several virus isolates recovered during disease episodes [6, 30, 34]), whereas a single CAAT box is usually present in avirulent viruses ($EIAV_{Pr}$, MA-1, and $EIAV_{19\text{-}2}$ [6, 31, 35]). Our LTR sequencing results showed that viruses containing a single CAAT box predominated during a disease episode and that a pathogenic stock of virus need not contain an LTR with two CAAT boxes. The macrophage- and B-lymphocyte-restricted *ets* transcription factor family member PU.1 binds the EIAV LTR and is important for high-level Tat-transactivated transcription in macrophages in vitro (7, 2). Since an LTR with only two *ets* core motifs was present in all of the samples that were sequenced, including those isolated from two independent disease episodes, the data suggest that three *ets* core motifs were not required for virus-induced disease. Payne et al. (30) also reported an LTR sequence containing one CAAT box and two core *ets* motifs that was isolated from a pony during a disease episode. However, this sequence may have been altered during biological cloning of the virus in FEK cells. The possibility that our samples contained a low abundance ($\leq$1 to 8%) of sequences with two CAAT boxes and/or three core *ets* motifs cannot be ruled out.

Our results showed that the LTR and *env* sequences of $EIAV_{PV}$ were similar to those found in the plasma of two different ponies during a disease episode and were significantly different from the corresponding sequences of nonpathogenic strains (Fig. 2 and 4). Thus, we reasoned that the predominant $EIAV_{PV}$ LTR and *env* sequences might be associated with disease. To test this hypothesis, we constructed chimeric viruses that differed from the parental nonpathogenic molecular clone $EIAV_{19\text{-}2}$ only in the U3-LTR or this region in combination with the gp90 variable region. Both chimeric viruses replicated in equine monocytes to the same extent as the parental virus. However, by 120 days postinfection, neither chimera had induced disease in Shetland ponies. While a limited number of ponies were used in this study, the amount of recombinant chimeric virus used for these infections ($10^{7.0}$ infectious center doses) was at least $10^{4.0}$-fold greater than the standard $EIAV_{PV}$ inoculum that produces disease in 100% of inoculated ponies. These data suggested that the viral determinants of pathogenesis may lie outside the regions used to make the chimeric viruses. Alternatively, development of disease may involve a mixture of viral quasispecies rather than a particular viral strain, perhaps in agreement with the lack of selection of particular $EIAV_{PV}$ quasispecies during the initial stages of infection and disease as observed here. Studies using chimeras between nonpathogenic and pathogenic molecular clones of simian immunodeficiency virus indicate that although the *env* gene is an important determinant of pathogenesis, it is not the only determinant (20, 26). These results indicate that other portions of the EIAV genome should be assessed for their roles in pathogenesis.

ACKNOWLEDGMENTS

We thank John Mellors and Sharon Harrold for helpful discussions and Phalguni Gupta for critically reading the manuscript.

This work was supported by National Institutes of Health grant 5R01CA49296 and by funds from the Lucille Markey Charitable Trust and the Kentucky Agricultural Experimental Station.

REFERENCES

1. **Alexandersen, S., and S. Carpenter.** 1991. Characterization of variable regions in the envelope and S3 open reading frame of equine infectious anemia virus. J. Virol. **65:**4255–4262.
2. **Baier, M., M. T. Dittmar, K. Cichutek, and R. Kurth.** 1991. Development *in vivo* of genetic variability of simian immunodeficiency virus. Proc. Natl. Acad. Sci. USA **88:**8126–8130.
3. **Ball, J. M., K. E. Rushlow, C. J. Issel, and R. C. Montelaro.** 1992. Detailed mapping of the antigenicity of the surface unit glycoprotein of equine infectious anemia virus by using synthetic peptide strategies. J. Virol. **66:**732–742.
4. **Bosch, M. L., A. C. Andeweg, R. Schipper, and M. Kenter.** 1994. Insertion of N-linked glycosylation sites in the variable regions of the human immunodeficiency virus type 1 surface glycoprotein through AAT triplet reiteration. J. Virol. **68:**7566–7569.
5. **Burns, D. P., and R. C. Desrosiers.** 1991. Selection of genetic variants of simian immunodeficiency virus in persistently infected rhesus monkeys. J. Virol. **65:**1843–1854.
6. **Carpenter, S., S. Alexandersen, M. J. Long, S. Perryman, and B. Chesebro.** 1991. Identification of a hypervariable region in the long terminal repeat of equine infectious anemia virus. J. Virol. **65:**1605–1610.
7. **Carvalho, M., and D. Derse.** 1993. The PU.1/Spi-1 proto-oncogene is a transcriptional regulator of a lentivirus promoter. J. Virol. **67:**3885–3890.
8. **Couillin, I., P. B. Culmann, E. Gomard, J. Choppin, J. P. Levy, J. G. Guillet, and S. Saragosti.** 1994. Impaired cytotoxic T lymphocyte recognition due to genetic variations in the main immunogenic region of the HIV-1 nef protein. J. Exp. Med. **180:**1129–1134.
9. **Gao, F., L. Yue, A. T. White, P. G. Pappas, J. Barchue, A. P. Hanson, B. M. Greene, P. M. Sharp, G. M. Shaw, and B. H. Hahn.** 1992. Human infection

by genetically diverse SIV_{SM}-related HIV-2 in west Africa. Nature (London) **358:**495–499.
10. **Genetics Computer Group.** 1994. Program manual for the Wisconsin Package. Genetics Computer Group, Madison, Wis.
11. **Grund, C. H., P. R. Calkins, and R. C. Montelaro.** Quantification of equine infectious anemia virus (EIAV) infectivity by an immunological plaque assay. Unpublished data.
12. **Hahn, B. H., G. M. Shaw, M. E. Taylor, R. R. Redfield, P. D. Markham, S. Z. Salahuddin, S. F. Wong, R. C. Gallo, E. S. Parks, and W. P. Parks.** 1986. Genetic variation in HTLV-III/LAV over time in patients with AIDS or at risk for AIDS. Science **232:**1548–1553.
13. **Ho, D. D., A. U. Neumann, A. S. Perelson, W. Chen, J. M. Leonard, and M. Markowitz.** 1995. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. Nature (London) **373:**123–126.
14. **Hussain, K. A., C. J. Issel, K. L. Schnorr, P. M. Rwambo, and R. C. Montelaro.** 1987. Antigenic analysis of equine infectious anemia virus (EIAV) variants by using monoclonal antibodies: epitopes of glycoprotein gp90 of EIAV stimulate neutralizing antibodies. J. Virol. **61:**2956–2961.
15. **Johnson, P. R., T. E. Hamm, S. Goldstein, S. Kitov, and V. M. Hirsch.** 1991. The genetic fate of molecularly cloned simian immunodeficiency virus in experimentally infected macaques. Virology **185:**217–228.
16. **Kono, Y., K. Kobayashi, and Y. Fukunaga.** 1973. Antigenic drift of equine infectious anemia virus in chronically infected horses. Arch. Virol. **41:**1–10.
17. **Koup, R. A., J. T. Safrit, Y. Cao, C. A. Andrews, G. McLeod, W. Borkowsky, C. Farthing, and D. D. Ho.** 1994. Temporal association of cellular immune responses with the initial control of viremia in primary human immunodeficiency virus type 1 syndrome. J. Virol. **68:**4650–4655.
18. **Lichtenstein, D. L., K. E. Rushlow, R. F. Cook, M. L. Raabe, C. J. Swardson, G. J. Kociba, C. J. Issel, and R. C. Montelaro.** 1995. Replication in vitro and in vivo of an equine infectious anemia virus mutant deficient in dUTPase activity. J. Virol. **69:**2881–2888.
19. **Malmquist, W. A., D. Barnett, and C. S. Becvar.** 1973. Production of equine infectious anemia antigen in a persistently infected cell line. Arch. Virol. **42:**361–370.
20. **Marthas, M. L., R. A. Ramos, B. L. Lohman, R. K. Van, R. E. Unger, C. J. Miller, B. Banapour, N. C. Pedersen, and P. A. Luciw.** 1993. Viral determinants of simian immunodeficiency virus (SIV) virulence in rhesus macaques assessed by using attenuated and pathogenic molecular clones of SIV_{mac}. J. Virol. **67:**6047–6055.
21. **Maury, W.** 1994. Monocyte maturation controls expression of equine infectious anemia virus. J. Virol. **68:**6270–6279.
22. **McNearney, T., Z. Hornickova, R. Markham, A. Birdwell, M. Arens, A. Saah, and L. Ratner.** 1992. Relationship of human immunodeficiency virus type 1 sequence heterogeneity to stage of disease. Proc. Natl. Acad. Sci. USA **89:**10247–10251.
23. **Meyerhans, A., R. Cheynier, J. Albert, M. Seth, S. Kwok, J. Sninsky, M. L. Morfeldt, B. Asjo, and H. S. Wain.** 1989. Temporal fluctuations in HIV quasispecies *in vivo* are not reflected by sequential HIV isolations. Cell **58:**901–910.
24. **Montelaro, R. C., J. M. Ball, and K. E. Rushlow.** 1993. Equine retroviruses, p. 257–360. *In* J. A. Levy (ed.), The Retroviridae, vol. 2. Plenum Press, New York.
25. **Montelaro, R. C., B. Parekh, A. Orrego, and C. J. Issel.** 1984. Antigenic variation during persistent infection by equine infectious anemia virus, a retrovirus. J. Biol. Chem. **259:**10539–10544.
26. **Novembre, F. J., P. R. Johnson, M. G. Lewis, D. C. Anderson, S. Klumpp, H. M. McClure, and V. M. Hirsch.** 1993. Multiple viral determinants contribute to pathogenicity of the acutely lethal simian immunodeficiency virus SIV_{smmPBj} variant. J. Virol. **67:**2466–2474.
27. **Orrego, A., C. J. Issel, R. C. Montelaro, and W. J. Adams.** 1982. Virulence and *in vitro* growth of a cell-adapted strain of equine infectious anemia virus after serial passage in ponies. Am. J. Vet. Res. **43:**1556–1560.
28. **Overbaugh, J., L. M. Rudensey, M. D. Papenhausen, R. E. Benveniste, and W. R. Morton.** 1991. Variation in simian immunodeficiency virus *env* is confined to V1 and V4 during progression to simian AIDS. J. Virol. **65:** 7025–7031.
29. **Payne, S., B. Parekh, R. C. Montelaro, and C. J. Issel.** 1984. Genomic alterations associated with persistent infections by equine infectious anaemia virus, a retrovirus. J. Gen. Virol. **65:**1395–1399.
30. **Payne, S. L., F. D. Fang, C. P. Liu, B. R. Dhruva, P. Rwambo, C. J. Issel, and R. C. Montelaro.** 1987. Antigenic variation and lentivirus persistence: variations in envelope gene sequences during EIAV infection resemble changes reported for sequential isolates of HIV. Virology **161:**321–331.
31. **Payne, S. L., J. Rausch, K. Rushlow, R. C. Montelaro, C. Issel, M. Flaherty, S. Perry, D. Sellon, and F. Fuller.** 1994. Characterization of infectious molecular clones of equine infectious anaemia virus. J. Gen. Virol. **75:**425–429.
32. **Payne, S. L., K. Rushlow, B. R. Dhruva, C. J. Issel, and R. C. Montelaro.** 1989. Localization of conserved and variable antigenic domains of equine infectious anemia virus envelope glycoproteins using recombinant *env*-encoded protein fragments produced in *Escherichia coli.* Virology **172:**609–615.
33. **Payne, S. L., O. Salinovich, S. M. Nauman, C. J. Issel, and R. C. Montelaro.** 1987. Course and extent of variation of equine infectious anemia virus during parallel persistent infections. J. Virol. **61:**1266–1270.
34. **Perry, S. T., M. T. Flaherty, M. J. Kelley, D. L. Clabough, S. R. Tronick, L. Coggins, L. Whetter, C. R. Lengel, and F. Fuller.** 1992. The surface envelope protein gene region of equine infectious anemia virus is not an important determinant of tropism in vitro. J. Virol. **66:**4085–4097.
35. **Rushlow, K., K. Olsen, G. Stiegler, S. L. Payne, R. C. Montelaro, and C. J. Issel.** 1986. Lentivirus genomic organization: the complete nucleotide sequence of the *env* gene region of equine infectious anemia virus. Virology **155:**309–321.
36. **Rwambo, P. M., C. J. Issel, W. J. Adams, K. A. Hussain, M. Miller, and R. C. Montelaro.** 1990. Equine infectious anemia virus (EIAV) humoral responses of recipient ponies and antigenic variation during persistent infection. Arch. Virol. **111:**199–212.
37. **Rwambo, P. M., C. J. Issel, K. A. Hussain, and R. C. Montelaro.** 1990. *In vitro* isolation of a neutralization escape mutant of equine infectious anemia virus (EIAV). Arch. Virol. **111:**275–280.
38. **Salinovich, O., S. L. Payne, R. C. Montelaro, K. A. Hussain, C. J. Issel, and K. L. Schnorr.** 1986. Rapid emergence of novel antigenic and genetic variants of equine infectious anemia virus during persistent infection. J. Virol. **57:**71–80.
39. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
40. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
41. **Tabor, S., and C. C. Richardson.** 1987. DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. Proc. Natl. Acad. Sci. USA **84:**1467–1471.
42. **Vartanian, J. P., A. Meyerhans, B. Asjo, and S. Wain-Hobson.** 1991. Selection, recombination, and G→A hypermutation of human immunodeficiency virus type 1 genomes. J. Virol. **65:**1779–1788.
43. **Wain-Hobson, S., P. Sonigo, M. Guyader, A. Gazit, and M. Henry.** 1995. Erratic G→A hypermutation within a complete caprine arthritis-encephalitis virus (CAEV) provirus. Virology **209:**297–303.
44. **Wei, X., S. K. Ghosh, M. E. Taylor, V. A. Johnson, E. A. Emini, P. Deutsch, J. D. Lifson, S. Bonhoeffer, M. A. Nowak, B. H. Hahn, M. S. Saag, and G. M. Shaw.** 1995. Viral dynamics in human immunodeficiency virus type 1 infection. Nature (London) **373:**117–122.