AIDS RESEARCH AND HUMAN RETROVIRUSES
Volume 15, Number 1, 1999, pp. 23–32
Mary Ann Liebert, Inc.

Genetic Characterization of the *nef* Gene from Human Immunodeficiency Virus Type 1 Group M Strains Representing Genetic Subtypes A, B, C, E, F, G, and H

VERONIQUE JUBIER-MAURIN,[1] SENTOB SARAGOSTI,[2] JEAN-LUC PERRET,[3] EITEL MPOUDI,[4] EKA ESU-WILLIAMS,[5] CLAIRE MULANGA,[1] FLORIAN LIEGEOIS,[1] MICHEL EKWALANGA,[6] ERIC DELAPORTE,[1] and MARTINE PEETERS[1]

ABSTRACT

Most efforts to characterize sequence variation of HIV isolates has been directed toward the structural envelope gene. Few studies have evaluated the sequence variability of auxiliary genes such as *nef*. In this study 41 new HIV-1 strains, representing the majority of the described envelope subtypes of HIV-1 (A to H), were genetically characterized in the *nef* region. Phylogenetic analysis showed that 34 strains could be classified in the same subtype in *nef* and *env*, and 7 (19%) of the 41 new viruses were recombinants. For two of the seven strains, recombination occurred upstream of the *nef* gene, whereas for five of the seven strains recombination occurred within the *nef* gene with a crossover close to the 5′ end of the LTR (long terminal repeat). The low intersubtype distance between subtype B and D in the *nef* gene confirms previous observations in the *pol*, *env*, and *gag* genes, which suggest a common ancestor for these subtypes. The majority of all the previously described functional domains in the *nef* gene were relatively conserved among the different subtypes, with only minor differences being observed. The myristoylation signal among the different subtypes, with only minor differences being observed. The myristoylation signal was less conserved for subtype C, with one or more amino acid changes being observed at positions 3, 4, and 5. The highly conserved acidic region (positions 62 to 65), critical for the enhancement of viral synthesis with an increased virus growth rate, was less conserved among the subtype G strains from our study. At least three epitopic regions of the *nef* gene have been defined and each can be recognized by CTLs under a variety of HLA restrictions; all were also relatively well conserved between the different genetic subtypes. Despite the relatively important genetic variation in *nef* sequences obtained among the different genetic subtypes, functional domains and CTL epitopes were relatively well conserved. *In vitro* and/or *in vivo* studies are necessary to study the relevance of the observed differences.

INTRODUCTION

SINCE THE DEMONSTRATION THAT *nef*-DEFECTIVE simian immunodeficiency virus (SIV) can cause attenuated disease in rhesus macaques, the *nef* gene has been the subject of many studies. The importance of *nef* in the pathogenesis of human immunodeficiency virus type 1 (HIV-1) disease was revealed when an individual with hemophilia who remained healthy despite prolonged HIV-1 infection (for more than 10 years) was found to be infected with a virus possessing deletions in the Nef-coding sequence and the U3 region of the long terminal repeat (LTR). This individual showed an exceptionally low level of ongoing HIV-1 replication and viral load and stable high $CD4^+$ T cell counts.[1] More recently, an unusual cohort of HIV-1-infected blood transfusion recipients was described whose members, as well as the blood donor who represented the com-

[1]Laboratoire Retrovirus, ORSTOM, 34032 Montpellier Cedex 1, France.
[2]IMEA-INSERM, 90518 Paris, France.
[3]Hôpital Laveran, HIA, 13978 Marseille, France.
[4]Military Hospital, BP 406 Yaounde, Cameroun.
[5]University of Calabar, P.O. Box 115, Calabar, Nigeria.
[6]INRB, Kinshasa, Democratic Republic of Congo.

mon source of the infections, were all found to harbor *nef*-defective viruses. These individuals showed no progression to AIDS after 10 to 14 years of infection and, like rhesus macaques infected with *nef*-defective mutants of SIV, have maintained low viral loads and normal CD4 counts.[2] In addition, animals immunized with a similar construct can be protected when challenged with a wild-type pathogenic SIV, and reversion of these *nef* mutants *in vivo* results in regain of high virus load and rapid disease progression.[3,4] These results demonstrate that viral defects can be responsible for, or at least can contribute to, the absence of disease progression in SIV-infected rhesus monkeys and HIV-1-infected humans. On the other hand, some studies do not support a strict correlation between the structure or the function of the *nef* gene and rates of disease progression.[5,6]

Nef is a myristoylated, membrane-associated protein, and is one of the earliest viral gene products expressed in newly infected target cells. Although Nef is clearly important for viral replication *in vivo*, it has been difficult to determine the basis for this. Several claims have been made on the function of the *nef* gene product, many of which have been disputed or unconfirmed, and yet the precise role of Nef has not been completely clarified. However, several consistent consequences of Nef expression have been defined; enhancement of viral replication, downregulation of cell surface CD4 and HLA class I molecules, and signal transduction in cells.[7–13]

Multiple HLA-restricted cytotoxic T lymphocyte (CTL) epitopes have been identified in Nef, and CTLs targeted against these epitopes have been postulated to play a role in protection from infection in HIV-exposed subjects.[14,15] Nevertheless, impaired CTL recognition due to genetic variations in the main immunogenic region of the HIV-1 Nef protein has been observed.[16,17]

The genetic variability of HIV-1 is a major challenge in the development of a globally effective vaccine. With the characterization of many isolates of HIV from Africa and from other regions of the world, numerous genetic subtypes of the virus have been identified.[18] Phylogenetic analysis revealed two groups of HIV-1 isolates: group M, the major group with at least 10 different genetic subtypes (A–J), and group O.[19] Analyses of subgenomic as well as full-length HIV-1 sequences identified a surprising number of HIV-1 strains that clustered in different subtypes in different parts of their genome.[20–23] All representatives of subtypes E and G that have been sequenced to date represent mosaic genomes, with parts of the viral genome clustering with subtype A. These mosaic viruses have clearly established infection based on their broad geographic spread.[21,23] The capacity for HIV subtypes to recombine is an important parameter of viral evolution, allowing dramatic and rapid genetic change. This genetic diversity can have implications for vaccine development, since vaccine formulations based on only one virus strain (or subtype) would not elicit a broad enough immune response to protect against members of other subtypes. The impact of the genetic variation on pathogenesis, disease progression, and virus transmission also remains to be elucidated, and the cross-reactivity of CTL epitopes to conserved regions and to what extent their role is critical for vaccine development must be studied further.[24]

Functional epitopes and CTL epitopes in Nef play a possible role in HIV-1 pathogenesis. To examine to what extent these sites are conserved, further *in vivo* sequence data must be accumulated from other HIV-1 genetic subtypes. The current database of Nef sequences contains almost exclusively subtype B viruses, some subtype E viruses, and a limited number of subtype A, C, D, F, G, and H viruses. In this study we characterized genetically the *nef* genes from 41 new HIV-1 viruses representing 7 different genetic subtypes.

MATERIALS AND METHODS

Virus strains

Forty-one new HIV-1 viruses, representing different genetic subtypes in the envelope region, were studied: 5 A, 4 B, 18 C, 3 E, 4 F, 5 G, and 2 H. The samples were collected from various geographic locales or were obtained from patients living in France but who became infected with HIV-1 after overseas employment as previously described.[25] Table 1 summarizes the origin and the *env* subtype of the various samples tested.

The genetic subtype of the envelope region was determined by the heteroduplex mobility assay (HMA) as described by Delwart *et al.*[26] and/or by sequencing of the C2V3 region of the envelope as previously described.[25]

Amplification and sequencing of the nef *gene*

DNA was extracted from primary or cultured peripheral blood mononuclear cells (PBMCs) by using the IsoQuick DNA extraction kit (Microprobe, Garden Cove, CA). A nested polymerase chain reaction (PCR) was done to amplify a 770-bp fragment containing the entire *nef* gene, using the following primers, which were previously described in a study of *nef* genes from subtype E-infected patients in Thailand[27]: outer primers outer-5-1e (5′ GTGCCTCTTCAGCTACCACCG) and outer-3-3e (5′ AGCATCTGAGGGTTAGCCACT) and inner primers inner-5-1e (5′ TGGACAGAYAGGGTTATAGAA) and inner-3-7e (5′ CACCTCCCCTGGAAAGTCCCC). The PCR reaction mixture contained a 0.2 mM concentration of each deoxynucleoside triphosphate, 1.5 mM $MgCl_2$, 50 mM KCl, 10 mM Tris-HCl (pH 9.0), 0.1% Triton X-100, 2.5 U of *Taq* polymerase (Promega, Madison, WI), and 25 pmol of each primer. The two amplification rounds were identical, with an initial denaturation step at 92°C for 5 min, followed by 27 cycles of 94°C for 45 sec, 55°C for 30 sec, and 72°C for 1 min 30 sec and a final extension of 5 min at 72°C.

Nucleotide sequences were obtained by direct sequencing of the PCR products. The amplified DNA was purified, after separation of the bands on low-melting-point agarose gels, by the use of the Wizard DNA clean-up system (Promega). Cycle sequencing by the dye-terminator method with an automated DNA sequencer (model 373 Stretch; Applied Biosystems, Foster City, CA) was performed as recommended by the manufacturer.

GenBank accession numbers for the HIV-1 *nef* sequences reported in this study are as follows: AJ 232956 to AJ 232996 for the 41 new HIV-1 strains and AJ 232997 to AJ 233048 for the 52 *nef* sequences obtained from 14 French patients (2–8 clones per patient).

Phylogenetic analysis

Nucleotide sequences were aligned using CLUSTAL W[28] with minor manual adjustments, bearing in mind the protein se-

TABLE 1. ORIGIN AND ENVELOPE SUBTYPE OF VARIOUS HIV-1 VIRUSES STUDIED IN *nef* GENE

Number	*Year of sample collection*	*Country of HIV infection*	*Country of residence*	*Env subtype C2–V3 or V3–V5*	nef *subtype*
MP20	1995	Djibouti	France	A	A
MP23	1995	Djibouti	France	A	A
MP33	1995	Djibouti	France	A	U/E
MP58	1995	CAR	France	A	A
MP117	1995	Guyane?	France	A	U/E
MP25	1995	Guyan	France	B	B
MP32	1995	Tchaad	France	B	B
MP53	1995	Djibouti	France	B	B
MP328	1996	Djibouti	France	B	B
MP3	1995	Djibouti	France	C	B
MP19	1995	Djibouti	France	C	C
MP37	1995	Djibouti	France	C	C
MP40	1995	Djibouti	France	C	C
MP41	1995	Djibouti	France	C	C
MP43	1995	Djibouti	France	C	C
MP83	1995	Djibouti	France	C	C
MP98	1995	Djibouti	France	C	C
MP102	1995	Djibouti	France	C	C
MP116	1995	Djibouti	France	C	U/G
MP129	1995	Djibouti	France	C	C
MP130	1995	Djibouti	France	C	C
MP148	1995	Djibouti	France	C	C
MP169	1995	Djibouti	France	C	C
MP197	1995	Djibouti	France	C	C
MP333	1995	Djibouti	France	C	C
MP348	1995	Djibouti	France	C	C
LBV10-5	1989	Gabon	Gabon	C	F/A
MP38	1995	Cambodia	France	E	E
MP59	1995	Cambodia	France	E	E
MP126	1995	Cambodia	France	E	E
MP255	1995	Cameroon	Cameroon	F	F
MP257	1995	Cameroon	Cameroon	F	F
MP411	1996	Tchaad?	France	F	F
MP535	1996	Cameroon	Cameroon	F	U
ML136	1995	Mali	Mali	G	G
ML150	1995	Mali	Mali	G	G
IKCSW22	1996	Nigeria	Nigeria	G	G
MACSW39	1996	Nigeria	Nigeria	G	G
VI526	1990	Gabon	Gabon	G	H/G
ZR21	1996	DRC	DRC	H	H
ZR23	1996	DRC	DRC	H	H

Abbreviations: CAR, Central African Republic; DRC, Democratic Republic of Congo (formerly Zaire); U, unknown subtype.

quences. Phylogenetic trees, created using the neighbor-joining method, and reliability of the branching orders, determined using the bootstrap approach, were implemented by using CLUSTAL W. Genetic distances were calculated by the Kimura two-parameter method.[29]

Identification of viruses recombinant in the *nef* region was done using the Recombinant Identification Program (RIP) from the Los Alamos Database;[30] by constructing the phylogenetic trees of 300 nucleotides, each overlapping the previous by 60 nucleotides; and by informative site analysis.[31] For the informative site analysis each putative recombinant sequence was aligned with the reference sequences. For RIP analysis, a consensus sequence for each subtype was derived from the viral epidemiology signature pattern analysis (VESPA) analysis,[32] using only those strains for which the subtype designation in *nef* was unambiguous.

RESULTS

Phylogenetic analysis of nef

The *nef* sequences analyzed in this study encompass the entire coding region of the *nef* gene, which is approximately 618 bp. Figure 1 shows the phylogenetic tree obtained with the various *nef* sequences from our study and with representatives of the different subtypes from the database.

Thirty-five of the 41 new HIV-1 strains studied clustered with high bootstrap values with reference strains from the database and were classified as follows: 3 A, 5 B, 15 C, 3 E, 3 F, 4 G, and 2 H (Fig. 1). The ML136 strain, which has a deletion of 66 nucleotides, was excluded from the phylogenetic analy-

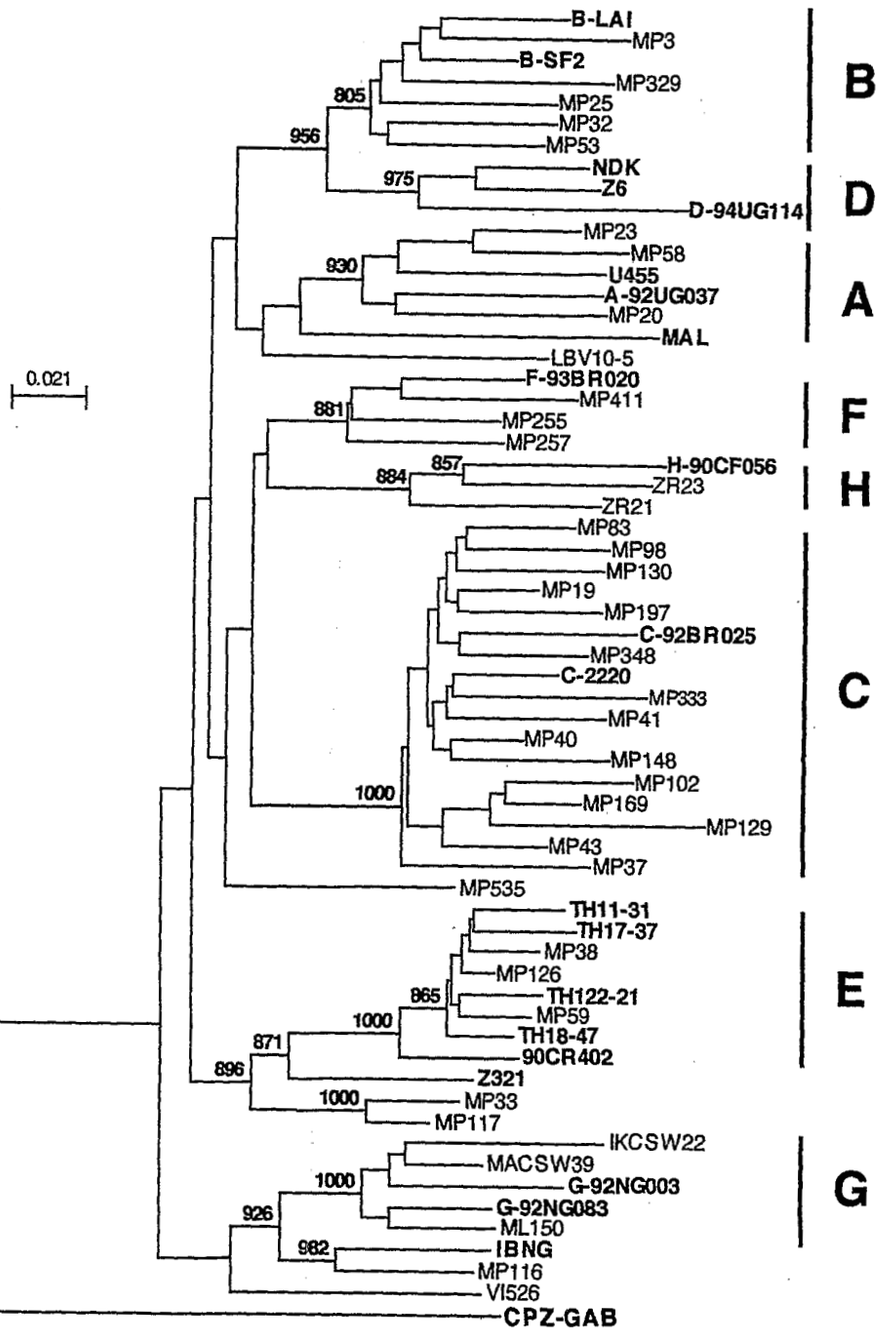


FIG. 1. Phylogenetic tree based on 618 unambiguously aligned nucleotides of the *nef* gene from 40 new HIV-1 isolates and reference strains representing the various genetic subtypes. The tree was rooted with the corresponding region of the chimpanzee SIV cpz-gab isolate as an outgroup. The analysis was performed as described in Materials and Methods.

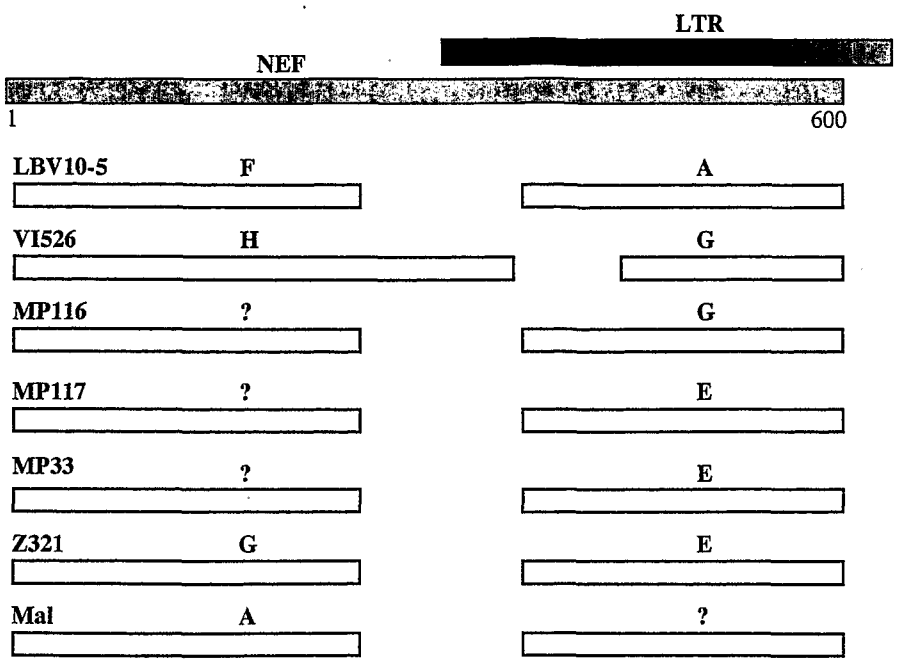


FIG. 2. Putative recombination break points within the *nef* gene were placed, taking into account the informative site analysis, the phylogenetic tree analysis of overlapping fragments, and the results obtained by RIP (samples with the break point upstream of the gene are not shown here).

sis shown in Fig. 1. Since the phylogenetic analysis excludes gaps in any sequence, inclusion of this sequence in the analysis would result in using a shorter alignment for the analysis. Two strains, MP535 and LBV10-5, formed each a separate single branch, whereas MP33 and MP117 clustered together with high bootstrap values, and formed a separate single cluster. In addition, the VI526 strain seemed only distantly related to the subtype G strains, and MP116 together with the IBNG reference strain formed a separate group within the subtype G strains.

Analysis for viruses recombinant in the nef *gene*

For the 41 new viruses, the genetic subtype was previously identified in the V3–V5 region of the envelope by heteroduplex mobility assay (HMA) and/or sequencing. As shown in Table 1, 34 of the 41 HIV-1 strains had the same subtype in *nef* and *env*. For the remaining seven strains (MP3, MP33, MP117, MP116, MP535, VI526, and LBV10-5) the subtype assignation in the envelope and the *nef* gene were different. Two of the five *env* subtype A strains (MP33 and MP117) formed a separate cluster. Three of the 18 *env* C samples clustered with other subtypes: MP3 with the subtype B strains, MP116 forming a separate group within subtype G, and LBV10-5 forming a separate single branch. MP535 (*env* F) formed a separate single branch in the *nef* region. The VI526 strains (*env* G) was only distantly related to the other G strains.

In addition, two *nef* sequences from the database, Z321 and MAL, both isolated from patients originating from the Democratic Republic of Congo (formerly Zaire), did not cluster with any of the known subtypes. The Z321 strain, which is a recombinant virus (*gag* G and *env* A),[33] is close to subtype E in *nef*, and the MAL strain, an A/D/I recombinant,[20] could not be classified at all.

All the strains with discordant subtypes between the *env* and the *nef* regions, and those that were only distantly related to a particular subtype in the phylogenetic tree, were further analyzed for recombination by RIP. This analysis showed that for two strains the recombination site is upstream of *nef*: the MP3 strain (*env* C) was entirely B in the *nef* region, and for the MP535 strain (*env* F) the entire *nef* gene sequence could not be classified. For the remaining samples, RIP analysis, followed by phylogenetic tree and informative site analysis, revealed that they were recombinant within the *nef* gene (Fig. 2). Two *env* subtype C strains, MP116 and LBV10-5, were identified as U (unknown)/G and F/A recombinants, respectively, in the *nef* gene. The MP116 strain forms a separate cluster together with the IBNG reference strain in the phylogenetic tree from the entire *nef* gene, and RIP analysis of these latter strains showed that MP116 and IBNG had a similar recombination pattern in *nef*. The *env* subtype A strains MP33 and MP117 did not belong to any of the known subtypes in the 5′ part of the *nef* gene, but belonged to subtype E in the 3′ part. This latter was supported with high bootstrap values in the phylogenetic tree of the corresponding region. The VI526 strain, which was distantly related to subtype G (Fig. 1), appeared to be a recombinant H/G virus. RIP analysis of the reference strains MAL and Z321 showed recombinant A/U and G/E viruses. All subtype E strains were in the *nef* gene as previously described for the reference subtype E strains, i.e., close to subtype A in the 5′ end and to subtype E in the 3′ end.

Inter- and intragenetic distances in nef

The genetic distances, calculated by the Kimura two-parameter method, are shown in Table 2. The strains that were recombinant in the *nef* gene, except for subtype E, were excluded from the inter- and intrasubtype distance calculations. Within

TABLE 2. INTER- AND INTRASUBTYPE DISTANCES[a] IN THE *nef* GENE, CALCULATED BY THE KIMURA TWO-PARAMETER METHOD[b]

	Subtype							
Subtype	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*
A	12.1							
B	21.1	9.8						
C	21.6	20.4	9.8					
D	22.8	14.4	21.8	10.4				
E	18.9	21.1	19.9	23.4	4.9			
F	19.1	16.5	16.8	18.9	21.3	9.3		
G	20.3	22.1	21.9	23.8	18.5	21.0	8.5	
H	21.8	20.9	21.9	21.7	21.1	16.7	22.2	11.8

[a]In percent.

[b]Distances were calculated using the same nonrecombinant *nef* sequences used to construct the phylogenetic tree in Fig. 1, including 5 subtype A, 7 subtype B, 17 subtype C, 3 subtype D, 8 subtype E, 4 subtype F, 5 subtype G, and 3 subtype H sequences.

subtypes, the mean intrasubtype distance was 9.6% (4.9 to 12.1%), with the lowest diversity within subtype E samples and the highest variation within subtype A. The mean intersubtype distance was 20.4% (14.4 to 23.8%). The lowest intersubtype distance (14.4%) was observed between *nef* sequences from subtypes B and D.

Pattern of amino acid conservation among Nef sequences

The 41 predicted amino acid Nef sequences were aligned, together with 21 reference sequences from the database, with a consensus Nef sequence derived from 52 sequences (14 French patients, 2–8 clones per patient) from subtype B-infected patients in France (S. Saragosti and I. Couillin, unpublished results, 1997). The consensus signature amino acids of each subtype, obtained using the VESPA program[32] with a threshold of 0.80, and the consensus amino acid sequence from the French B isolates, are represented in Figure 3. We examined sequences for the presence of domains that have been shown to contain important biological characteristics. The positions of the amino acids correspond to the LAI sequence. The myristoylation signal (GGKWSK), residues 2 to 7, is relatively conserved among the different genetic subtypes. The glycine at position 2, the serine at position 6, and the lysine at position 7 were conserved in all of the strains analyzed. Within the myristoylation signal the highest variation was seen among subtype C samples, with 12 of the 15 samples showing 1 or more amino acid changes at positions 3 and/or 4 and/or 5. Strain ML136, identified as subtype G in the envelope and *nef* region, shows an important deletion of 22 amino acids as compared with the consensus B sequences, starting immediately after the myristoylation signal, while two other strains (ZR21 and ZR23) as well as Z321 show a small deletion (2–4 amino acids). Interestingly, these three strains originate from the Democratic Republic of Congo (formerly Zaire), and the Z321 (collected in 1976) and ZR21 (collected in 1996) strains have the same deletion.

The myristoylation signal is followed by a region of extensive variability and polymorphism compared with other regions of the Nef sequence. However, the previously described, highly conserved glycine or glutamic acid at position 12 was conserved among all the strains; 59 had glycine and 2 had glutamic acid at this position. The highly conserved tryptophan at position 13 was conserved in 57 of the 61 strains from our alignment; 2 strains (IKCSW22 and C-92BRO25) had an arginine at this position, strain ZR21 had a cysteine at this position, and this region is deleted in the ML136 strain. The methionine at position 20, which has been shown to serve as a site for internal initiation, was observed only in 27 of the 61 strains; it was present in all of the subtype B strains. This methionine was replaced by isoleucine in 23 strains and by a leucine in the LBV-10-5 strain. The potential recognition site for phosphorylation (RPMTYK) at positions 77 to 82 is highly conserved among the different subtypes. In particular, the RPM—K amino acid motif was present in all of the aligned sequences, whereas the threonine (T) was replaced by serine (S) in four strains and the tyrosine (Y) was replaced by phenylalanine (F) in nine strains. The acidic region, mainly composed of glutamic acid residues and located at positions 62 to 66, was less conserved among subtype G strains. In subtype G strains, only two of the four glutamic acid molecules were present; one was replaced by aspartic acid and one by serine. The $(PXXP)_3$ motif at positions 69 to 78 was highly conserved. The first three amino acids from the GPG(I/V) pattern at positions 130 to 133, which is highly predictive of a β turn, is conserved in all of the strains, whereas the fourth amino acid was always hydrophobic: isoleucine for 22 samples and valine for 20 samples (including almost all of the subtype C strains), and threonine in 14 samples. All of the subtype E strains had a deletion of two amino acids at positions 48–49, and this deletion was also observed in the few E sequences from the database[34] and MP20. In addition, in all subtype E samples a cysteine (C) was present at position 138; this cysteine was also present in the recombinant *nef* strains MP33, MP117, and Z321, which belonged to subtype E in that part of the *nef* gene. In addition, the C-terminal cysteine is missing in all subtype F sequences.

Another important feature of the Nef protein is the presence of several epitopes recognized by CTLs. At least three epitopic regions of the Nef protein have been defined and each can be recognized by CTLs under a variety of HLA restrictions. The first two epitopes are present at positions that overlap the proline-rich region $(PXXP)_3$ and the potential recognition site for

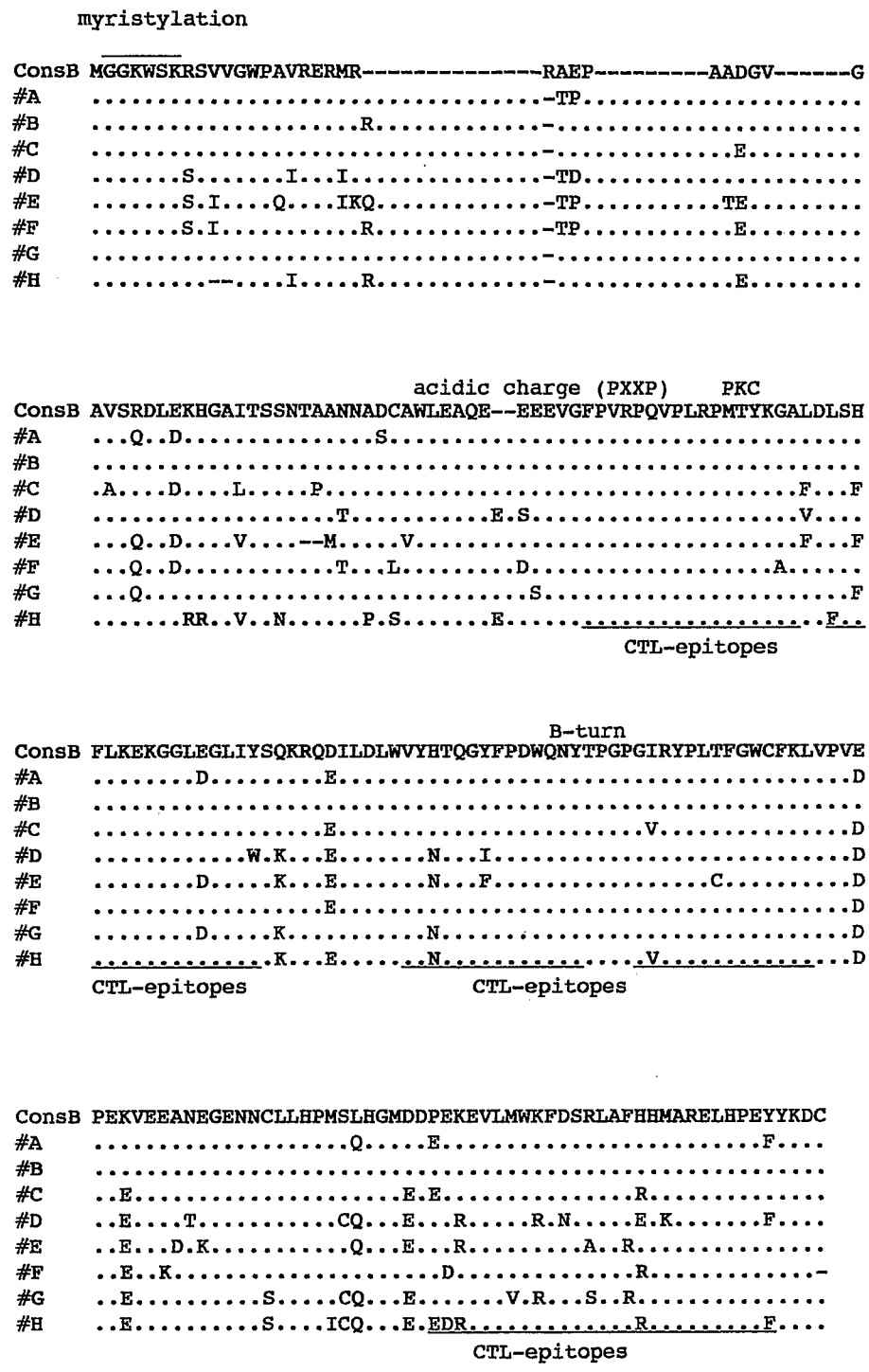


FIG. 3. HIV-1 Nef consensus sequences. The 41 predicted amino acid Nef sequences were aligned, together with the 21 reference sequences from the database included in the phylogenetic analysis, against a consensus Nef sequence generated from 52 subtype B sequences derived from subtype B-infected patients sampled in France in 1994 (14 French patients, 2–8 clones per patient). Consensus sequences have been generated for each of the defined subtypes, using the VESPA program.[32] Only the signature amino acids that are conserved in at least 80% of the sequences in each of the defined subtypes are indicated. Dots represent identity and dashes represent gaps as compared with the consensus B sequences.

phosphorylation and are therefore also highly conserved. The third CTL epitope, located between the β turn and another proline-rich region, is also relatively well conserved between the different genetic subtypes.

DISCUSSION

Most efforts to characterize sequence variation of HIV isolates has been directed toward the structural envelope gene. Few

studies have evaluated the sequence variability of auxiliary genes such as *nef*. In this study 41 new HIV-1 strains, representing the majority of the described envelope subtypes of HIV-1 (A to H), were genetically characterized in the *nef* region. Phylogenetic analysis showed that 34 strains could be classified in the same subtype in *nef* and *env* (this was supported by high bootstrap values), and 7 (19%) of the 41 new viruses were recombinants. This confirms data from previous reports that have shown that 10 to 30% of the HIV-1 sequences contained interspersed segments of the genetic material from two or more different genetic subtypes.[20–22] For two of the seven strains, recombination occurred upstream of the *nef* gene, whereas more detailed analysis showed that for five of the seven strains recombination occurred within the *nef* gene with a crossover close to the 5′ end of the 3′ long terminal repeat (LTR). In addition to these seven recombinant strains, all of the E strains from our study, which originate from Cambodia, clustered together with previously described E strains from Thailand in the *nef* gene.[23,27] All of these reference strains are known to be recombinant in the *nef* gene, with the 5′ end close to subtype A and the second part of the *nef* gene, which overlaps with the LTR, belonging to subtype E.[23]

Phylogenetic analysis and calculation of the distances between the different strains from our study show a relatively important genetic variability in *nef*, with a mean intersubtype variation of 20.4% and an intrasubtype variation of 9.6%. It is interesting to note the low intersubtype distance between subtypes B and D in the *nef* gene. This confirms previous observations of subtype B and D strains in the *pol, env*, and *gag* genes, suggesting a common ancestor for these subtypes.[35,36]

The purpose of this study was also to determine to what extent functionally important domains of the *nef* gene are conserved among different subtypes. Conservation of the myristoylation signal was expected since the subcellular targeting of Nef proteins to cytoplasmic membranes depends on the presence of an intact myristoylation signal. The amino acid sequence encoding the myristoylation signal was almost invariant among our HIV-1 Nef protein sequences, except for subtype C. The functional implications of these differences among subtype C strains must be further examined.

The methionine at position 20, which has been shown to serve as a site for internal initiation, resulting in a nonmyristoylated and truncated 25-kDa Nef protein rather than the full-length 27-kDa Nef protein,[37,38] was conserved in all of the subtype B strains, but in the majority of the other subtype Nef sequences it was replaced by isoleucine. The absence of this methionine was also described for SIV Nef sequences.[11]

Almost all of the Nef sequences contained the highly conserved sequence RPMTYK at positions 77 to 82, which is a potential recognition site for phosphorylation by protein kinase C.[39]

The highly conserved acidic region, located at positions 62 to 65 and critical for the enhancement of viral synthesis,[40] was less conserved among the subtype G strains from our study; the implications for viral growth, however, must be further studied *in vitro* and/or *in vivo*.

The proline-rich region, containing the motif $(PXXP)_3$ at positions 69 to 78, was highly conserved among all of the Nef sequences from our study. This repeat was also highly conserved among Nef sequences from SIV and HIV-2 isolates. Intact Nef *PXXP* motifs are dispensable for Nef-induced CD4 downregulation, but are required for the higher *in vitro* replicative potential of Nef^+ viruses. Thus, CD4 downregulation and promotion of viral growth are two distinct functions of Nef, and the latter is mediated via SH3 binding.[41] Although the significance of this repeat sequence is not entirely understood, it is known to correspond to a region recognized by CTLs.

The highly conserved GPG(I/V) sequence located at positions 130 to 132, which is highly predictive of a β turn, was observed among all of the subtypes studied.

Zazopoulos and Haseltine[42] have shown that under nonreducing conditions, alternative pairing of the cysteines occurs and that differences in the observed biological activity of *nef* alleles may be attributed, at least in part, to differences in the secondary structure of the proteins. Interestingly, all subtype E Nef proteins have an extra cysteine that could play a role in the biological activity of these isolates by inducing a different folding of the protein, as well as a small deletion.[27] It has been suggested that the presence of this cysteine may attenuate Nef function[34] in a long-term nonprogressor (LTNP; patient Sur25) as well as in African green monkeys, in which no known disease is associated with this virus in the wild. However, the presence of a cysteine at this position in LTSs as well as in rapid-replicating viruses invalidated this hypothesis.[43] Interestingly, to the best of our knowledge, the Sur25 Nef sequence is the only non-E sequence (subtype B) that has the two subtype E signatures (Cys-138 and the two-amino acid deletion at position 47–48), while the Nef sequences described by Blaak *et al.*[43] for LTNP and rapid progressors have only the Cys-138 signature.

HIV-1 viruses possessing deletions in the Nef-coding sequence and the U3 region of the LTR were associated with long-term nonprogression.[1,2] In our study, only one strain (ML136) showed an important deletion in the *nef* gene, i.e., that portion encoding 22 amino acids just after the myristoylation site. No detailed clinical information was available on this patient, but in contrast to the *nef* deletion described in LTNP, the deletion we observed was in the 5′ part of the *nef* gene and not in the 3′ part corresponding to the 3′ LTR.[1,2,44]

Individuals infected with HIV-1 generate a great diversity of CTLs, which are directed against several structural and regulatory proteins. The rapid elimination of infected cells by CTLs directed against proteins synthesized early in infection is an important defense against viral infection. CTLs directed against the early-expressed Nef protein of HIV could therefore provide protection *in vivo* by eliminating virus-infected cells before there is substantial increase in viral replication. Therefore a vaccine should elicit a broadly reactive CTL response. The previously described CTL epitopes encoded in the *nef* region seem relatively conserved among the various genetic subtypes. This is probably because they overlap with some important and conserved functional domains in the *nef* gene. Several amino acid substitutions were observed, but no pattern specific to a certain subtype could be identified. Some studies report a lack of CTL recognition owing to mutations in the main immunogeneic region of the Nef protein, and there is clear evidence that CTLs exert selection pressure on the viral population, during and after seroconversion. HIV-1 variants in Nef CTL epitopes were selected and were less or not all recognized by CTLs.[17] Frequently, single or double mutations found in natural epitopes are well tolerated and peptides bearing these substitutions could

stimulate generation of CTL clones that had initially responded to the wild-type form, although sometimes with reduced potency. Indeed, some studies report cross-clade recognition of CTLs directed against Gag and Env proteins, and emphasize the importance of including some of the more conserved proteins in vaccine constructs.[45–47]

Although the correlates of immune protection in HIV-1 infection are not known, the induction of CTL responses in addition to neutralizing antibodies would be an important component of an effective vaccine. However, the role of CTLs could be weakened by Nef activity. Indeed, CTLs lyse virus-infected cells that display viral peptide epitopes in association with major histocompatibility complex (MHC) class I molecules on the cell surface. Nef, by downmodulating surface expression of MHC class I molecules in infected cells, reduces epitope density on their surface, impairing lysis by cytotoxic T lymphocytes.[48,49]

In conclusion, we observed a relatively important genetic variation among *nef* sequences of various genetic subtypes. About 20% of the strains studied were recombinants, with either a crossover point upstream of the *nef* gene, or a crossover point around the 5′ end of the 3′ LTR-overlapping fragment. Despite the important genetic variation, the majority of all of the previously described functional domains in the *nef* gene were relatively conserved among the different subtypes; only minor differences were observed. *In vitro* and/or *in vivo* studies are necessary to determine their relevance.

ACKNOWLEDGMENTS

This study was cosponsored by grants from the Agence National de Recherches contre le SIDA (ANRS), the INCO-DC program from the European Union (contract IC18-CT96-0110 and IC18-CT97-0216), and by the Joint United Nations Programme on HIV/AIDS (UNAIDS) (contract refMolecu07). M.E. was supported by a fellowship from Sidaction.

REFERENCES

1. Kirchhoff F, Greenough TC, Brettler DB, Sullivan JL, and Desrosiers RC: Absence of intact Nef sequences in a long-term survivor with nonprogressive HIV-1. N Engl J Med 1995;332: 228–232.
2. Deacon NJ, Tsykin A, Solomon A, Smith K, Ludford-Menting M, Hooker DJ, McPhee DA, Greenway AL, Ellett A, Chatfield C, Lawson VA, Crowe S, Maerz A, Sonza S, Learmont J, Sullivan JS, Cunningham A, Dwyer D, Dowton D, and Mills J: Genomic structure of an attenuated quasi species of HIV-1 from a blood transfusion donor and recipients. Science 1995;270:988–991.
3. Kestler HW, Ringler DJ, Mori K, Panicali DL, Sehgal PK, Daniel MD, and Desrosiers RC: Importance of the Nef gene for maintenance of high virus loads and for development of AIDS. Cell 1991;65:651–662.
4. Daniel MD, Kirchhoff F, Czajak SC, Sehgal PK and Desrosiers, RC: Protective effects of a live attenuated SIV vaccine with a deletion in the Nef gene. Science 1992;258:1938–1941.
5. Michael NL, Chang G, d'Arcy LA, Tseng CJ, Birx DL, and Sheppard HW: Functional characterization of human immunodeficiency virus type 1 Nef genes in patients with divergent rates of disease progression. J Virol 1995;69:6758–6769.
6. Huang Y, Zhang L, and Ho DD: Biological characterization of Nef in long-term survivors of human immunodeficiency virus type 1 infection. J Virol 1995;69:8142–8146.
7. Jamieson BD, Aldrovandi GM, Planelles V, Jowett JB, Gao L, Bloch LM, Chen IS, and Zack JA: Requirement of human immunodeficiency virus type 1 Nef for in vivo replication and pathogenicity. J Virol 1994;68:3478–3485.
8. Miller MD, Warmerdam MT, Gaston I, Greene WC, and Feinberg MB: The human immunodeficiency virus-1 Nef gene product: A positive factor for viral infection and replication in primary lymphocytes and macrophages. J Exp Med 1994;179:101–113.
9. Zazopoulos E and Haseltine WA: Effect of Nef alleles on replication of human immunodeficiency virus type 1. Virology 1993; 194:20–27.
10. Goldsmith MA, Warmerdam MT, Atchison RE, Miller MD, and Greene WC: Dissociation of the CD4 downregulation and viral infectivity enhancement functions of human immunodeficiency virus type 1 Nef. J Virol 1995;69:4112–4121.
11. Shugars DC, Smith MS, Glueck DH, Nantermet PV, Seillier-Moiseiwitsch F, and Swanstrom R: Analysis of human immunodeficiency virus type 1 Nef gene sequences present in vivo. J Virol 1993;67:4639–4650.
12. Anderson S, Shugars DC, Swanstrom R, and Garcia JV: Nef from primary isolates of human immunodeficiency virus type 1 suppresses surface CD4 expression in human and mouse T cells. J Virol 1993;67:4923–4931.
13. Brady HJ, Pennington DJ, Miles CG, and Dzierzak EA: CD4 cell surface downregulation in HIV-1 Nef transgenic mice is a consequence of intracellular sequestration. EMBO J 1993;12:4923–4932.
14. Culmann-Penciolelli B, Lamhamedi-Cherradi S, Couillin I, Guegan N, Levy JP, Guillet JG, and Gomard E: Identification of multirestricted immunodominant regions recognized by cytolytic T lymphocytes in the human immunodeficiency virus type 1 Nef protein. J Virol 1994;68:7336–7343.
15. Lamhamedi-Cherradi S, Culmann-Penciolelli B, Guy B, Kieny MP, Dreyfus F, Saimot AG, Sereni D, Sicard D, Levy JP, and Gomard E: Qualitative and quantitative analysis of human cytotoxic T-lymphocyte responses to HIV-1 proteins. AIDS 1992;6:1249–1258.
16. Couillin I, Culmann-Penciolelli B, Gomard E, Choppin J, Levy JP, Guillet JG, and Saragosti S: Impaired cytotoxic T lymphocyte recognition due to genetic variations in the main immunogenic region of the human immunodeficiency virus 1 NEF protein. J Exp Med 1994;180:1129–1134.
17. Price DA, Goulder PJ, Klenerman P, Sewell AK, Easterbrook PJ, Troop M, Bangham CR, and Phillips RE: Positive selection of HIV-1 cytotoxic T lymphocyte escape variants during primary infection. Proc Natl Acad Sci USA 1997;94:1890–1895.
18. Louwagie J, McCutchan FE, Peeters M, Brennan TP, Sanders-Buell E, Eddy G, van der Groen G, Fransen K, Gershy-Damet GM, Deleys R and Burke D: Phylogenetic analysis of *gag* genes from 70 international HIV-1 isolates provides evidence for multiple genotypes. AIDS 1993;7:769–780.
19. Myers G, Hahn B, Mellors JW, Henderson LE, Korber B, Jeang K-T, McCutchan FE, and Pavlakis GN: *Human Retroviruses and AIDS Database: A Compilation and Analysis of Nucleic Acid and Amino Acid Sequences*. Los Alamos National Laboratory, Los Alamos, New Mexico, 1995.
20. Robertson DL, Gao F, Hahn BH, and Sharp PM: *Human Retroviruses and AIDS Database: Intersubtype Recombinant HIV-1 Sequences*. Los Alamos National Laboratory, Los Alamos, New Mexico, 1997.
21. McCutchan FE, Salminen MO, Carr JK, and Burke DS: HIV-1 genetic diversity. AIDS 1996;10(Suppl. 3):S13–20.
22. Cornelissen M, Kampinga G, Zorgdrager F, and Goudsmit J: Human immunodeficiency virus type 1 subtypes defined by *env* show high frequency of recombinant *gag* genes. The UNAIDS

Network for HIV Isolation and Characterization. J Virol 1996;70: 8209–8212.
23. Gao F, Robertson DL, Morrison SG, Hui H, Craig S, Decker J, Fultz P, Girard M, Shaw GM, Hahn BH, and Sharp PM: The heterosexual human immunodeficiency virus type 1 epidemic in Thailand is caused by an intersubtype (A/E) recombinant of African origin. J Virol 1996;70:7013–7029.
24. European Commission (DG XII, INCO-DC) and the Joint United Nations Programme on HIV/AIDS: HIV-1 subtypes: Implications for epidemiology, pathogenicity, vaccines and diagnostics. AIDS 1997;11:UNAIDS17–UNAIDS36. [Workship report]
25. Lasky M, Perret JL, Peeters M, Bibollet-Ruche F, Liegeois F, Patrel D, Molinier S, Gras C, and Delaporte E: Presence of multiple non-B subtypes and divergent subtype B strains of HIV-1 in individuals infected after overseas deployment. AIDS 1997;11:43–51.
26. Delwart EL, Shpaer EG, Louwagie J, McCutchan FE, Grez M, Rübsamen-Waigmann H, and Mullins JI: Genetic relationships determined by a DNA heteroduplex mobility assay: Analysis of HIV-1 *env* genes. Science 1993;262:1257–1261.
27. Artenstein AW, Hegerich PA, Beyrer C, Rungruengthanakit K, Michael NL, and Natpratan C: Sequences and phylogenetic analysis of the Nef gene from Thai subjects harboring subtype E HIV-1. AIDS Res Hum Retroviruses 1996;12:557–560.
28. Thompson J, Higgins D, and Gibson T: CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acid Res 1994;22:4673–4680.
29. Kimura M: A simple method for estimating evolutionary rate of base substitution through comparative studies of nucleotides sequences. J Mol Evol 1980;16:111–120.
30. Siepel AC, Halpern AL, Macken C, and Korber BT: A computer program designed to screen rapidly for HIV type 1 intersubtype recombinant sequences. AIDS Res Hum Retroviruses 1995;11: 1413–1416.
31. Robertson D, Hahn B, and Sharp P: Recombination in AIDS viruses. J Mol Evol 1995;40:249–259.
32. Korber B and Myers G: Signature pattern analysis: A method for assessing viral sequence relatedness. AIDS Res Hum Retroviruses 1992;8:1549–1560.
33. Choi DJ, Dube S, Spicer TP, Slade HB, Jensen FC, and Poiesz BJ: HIV type 1 isolate Z321, the strain used to make a therapeutic HIV type 1 immunogen, is intersubtype recombinant. AIDS Res Hum Retroviruses 197;13:357–361.
34. Premkumar DR, Ma XZ, Maitra RK, Chakrabarti BK, Salkowitz J, Yen-Lieberman B, Hirsch MS, and Kestler HW: The Nef gene from a long-term HIV type 1 nonprogressor. AIDS Res Hum Retroviruses 1996;12:337–345.
35. Cornelissen M, van den Burg R, Zorgdrager F, Lukashov V, and Goudsmit J: Pol gene diversity of five human immunodeficiency virus type 1 subtypes: Evidence for naturally occurring mutations that contribute to drug resistance, limiting recombination patterns, and common ancestry for subtypes B and D. J Virol 1997; 71:6348–6358.
36. Zhu T, Korber BT, Nahmias AJ, Hooper E, Sharp PM, and Ho DD: An African HIV-1 sequence from 1959 and implications for the origin of the epidemic. Nature (London) 1998;391:594–597.
37. Kaminchik J, Bashan N, Itach A, Sarver N, Gorecki M, and Panet A: Genetic characterization of human immunodeficiency virus type 1 Nef gene products translated in vitro and expressed in mammalian cells. J Virol 1991;65:583–588.
38. Zweig M, Samuel KP, Showalter SD, Bladen SV, Dubois GC, Lautenberger JA, Hodge DR, and Papas TS: Heterogeneity of Nef proteins in cells infected with human immunodeficiency virus type 1. Virology 1990;179:504–507.
39. Kennelly PJ and Krebs EG: Consensus sequences as substrate specificity detreminants for protein kinases and protein phosphatases. J Biol Chem 1991;266:15555–15558.
40. Kawano Y, Tanaka Y, Misawa N, Tanaka R, Kira JI, Kimura T, Fukushi M, Sano K, Goto T, Nakai M, Kobayashi T, Yamamoto N, and Koyanagi Y: Mutational analysis of human immunodeficiency virus type 1 (HIV-1) accessory genes: Requirement of a site in the Nef gene for HIV-1 replication in activated CD4+ T cells in vitro and in vivo. J Virol 1997;71:8456–8466.
41. Saksela K, Cheng G, and Baltimore D: Proline-rich (PxxP) motifs in HIV-1 Nef bind to SH3 domains of a subset of Src kinases and are required for the enhanced growth of Nef+ viruses but not for down-regulation of CD4. EMBO J 1995;14:484–491.
42. Zazopoulos E and Haseltine WA: Disulfide bond formation in the human immunodeficiency virus type 1 Nef protein. J Virol 1993; 67:1676–1680.
43. Blaak H, Brouwer M, Ran LJ, de Wolf F, and Schuitemaker H: In vitro replication kinetics of human immunodeficiency virus type 1 (HIV-1) variants in relation to virus load in long-term survivors of HIV-1 infection. J Infect Dis 1998;177:600–610.
44. Saksena NK, Ge YC, Wang B, Xiang SH, Ziegler J, Palasanthiran P, Bolton W, and Cunningham AL: RNA and DNA sequence analysis of the Nef gene of HIV type 1 strains from the first HIV type 1-infected long-term nonprogressing mother–child pair. AIDS Res Hum Retroviruses 1997;13:729–732.
45. Cao H, Kanki P, Sankale JL, Dieng-Sarr A, Mazzara G, Kalams S, Korber B, Mboup S, and Walker B: Cytotoxic T-lymphocyte cross-reactivity among different human immunodeficiency virus type 1 clades: Implications for vaccine development. J Virol 1997; 71:8615–8623.
46. Bertoletti A, Cham F, McAdam S, Rostron T, Rowland-Jones S, Sabally S, Corrah T, Ariyosha K, and Whittle H: Cytotoxic T cells from human immunodeficiency virus type 2-infected patients frequently cross-react with different human immunodeficiency virus type 1 clades. J Virol 1998;72:2439–2448.
47. Betts MR, Krowka J, Santamaria C, Balsamo K, Gao F, Mulundu G, Luo C, N'Gandu N, Sheppard H, Hahn B, Allen S, and Frelinger JA: Cross-clade human immunodeficiency virus (HIV)-specific cytotoxic T-lymphocyte responses in HIV-infected Zambians. J Virol 1997;71:8908–8911.
48. Schwartz O, Marechal V, Le Gall S, Lemonnier F, and Heard JM: Endocytosis of major histocompatibility complex class I molecules is induced by the HIV-1 Nef protein. Nature Med 1996;2:338–342.
49. Collins KL, Chen BK, Kalams SA, Walker BD, and Baltimore D: HIV-1 Nef protein protects infected primary cells against killing by cytotoxic T lymphocytes. Nature (London) 1998;391:397–401.

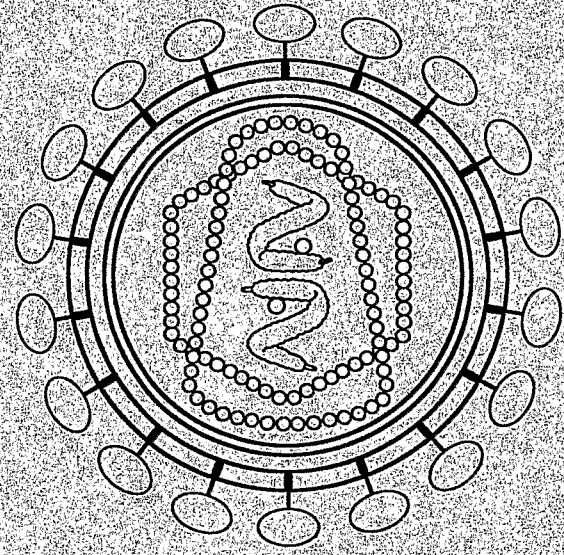
Address reprint requests to:
Martine Peeters
Laboratoire Retrovirus
ORSTOM
911 Avenue Agropolis
BP 5045
34032 Montpellier Cedex 1, France
E-mail: Martine.Peeters@mpl.orstom.fr

Volume 15 Number 1 January 1, 1999

ISSN: 0889-2229

AIDS RESEARCH AND HUMAN RETROVIRUSES

Editor: DANI BOLOGNESI, Ph.D.



Mary Ann Liebert, Inc. publishers