Natural variation of the *nef* gene in human immunodeficiency virus type 2 infections in Portugal

Elizabeth Padua,¹ Adrian Jenkins,² Stuart Brown,² Janet Bootman,² Maria Teresa Paixao,¹ Neil Almond² and Neil Berry²

¹AIDS Reference Laboratory, National Institute of Health, Av. Padre Cruz, 1649-016 Lisbon, Portugal

²Division of Retrovirology, National Institute for Biological Standards and Control, Blanche Lane, South Mimms, Potters Bar, Hertfordshire EN6 30G, UK

Human immunodeficiency virus type 2 (HIV-2) infections cause severe immunodeficiency in humans, although HIV-2 is associated frequently with reduced virulence and pathogenicity compared to HIV-1. Genetic determinants that play a role in HIV pathogenesis are relatively poorly understood but nef has been implicated in inducing a more pathogenic phenotype in vivo. However, relatively little is known about the role of nef in HIV-2 pathogenesis. To address this, the genetic composition of 44 nef alleles from 37 HIV-2-infected individuals in Portugal, encompassing a wide spectrum of disease associations, CD4 counts and virus load, has been assessed. All nef alleles were subtype A, with no evidence of gross deletions, truncations or disruptions in the nef-encoding sequence; all were full-length and intact. HIV-2 long terminal repeat sequences were conserved and also indicated subtype A infections. Detailed analysis of motifs that mediate *nef* function in HIV-1 and simian immunodeficiency virus, such as CD4 downregulation and putative SH2/SH3 interactions, revealed significant natural variation. In particular, the central P¹⁰⁴xxPLR motif exhibited wide interpatient variation, ranging from an HIV-1-like tetra-proline structure (PxxP)₃ to a disrupted minimal core motif ($P^{104}xxQLR$). The $P^{107} \rightarrow Q$ substitution was associated with an asymptomatic phenotype (Fisher's exact test, P = 0.026) and low virus loads. These data indicate that discrete differences in the nef gene sequence rather than gross structural changes are more likely to play a role in HIV-2 pathogenesis mediated via specific functional interactions.

Correspondence Neil Berry nberry@nibsc.ac.uk

Received 17 October 2002 Accepted 27 January 2003

INTRODUCTION

Human immunodeficiency virus type 2 (HIV-2) is the second human retrovirus to be causally linked to AIDS (Clavel *et al.*, 1986). HIV-2 has remained restricted largely to west Africa (reviewed by Schim van der Loeff & Aaby, 1999; Bock & Markovitz, 2001), unlike HIV-1, which has spread rapidly to all parts of the world. HIV-2 has been associated especially with former Portuguese colonies, particularly Guinea Bissau, which reports the highest prevalence of HIV-2 in west Africa: up to 8 % of the population is infected with HIV-2 but is accompanied by relatively low rates of excess mortality (Poulsen *et al.*, 1997; Berry *et al.*, 2002). In Europe, Portugal reports the highest prevalence of HIV-2 infections, with 4 % of notified AIDS cases due solely to HIV-2 (National AIDS Commission, 2002; Soriano *et al.*, 2000).

Differences in the natural history between HIV-1 and HIV-2 are well documented (Marlink *et al.*, 1994; Whittle *et al.*, 1994), although the mechanism(s) that governs the maintenance of an attenuated phenotype for extended

time-periods in a large proportion of HIV-2-infected individuals is not fully understood. Differences in virus subtype may provide one explanation, in particular of HIV-2 variants that resemble simian immunodeficiency virus (SIV_{SM}) infections, which appear apathogenic (Chen *et al.*, 1997; Gao *et al.*, 1994; Wolf *et al.*, 2001; Yamaguchi *et al.*, 2000), but these are few in number and distinct from infections due to the more pathogenic subtypes A and B. Isolation of virus from asymptomatic HIV-2-infected individuals is more difficult than for HIV-1 (Schulz *et al.*, 1990; Simon *et al.*, 1993); many strains are characterized by low levels of cytopathogenicity and reduced virulence. HIV-2 strains, however, exhibit the capacity to utilize a broad spectrum of co-receptors and do not appear to be restricted in their cell tropism (reviewed by Reeves & Doms, 2002).

The failure of HIV-2 to develop into a pandemic and the more attenuated phenotype *in vivo* is correlated with low levels of plasma HIV-2 RNA (Berry *et al.*, 1998, 2002; Andersson *et al.*, 2000; Ariyoshi *et al.*, 2000; Popper *et al.*, 1999) and when associated with high CD4 counts, low

plasma viraemia is predictive of normal survival (Berry *et al.*, 2002). Conversely, advanced HIV-2 infections and disease are associated frequently with high virus load (Ariyoshi *et al.*, 2000; Berry *et al.*, 1998, 2002), with HIV-2 plasma RNA levels representing an independent predictor of disease progression. Hence, host or viral genetic factors that increase virus replication/virus load are likely to have a direct impact on the pathogenesis of HIV-2. Unlike HIV-1 and SIV, the role of *nef* in the differing phenotypes of HIV-2 infections has not been studied as extensively.

Nef, an auxiliary protein that is highly conserved among different primate lentiviruses, regulates virus replication and acts as an immune modulator (Collette, 1997). Genetically and structurally, HIV-2 nef is more similar to SIV_{MAC/SM} nef than to HIV-1 nef and is approximately 260 amino acids in length. In vivo studies of SIV_{MAC} in Asian macaques indicate that an intact *nef* gene is essential for full disease induction and maintenance of high virus load (Kestler et al., 1991). Virus populations attenuated due to the disruption of nef have been demonstrated in both SIV-infected macaques (Whatmore et al., 1995) and long-term survivors with HIV-1 (Deacon et al., 1995; Michael et al., 1995; Geffin et al., 2000; Learmont et al., 1999). However, no consistent relationship between *nef* disruption and prolonged survival in HIV-1-infected individuals has been shown (Carl et al., 2000a, b; Kirchhoff et al., 1999; Mourich et al., 1999), suggesting a complex role for *nef* in the development of human AIDS. Studies of transgenic mice have also indicated a major role for *nef* in inducing a pathogenic phenotype (Hanna et al., 1998, 2001). One previous report based on a mixture of HIV-2 subtypes indicated a higher proportion of truncated nef alleles in asymptomatic HIV-2 infections than expected for HIV-1, an observation that might explain differences in pathogenesis (Switzer et al., 1998) and which, therefore, warrants further investigation.

Nef accomplishes different functions for virus infectivity and virulence (reviewed by Piguet & Trono, 1999; Fackler & Baur, 2002). HIV-1 nef downregulates cell surface CD4 expression and probably MHC class I expression (Greenberg et al., 1997; Mangasarian et al., 1999), connecting the viral receptor and cellular components by adaptor protein (AP)-forming complexes with clathrin-coated pits (Foti et al., 1997), which mediate endocytosis (Piguet et al., 1998). Di-leucine-based motifs in the C-terminal part of nef are important in this process (Craig et al., 1998). In SIV nef, both tyrosine- and leucine-based motifs are involved in CD4 downmodulation (Bresnahan et al., 1999). Although AP interactions vary for different virus-host interactions among different primate lentiviruses (Greenberg et al., 1997; Karn et al., 1998; Lock et al., 1999), combined data from functional studies indicate the same net effect of nef-induced CD4 downregulation. Other cellular partners such as V-ATPase and β -COP (Piguet *et al.*, 1999) may also be involved in these complex processes.

Nef augments virus infectivity *in vivo* by activating target cells through T-cell signalling and signal transduction

pathways (Bell *et al.*, 1998; Simmons *et al.*, 2001; Manninen *et al.*, 1998; Fackler *et al.*, 2001; Fackler & Baur, 2002). Canonical proline-rich (PxxP) regions in *nef* (Saksela *et al.*, 1995) mediate interactions with Src homology region 3 (SH3)-binding domains for Src family kinases such as Hck, Lck and Fyn (Cheng *et al.*, 1999; Lee *et al.*, 1995; Collette *et al.*, 2000). In HIV-2 and SIV *nef*, the core SH3binding domain is usually represented by a minimal binding consensus motif ($P^{104}xxPLR$), although this is typically a tetra-proline structure (PxxP)₃ in HIV-1.

In this study of HIV-2 infections, we describe in detail HIV-2 *nef* sequence variability in Portugal and its relation to putative *nef* function and pathogenesis.

METHODS

Study population. Whole blood was collected between 1994 and 2000 from 37 individuals living in Lisbon, Portugal. These individuals were identified as having HIV-2 infection using a range of serological and genome detection methods and represented a broad cross-section of HIV-2 infections in this country. A wide range of epidemiological and clinical data was collected, including age, gender, ethnic status, route of transmission and country of infection (summarized in Table 1). Of the 37 individuals studied, only seven received anti-retroviral therapy, three (117, 281 and 956) to treat progressive HIV-2 infection and the other four (1227, 1570, 1428 and 1378), who received transient anti-retroviral therapy, to prevent vertical transmission around the time of birth. All four women receiving therapy to prevent HIV transmission were asymptomatic. HIV-2 viral RNA load was only measured subsequently in three women who had received transient drug therapy (1227, 1570 and 1428) and from whom plasma was available.

Amplification of HIV-2 nef and nef/LTR. Peripheral blood mononuclear cells (PBMCs) were isolated by Ficoll-Hypaque density gradient centrifugation. Cellular DNA was prepared by proteinase K treatment, phenol/chloroform purification and ethanol precipitation. DNA resuspended in nuclease-free water was quantified by spectrophotometry. HIV-2 proviral sequences were amplified initially using primers located in the LTR and pol (integrase), as described previously (Berry et al., 1994). nef and nef/U3 LTR regions were amplified using different primer combinations. HIV-2 nef was amplified using outer primers [sense, 5'-GAGCTCGGTACCC-GGGATGGGTGCGAGTGGATCCAA-3' (nt 8503-8538); antisense, 5'-TCTAGAGTCGACCTGCAGGGTATCCCTCTTGCTTTC-3' (nt 9262-9287)], numbering based on HIV-2_{ISY} (GenBank accession no. J04498), and the nef/LTR to the R region was amplified using a nested PCR protocol with outer [sense, 5'-AAGGGC-TATAGGCCWGTDTTCT-3' (nt 8234-8254); antisense, 5'-TGGTGAGAGTCTAGCAGGG-3' (nt 9534-9552)] and inner [sense, 5'-GCCCAACTGCAATATGGGTGCGAGT-3' (nt 8506-8530); antisense, 5'-AACACCCAGGCTCTACCTGCT-3' (nt 9513-9533)] primers.

Amplifications were performed in 100 mM KCl, 20 mM Tris/HCl (pH 9), 0·1 mM EDTA, 1 mM DTT, 0·5 % Tween 20, 50 % (v/v) glycerol, 1·5 mM MgCl₂, 200 μ M of each dNTPs (dATP, dTTP, dCTP and dGTP), 300 nM of each internal primer pair and either 2·5 units of Ampli*Taq* Gold DNA polymerase (Perkin-Elmer) or 2·6 units of *Taq/Pwo* using the Expand High Fidelity PCR system enzyme mixture (Roche). Primary amplifications were performed as follows: 94 °C for 10 min for 1 cycle and 94 °C for 30 s, 53 °C for 45 s and 72 °C for 1 min 30 sec for 35 cycles. Nested reactions with inner primers were 94 °C for 15 s, 50 °C for 45 s, 72 °C for 1 min 30 s for 35 cycles with a final extension step at 72 °C for 7 min. Products were visualized on 1·75 %

Table 1. Epidemiological, immunological and virological characteristics of the 37 HIV-2-infected individuals studied

NA, Not available.

Reference (individual)	Country of infection	Race	Sex	Age	Route of infection	Clinical symptoms	Anti-retroviral therapy	Plasma HIV-2 viral RNA load (copies ml ⁻¹)	CD4 count (cells mm ⁻³)	Additional information	
268	Cape Verde	В	М	55	Heterosexual	Yes	No	200	340	Lung cancer	
423	Guinea Bissau	В	М	47	Heterosexual	Yes	No	< 200	966	Tuberculosis	
379	Portugal	W	М	37	Heterosexual	Yes	No	62 950	366	Tuberculosis	
546	Portugal	W	F	65	Blood transfusion	Yes	No	< 200	297	Infected 1987/Died 2000	
1069	Africa	В	F	67	Blood transfusion	Yes	No	34 750	NA	Infected 1979/Died 1999	
138	Africa	В	М	44	Heterosexual	Yes	No	NA	NA	Tuberculosis	
984	Portugal	W	F	45	Blood transfusion	Yes	No	8 830	195	Infected in 1978	
117	Guinea Bissau	В	М	8	Blood transfusion	Yes	Yes	NA	199	Infected in 1986	
281	Guinea Bissau	В	М	42	Heterosexual	Yes	Yes	NA	46	Tuberculosis	
956	Guinea Bissau	В	F	32	Heterosexual	Yes	Yes	NA	NA	Died in 1999	
280	Guinea Bissau	W	М	48	Heterosexual	Yes	No	NA	NA		
120	Angola	В	F	32	Blood transfusion	No	No	NA	NA		
MP	Portugal	В	F	30	Heterosexual	No	No	< 200	610	Prostitute	
741	Portugal	W	F	27	Heterosexual	No	No	< 200	NA		
1227	Africa	В	F	NA	Heterosexual	No	Yes	< 200	NA	Study of HIV vertical transmission	
1268	Guinea Bissau	В	F	21	Heterosexual	No	No	3 7 3 0	529		
BL	Guinea Bissau	В	М	48	Heterosexual	No	No	< 200	288		
794	Cape Verde	В	F	30	Heterosexual	No	No	< 200	NA		
293	Guinea Bissau	В	М	32	Heterosexual	No	No	NA	519	Diabetic	
483	Africa	В	F	24	Heterosexual	No	No	200	NA		
1139	Portugal	W	F	NA	Heterosexual	No	No	< 200	NA		
1147	Cape Verde	В	М	53	Heterosexual	No	No	2 259	407		
1570	Guinea Bissau	В	F	NA	Heterosexual	No	Yes	< 200	427	Study of HIV vertical transmission	
1215	Portugal	W	F	NA	Heterosexual	No	No	< 200	430		
1320	Africa	В	F	NA	Heterosexual	No	No	200	444		
1544	Cape Verde	В	М	63	Heterosexual	No	No	NA	730		
1428	Portugal	W	F	NA	Heterosexual	No	Yes	< 200	NA	Study of HIV vertical transmission	
292	Portugal	W	F	45	Heterosexual	No	No	< 200	NA		
1096	Portugal	W	М	16	Blood products	No	No	< 200	902	Haemophiliac	
1543	Guinea Bissau	В	F	28	Heterosexual	No	No	< 200	993		
LF	Portugal	W	F	44	Heterosexual	No	No	400	437		
EP	Portugal	W	F	52	Heterosexual	No	No	< 200	490		
1567	Portugal	W	F	26	Heterosexual	No	No	3 4 9 1	1346	IV drug user	
PL	Guinea Bissau	В	М	45	Blood transfusion	No	No	6920	127	Infected in 1973	
511	Portugal	W	F	36	Heterosexual	No	No	500	473		
223	Guinea Bissau	В	F	40	Heterosexual	No	No	NA	161		
1378	Africa	В	F	NA	Heterosexual	No	Yes	NA	NA	Study of HIV vertical transmission	

HIV-2 nef sequence variation in Portugal

agarose gel containing ethidium bromide (0.5 µg ml⁻¹). Aliquots (2–10 µl) containing 0.6 µg extracted genomic DNA were used in primary amplifications and 2 µl of first-round product was used as template for second-round PCR.

Purification of PCR fragments and sequencing. PCR products were purified for sequencing reactions by precipitation with 20% PEG₈₀₀₀ in 2.5 M NaCl. Alternatively, oligonucleotide primers were separated from the PCR product by gel filtration using a Sepharose CL4B column (Pharmacia) under gravity (Almond et al., 1992). Direct sequence analysis of both strands was determined using a modified Sequenase (Amersham) protocol and T7 polymerase (Pharmacia). Terminations were carried out at 50 °C. ABI PRISM Big Dye Terminator kits were used with a cycling profile recommended by the manufacturer. Additional sequencing primers to those used to generate the product were at conserved sites within nef in either the sense orientation [5'-GGGCTACGAGAGAGACTCTT-3' (nt 8561-8580), 5'-GCGAGTGGATCCAAGAAG-3' (nt 8525-8542), 5'-GATAAGGGGGGGGACTGGAAGGGATG-3' (nt 8891-8914), 5'-CAGCAAGGAGACTTTATG-3' (nt 8700-8717) and 5'-TGGCT-ATGGAAGCTAGT-3' (nt 9035-9051)] or the antisense orientation [5'-TGCTTTCAGTTTTGCCTTC-3' (nt 9250-9268), 5'-GTACTAG-CTTCCATAGCC-3' (nt 9036-9053), 5'-CATCATCTGAATCTACAT-3' (nt 8785-8801) and 5'-GTTCTCCATGGGGTATTC-3' (nt 8716-8733)], also numbered according to HIV-2_{ISY}. Sequence contigs were generated using the Gap4 program (Staden software package; Bonfield et al., 1995) from a consensus generated from both sense and antisense strands and analyses were performed using the GCG software package and the CLUSTAL W program to generate multiple sequence alignments (Thompson et al., 1994).

Cloning of *nef* **alleles.** PCR products from a selected number of individuals were cloned into the pGEM-T plasmid vector with a TA Cloning kit (Promega). Ligated DNA was transfected into JM109 cells (Promega) and plasmid DNA purified using QIAprep kits (Qiagen). Clones were sequenced as described above.

Phylogenetic analysis. Nucleotide sequences were aligned and evaluated with CLUSTAL W and phylogenetic trees were constructed from the matrix built by the neighbour-joining method (NEIGHBOR, version 3.5c; Felsenstein, 1989). Nucleotide substitution rates per site were estimated for each pairwise sequence comparison on the basis of the Kimura two-parameter model using DNADIST. Bootstrap analysis was performed with SEQBOOT and CONSENSE with 100 re-samplings. Trees were rooted with SIV_{MM239} and drawn with NJPLOT; reference sequences were obtained from the Los Alamos database. Nucleotide sequences reported in this study have been submitted to the EMBL nucleotide sequence database (accession nos AJ344369–AJ344415). Epidemiological and clinical aspects were evaluated statistically with SPSS (statistical package for social science).

Quantification of HIV-2 RNA levels in plasma. EDTA-treated plasma samples were stored at -80 °C until levels of HIV-2 viral RNA were assessed using a quantitative RT-PCR assay (Berry *et al.*, 1998). In brief, RNA was extracted according to the method of Boom *et al.* (1990) and RT-PCR performed, in duplicate, with HIV-2-specific LTR oligonucleotide sequences in single-tube assays using Titan reagents (Roche). Amplification products were quantified by chemiluminescence using an internal HIV-2 probe linked to alkaline phosphatase (Oswel). The reliable level of sensitivity of the assay was 200 copies ml⁻¹ using an input plasma volume of 100 µl. For statistical analyses, plasmas with <200 copies ml⁻¹ were assigned an arbitrary value of 20 copies ml⁻¹.

RESULTS

Epidemiological, clinical and virological characteristics of the study population

Of the 37 HIV-2-infected individuals studied (Table 1), 24 (64.9%) were female and 13 (35.1%) were male, with a mean overall age of 39.4 years (n=37, range 8–67 years). The main route of transmission was heterosexual (30/37, 81.2%), the remaining seven being transfusion-associated infections. HIV-2 infection was acquired in Portugal in 14 (37.8%) individuals; the other 23 (62.2%) infections were most likely acquired in Africa: Guinea Bissau (n=12, n=12)32.4%), Cape Verde (n=4, 10.8%), Angola (n=1, 2.7%) or an unidentified African country (n=6, 16.2%). Following clinical examination, 11/37 (29.7%) were symptomatic (HIV-related illness/AIDS) and 26 (70.3%) were asymptomatic, including three individuals (1069, PL and 280) who had been infected for 20 years or more. CD4 lymphocyte counts and HIV-2 viral RNA loads were available for the majority of, but not all, individuals. In the asymptomatic group, mean CD4 counts were 548 (n=17, range 161-1346 cells mm⁻³) and mean HIV-2 RNA levels were 817 copies ml⁻¹ (n=21, range < 200–6920). In the symptomatic group, mean CD4 counts were correspondingly lower at 328 (n=8, range 46–966 cells mm⁻³) and the mean HIV-2 viral RNA load was 20 times higher than that in the asymptomatic group at 17 051 copies ml^{-1} (n=7, range < 200-62 950).

HIV-2 *nef* sequences in Portugal are subtype A and full-length

A total of 44 nef/LTR proviral sequences spanning the entire nef gene were amplified from 37 HIV-2-infected individuals (792 bp, Fig. 1). Consensus sequences derived from each of the 37 individuals from bulk product were substantially full-length with little evidence of deletions, truncations or other disruptions to nef. Predicted amino acid sequences (see below) suggested a functional HIV-2 Nef protein to be present in all individuals studied, irrespective of any other factor. Subsequent bleeds from six individuals yielded a nef sequence with few nucleotide or amino acid differences from the first sequence; hence, comparisons of nef variation were made on the first, bulk amplification of each patient. This observation was confirmed further by dilution experiments of template DNA. Phylogenetic relationships between Portuguese nef sequences were compared with HIV-2 subtype A (HIV-2_{ROD}, HIV-2_{BEN}, HIV-2_{ALI}, HIV-2_{ISY} and HIV-2_{ST}), subtype B (HIV-2_{D205/ALT}, HIV-2_{UC1} and HIV-2_{EHO}) and SIV_{MM239} (Fig. 1). All Portuguese sequences were HIV-2 subtype A, with no other HIV-2/SIV_{SM} subtypes (B-G) identified. No phylogenetic relationship was established between nef sequence and disease status and no statistical link between HIV-2 nef sequence and route of transmission, either sexual or blood transfusion, was apparent (data not shown).





Fig. 1. Phylogenetic tree based on 44 nef/LTR sequences derived from proviral DNA fragments amplified by PCR from PBMCs of 37 individuals, five reference HIV-2 subtype A sequences (HIV-2_{BEN}, HIV-2_{ALI}, HIV-2_{ISY}, HIV-2_{ROD} and HIV- 2_{ST}) and three reference sequences for subtype B (HIV- 2_{EHO} , HIV-2_{ALT/D205} and HIV-2_{UC1}). SIV_{MM239} was used as the outgroup sequence. A second sequence obtained from each of six individuals is indicated (1139-1/1139-2, 1147-1/1147-2, LF-1/LF-2, MP-1/MP-2, 281-1/281-2 and 117-1/117-2). Prefixes of 'a' and 's' indicate asymptomatic and symptomatic, respectively. Samples suffixed 'dil' indicate that the sequence was derived from a 1/10 dilution of the original DNA material. Bootstrap values above 80%, based on 100 replicates, are indicated at the branching nodes. Evolutionary distances were estimated using the Kimura two-parameter model and genetic relationships were determined using the neighbour-joining method. Bar, 10% sequence divergence.

Analysis of LTR sequences

Sequence data for the non-overlapping U3 region of the LTR was available for 31 individuals, indicating no disruptions to the structure of the LTR. Comparisons of the core promoter and enhancer elements important in the transcriptional activity of the LTR, including PuB-1, PuB-2 p-*ets*, peri- κ B and NF- κ B sites up to the Sp1 sites and TATA box, indicated further sequences to be subtype A with high levels of sequence conservation *in vivo*. Sequences are not shown but have been filed in GenBank.

HIV-2 nef amino acid variation

Alignment of predicted amino acid sequences for the entire nef ORF for all 37 Portuguese HIV-2-infected individuals is shown in Fig. 2; the first sequence obtained for each individual is presented. The Portuguese HIV-2 nef consensus sequence was also compared with other lentivirus nef sequences, including a prototypic HIV-1 group M virus (HIV-1_{LAI}), SIV_{CPZ} (Gab), from chimpanzees, and SIV_{MM239}, an experimental infection of Asian macaques derived originally from SIV_{SM}. Predicted coding sequences were compared for 10 putative regions considered important for HIV-1, HIV-2 or SIV nef function (Fig. 2). Throughout nef, a number of key regions were conserved among the Portuguese sequences, irrespective of clinical status, virus load or CD4 count, and deletions that might disrupt such sites were not evident. Moreover, only one patient showed evidence of a deleted *nef* sequence: patient 379 (symptomatic/AIDS) exhibited a nef deletion of 8 aa downstream of a suggested N-terminal RNA-binding motif (Echarri *et al.*, 1997).

The myristoylation signal (aa 1–6, HIV-2 consensus GASGSK) was 100 % conserved and GPG(V/T) sequences predictive of a β -turn were also conserved, with a T present in one AIDS patient (281) and in six asymptomatic individuals. Comparative analyses of other regions identified as being relevant to *nef* function exhibited a more complex pattern of variation, particularly the SH3-binding PxxP motif, the di-leucine-based region important for downregulation of CD4 (in HIV-1 the ENTSLL motif) and sites for putative SH2 interactions based on tyrosine-sorting signals.

(i) Analysis of HIV-2 nef for putative tyrosine-based sorting motifs. The canonical tyrosine-based sorting motif $Yxx\phi$, present as Y^{28} GRL and Y^{39} SQS in SIV_{MAC239} and Y³⁹SRF in HIV-2_{ROD}, is Y³⁹LQS in these Portuguese sequences, with changes to this basic configuration infrequent. Patient 984, an AIDS patient with a CD4 count of 195 and viral RNA levels of 8830 copies ml^{-1} , exhibited a $R \rightarrow Y$ mutation at position 24; patients LF and PL exhibited $C \rightarrow Y$ mutations at position 28. However, additional tyrosine-based motifs in HIV-2, such as YE nef described for SIV_{MAC239}, were not present in any HIV-2 sequences, including those in the symptomatic group. Interestingly, the most conserved tyrosine residue at position 39 was mutated in two AIDS patients (117-1 and 1378) to phenylalanine, serine or cysteine. There was no apparent relationship between disease status and tyrosine-based motifs.

(ii) Di-leucine-based motifs for CD4 downregulation. Cellular di-leucine-based sorting pathways required for CD4 downregulation in HIV-1 are characterized by a core hexameric motif ENTSLL (ExxxLL), which conforms to the consensus E/DXXXL ϕ found in cellular transmembrane proteins (Craig *et al.*, 1998). HIV-2 *nef* sequences analysed for a comparable motif (L¹⁹⁸L¹⁹⁹) indicated variation at only L¹⁹⁹, present as leucine, valine or

	1 2	3		4 5	6 7
HIV-1 LAI	MCGKWSKSSVIGWPTVRERMRAEP	AADRVGAASRDLEKHGAITSSNTAA	TNAAC***********************************	********AWLE AQ-EEE-EVG I	FPVTPQVPLRPMTYKAAVDLSHFLKE KGGLEGLIHSQRRQDILDLWIYHTQ
SD MM 239	MGGAISMRRSRPSGDLRORLLRARG	ETYGRLLGEVEDGYSOSPGGLDKGL	SSLSCEGOKYNOGOYMNTPWRNPAE	EREKLAYRKONMDDIDESDDDLVGV	SVRPKVPLRTMSYKLAIDMSHFIKE KGGLEGIYYSARRHRILDIYLEKEE
HIV-2ROD	MGASGSKKHSRPPRGLQERLLRARA	GACGGYWNESGGEYSRFQEGSDREQ	KSPSCEGRQYQQGDFMNTPWKDPAA	EREKNLYRQQNMDDVDSDDDDQVRV	SVTPKVPLRPMTHRLAIDMSHLIKT RGGLEGMFYSERRHKILNIYLEKEE
PORT-CON	MGASGSKKRSRPSRGLQERLLRARG	GTCGGHWDESGGGYLQSQEGSGREQ	-SPSCEGQRYQQGDFMNTPWRTPAA	EGEKNLYRQQNMDDVDSDDDDIVGV	PVTPRVPLRAMTYRLAIDMSHRIKE KGGLEGMFYSERRHRILDIYLEKEE
280	KSD	-PRNEG	/ 5 NNG	-EKENO	SKK
223	RQ-H-	SQ-GEG	RHISPNT	-EKSQ	SK
281-1	RSLQ	ERCNEH	RT	-REKLQF	KLIHKL
MP-1	RPSLQE-RFA-	ERCNK-EFHD	RGV	-RE-HKL-N-QF	KPKMLLQELQEL
956	EQR-R	ECSQESHGD-G-	NA	ER-KS	PK
741	A	E-SGYKDLEE	RYRKD-T-	KE-SF	VQERRDQNQN
268	PA	E-S-EG-R-LE-SRE	R-RKDP RKD-S-	VRKE-S	SERLD
1570	A	-AYLE	KLT-	R-A-KQ	SDKDK
1215	A	-AY-N-LE-S	KQ	KAE-Q	SDKDK
1544	н-коа	-AV-SR	KN-T	-KIE	H-KKN
483	A	Y-E-L-EEG-	KKAI	RSA-KQ	HLD
293	A	-AG-	KET	R-A-KQ	K
1139-1	A	-AEQLEE-D-G-	ККГ	RKA-K0	S-R-K-Q
BL	A	-ALK-SR	КН	s	SKTEDD
794	QQA	EY-N-LE	RD	S	ТQЕD
423	A	YIO-S-FE-D	ККОГ	PTNKN	RDNN
1428	A	-AG-LE-E-S-F-GE	КТ	Q	SELL
292	A	Y-KQ-SRYE-D	RBNT	Q	SSKKKK
984	Y-	EQHGG	NRR	-KEN	SL
1543	LQ	QCSG-EK	N	-K-NQKQ	QQKQFK-
LF-1	K-LQRQ	E-YRCGG-DFH-E	N	-REM-KNQ	SKLD
138 EP	Q RQ	ER-C-G-E-RRF	NLA-T NIN-T	DERS	HKKK
1268	KE	EY-S-QG-	RKVAT	-R-NKQ-K	SD
1567	RT-E	GEGEGEGEGEGEGEGEGE	S	-REK	QЕQ_
PL 546	R	E-YRCNG-E-ECECE	K-LKD-TS	-QKQ	S-VQIKDCK
511	KR	-AR-CNG-E-E-S-FLGE-DKG-	NQQ	-KAK	PKMI-HE-Q
1069	Q	RSNE-EGE	NG	-R-RAM-KAE	RTK
117-1	H-KQ-R	ES-G-AERCEG-	NG	.R-GEKC	S-PML FMN
10/0	1 11 210	REAL REAL POPUL OF		Dit 11 11	
		8 9	10		
HIV-1 LAI	GYFPDSQNYTPGPGVRYPLTFGWCY	8 9 KLVPVEPDKIEEANKG**ENTSLLH	10 PVSLHGMDDPE****REVLEWRFDS	RLAFHHVARELHPEYF****KNC\$	
HIV-1 LAI SIVCPZGAB	GYFPDSQNYTPGPGVRYPLTFGWCY GFFPDWQNYTTGPGTRFPLCFGWCF	8 9 KLVPVEPDKIEBANKG**ENTSLH KLVPLTEEQVEOANEG**DNNCLLH	10 PVSLHGMDDPE****REVLEWRFDS PICQHGMEDED****KEVLHWQFDR	RLAFHHVARELHPEYF****KNC\$ RLALRHIAREQHPEYY****KD\$	
HIV-1 LAI SIVCPZGAB SD MM 239 HIV-2ROD	GYFPDSQNYTPGPGVRYPLTFGWCY GFFPDWQNYTTGPGTRFPLCFGWCF GIIPDWQDYTSGPGIRYPKTFGWLW GIIADWONYTHGPGURYPMFFGWLW	8 9 KLVPVEPDKIEBANKG**ENTSILH KLVPUTTEEQVEQANEG**DNNCLLH KLVPVNVSDEAQED***EEHYLMH	10 PVSLHGMDDPE****REVLEWRFDS PICQHGMEDED****KEVLHWQFDP PAQTSQWDPWGEVLAWKFDPTLAY PAOTSKFDDPHGETLVWFEPDFLAY	RLAFHHVARELHPEYF****KNC\$ RLALHHIAREQHPEYY****KNS TYEAYVRYPEFGSKSGLSEEVRR SVEAFTRYPFFGUKSGLSEEFWRA	RLTARGLINMADKKET
HIV-1 LAI SIVCPZGAB SD MM 239 HIV-2ROD PORT-CON	GYFPDSQNYTPGPGVRYPLTFGWCY GFPFDWQNYTTGPGTRPPLCFGWCF GIIPDWQDYTSGPGIRYPKTFGWLW GIIADWQNYTHGPGVRYPMFFGNLW GIIPDWQNYTHGPGIRYPKFFGULW	89 KLVPVEPDKIEBANKG**ENTSLH KLVPLTEEQVEQANEG**DNNCLH KLVPVNVSDEAQED****EEHYLMH KLVPVDVVQ***EGEDTETHCL*H	10 PVSLHGMDDPE****REVLEWRFDS PICQHGMEDED****KEVLHWQFDR PAQTSKFDDPMGEVLAWKFDPTLAY PAQTSKFDDPHGETLVWRFDPLLAY PAQTSKFDDPHGETLVWRFDPMLAY	RLAFHHVARELHPEYF****KNC\$ RLALRHIAREQHPEYY****KD\$ TYEAYVRYPEEFGKSGLSEEEVRR SYEAFIRYPEEFGHKSGLPEEEWKA DYTAFVRYPEFGHKSGLPEEEWKA	RLTARGLIAMADKKET RLKARGIPFS RLKARGIPFSSTLR
HIV-1 LAI SIVCPZGAB SD MM 239 HIV-2ROD PORT-CON	GYPPDSQNYTPGPGVRYPLTFGWCY GFPPDWQNYTTGPGTRPLCFGWCP GIIPDWQDYTSGPGIRYPKTFGWLW GIIPDWQNYTHGPGVRYPMFFGWLW GIIPDWQNYTHGPGIRYPKFFGWLW 175	8 9 KLVPVEPDLEANKG**ENTSLLH KLVPLTEEQVEOANEG**DNNCLLH KLVPVNVSDEAQED****EEHYLMH KLVPVD VPQ***EGEDTETHCUH KLVPVDTVPQ***EGEDTETHCUH KLVPVDTVPQ***EGEDTETHCUH	10 PVSLGMDDPE****REVLEWRFDS PUCQHGMDSDD****KEVLHWQFDR PAQTSQWDDPWGEVLAWKFDPTLAY PAQTSKFDDPHGETLWWFDPLLAY PAQTSKFDDPHGETLWWFDPLLAY PAQTSKFDDPHGETLWWFDPLLAY 225	RLAFHHVARELHPEYF****KNC\$ RLALRHIAREQHPEYY****KD\$ TYEAYVRYPEEFGSKSGLSEEVRR SYEAFIRYPEEFGHKSGLPEEBWKA DYTAFVRYPEEFGHKSGLPEEBWKA 250	RLTARGLLNMADKKET RLKARGIPFSSTLR Z64
HIV-1 LAI SIVCPZGAB SD MM 239 HIV-2ROD PORT-CON 280 223	GYPPDSQNYTPGPGVRYPLTFGWCY GFPPDWQNYTTGPGTRPLCFGWCP GIIPDWQDYTSGPGIRYPKTFGWLW GIIPDWQNYTHGPGIRYPKFGWLW GIIPDWQNYTHGPGIRYPKFGWLW 175 LYVT	8 8 8 8 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9	10 PVSLHGM PE**** REVLEWRFDS PICOHGMUDD*** KEVLHWQFDR PAQTSQWDDWGEVLAWKFDPTLAY PAQTSRFDPHGETLUWRFDPLLAY PAQTSRFDPHGETLUWRFDPLLAY PAQTSRFDPHGETLWRFDPMLAY 225 	RLAFHHVARELHPEYF****KNC\$ RLALHHIAREQHPEYY****KD\$ TYEAYURYPEFGSKSGLSEEVRR SYEAPIRYPEFGHKSGLPEEWKA DYTAFVRYPEEFGHKSGLPEEBWKA 250 KQLYQA-K XE-TYM	RLTARGLIAMADKKET RLKARGIPFS RLKARGIPFSSTLR 264 KCR
HIV-1 LAI SIVCPZGAB SD MM 239 HIV-2ROD PORT-CON 280 223 281-1	GYPPDSQNYTPGPGVRYPL/TFGWCY GFPPDWQNYTTGPOTRPL/CFGWCP GIIPDWQDYT5GPOTRPL/CFGWCW GIIPDWQNYTHGPORYPWPFGWLW GIIPDWQNYTHGPGIRYPKFFGWLW -LYVT	8 8 9 9 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	10 PVSLHGMDPF****REVLEWRFDS PICQHGMBDPGE****REVLHWQFDR PAQTSQMDDPMGEVLAWRFDPTLAY PAQTSRFDDPGETLVWRFDPLLAY PAQTSRFDDPHGETLVWRFDPLLAY 225	RLAFHHVARELHPEYF****KNC\$ RLALHHIAREQHPEYY****KD\$ TYEAVYNYPEEFGKSGLSEEVFRA DYTAFVTYPEEFGHKSGLPEEFWRA DYTAFVTYPEEFGHKSGLPEEFWRA 250 KQLYQA-K K-ETYQA-K NK	RLTARGLLNMADKKET RLKARGIPFS RLKARGIPFSSTLR 264 KCR KCR
HIV-1 LAI SIVCPZGAB SD MM 239 HIV-2ROD PORT-CON 280 223 281-1 120	GYPPDSQNYTPGPGVRYPLTFGWCY GFPPDWQNYTTGPGTRPLCFGWCF GIIPDWQDYTGGPGIRYPKTFGWLW GIILDWQNYTHGPGIRYPKFFGWLW 175 LYVT AYW A	8 9 KLUPPUEPDKI E ANKG**ENCLH KLUPPUNESDAQE KLUPVDUSDAQE KLUPVDUVSSAQE KLUPVDTVPC**ECENTETHC.H 200 LPQREN-L- VSYM-L VSY-M-L CRL-N-AN-L	10 PVSLHGMDPE****REVLEWRPDS PVSLHGMDPE****REVLEWRPDS PAQTSQNDDPMGEVLAWRPDPLLAY PAQTSRPDPHGETLWWRPDPLLAY PAQTSRPDPHGETLWWRPDPLLAY 225	RLAFHHVARELHPEYF****KNC\$ RLALHHIAREQHPEYY****KD\$ TYEAYURYPEEFGKSGLSEEEVKR SYEAPIRYPEEFGHKSGLPEEEWKA DTAFVRYPEEFGHKSGLPEEEWKA 250 KQLYQA-K K-ETYQA-K NK	RLTARGLINMADKKET RLKARGIPFS RLKARGIPFSSTLR 264 KCR KCR CR CR
HIV-1 LAI SIVCPZGAB SD MM 239 PORT-CON 280 223 281-1 120 MP-1 956	GYPEDSQNYTPGPGVRYELTFGWCY GPPPDWQNYTTGPGTRFPLCPGWCF GIIPDWQDYTSGPGIRYEKTFGWLW GIIPDWQNYTHGPGIRYEKFFGWLW GIIPDWQNYTHGPGIRYEKFFGWLW 	8 8 8 KLUPVEPDKI BANKG**ENS H KLUPVDTEEQVMANEG**DANC H KLUPVDVERAO ***EGT FEHCL VH KLUPVDVPV**EGT FEHCL VH KLUPVDTVPC**EGT FEHCL VH 200 LPQREN-AN-L Y-M Y-M 	10 PV5LHGMD7P5****REVLEWRP5G PICQHGM8D20****REVLEWQFDR PAQTSQMD7WGEVLAWRFD7PILAY PAQTSRPD0FGETLWRFD7PLAY PAQTSRPD0PHGETLWRFD7PLAY 225	RLAFHHVARELHPEYF****KNC\$ RLALHHIAREQHPEYY****KD\$ TYBAYURYPEFOSKSGLSEEVKR DYTAFURYPEFOSKSGLPEEWKA DYTAFURYPEFOSKSGLPEEWKA DYTAFURYPEFOSKSGLPEEWKA K.ETYM NYK NYK NNK NNK NNK NNK NNK NNK NNK NNNN	RLTARGLINMADKKET RLKARGIPFSSTLR 264 KCR CR CR CR CR
HIV-1 LAI SIVCPZGAB SD MM 239 HIV-2ROD PORT-CON 223 281-1 120 MP-1 956 741	GYPPDSQNYTPGPGVRYPL/TFGWCY GFPPDWQNYTTGPGTRPL/CFGWCP GIIPDWQDYTGGPGIRYPKFFGWLW GIIPDWQNYTHGPCNIPWFFGWLW GIIPDWQNYTHFGWLW GIIPDWQNYTHGPCNIPWFFGWLW GIIPDWQNYTHGPCNIPWFFGWLW GIIPDWQNYTHGPCNIPWFFGWLW GIIPDWQNYTHGPCNIPWFFGWLW GIIPDWQNYTHGPCNIPWFFGWLW GIIPDWQNYTHGPCNIPWFFGWLW GIIPDWQNYTHGPCNIPWFFGWLW GIIPDWQNYTHGPCNIPWFFGWLW GIIPDWQNYTHGPCNIPWFFGWLW GIIPDWQNYTHFGWLW GIIPDWQNYTHFGWLW	8 9 KLVPVEDEBANKG**ENNCLH KLVPVUNSDEAO KLVPVUNSDEAO ***EENY EM KLVPVUNSDEAO ***EENY EM KLVPUNSDEAO ***EENY EM ***EENY EM ***EENY EM ****EENY EM ***EENY EM ************************************	10 PV5LHGMDTPF****REVLEWRPDS PICQHGM2050 P1CQHGM2050 P1CQHGM2050 PAQTSQMDTPMGSTLJWRFDPLAY PAQTSRFDDTPHCSTLJWRFDPLAY PAQTSRFDDTPHCSTLJWRFDPLAY PAQTSRFDDTPHCSTLJWRFDPLAY PAQTSRFDDTPHCSTLJWRFDPLAY PAQTSRFDDTPHCSTLJWRFDPLAY PAQTSRFDTPHCSTLJWRFDPLAY PAQTSRFDTPHCSTLJWRFDTPHCSTLJWRFDPLAY PAQTSRFDTPHCSTLJWRFDTPHCSTLJWRFDPLAY PAQTSRFDTPHCSTLJWRFDTJW	RLAFHHVARELHPEYP****KNC\$ RLALRHIAREQHPEYY****KD\$ TYEAVVRYPEFOSKSGLSEEVRA DYTAFVRYPEEFOHKSGLPEEWKA DYTAFVRYPEEFOHKSGLPEEWKA DYTAFVRYPEEFOHKSGLPEEWKA STATUTAFVRYPEFOHKSGLPEEWKA STATUTAFVRYPEFOHKSGLPEEWKA STATUTAFVRYPEFOHKSGLPEEWKA STATUTAFVRYPEFOHKSGLPEEWKA STATUTAFVRYPEFOHKSGLPEEWKA STATUTAFVRYPEFOHKSGLPEEWKA STATUTAFVRYPEFOHKSGLPEEWKA STATUTAFVRYPEFOHKSGLPEEWKA STATUTAFVRYPEFOHKSGLPEEWKA STATUTAFVRYPEFOHKSGLPEEWKA STATUTAFVRYPEFOHKSGLPEEWKA STATUTAFVRYPEFOHKS	RLTARGLLAMADKKET RLKARGIPFSSTLR 264 KCR KCR
HIV-1 LAI SIVCPZGAB SD MM 239 HIV-2ROD PORT-CON 223 281-1 120 MP-1 956 741 1227	GYPPDSQNYTPGPGVRYPL/TFGWCY GFPPDWQNYTTGPGTRPL/CFGWCP GIIPDWQNYTGGPGTRPFKTFGWLW GIIPDWQNYTHGPGVRYPWFFGWLW GIIPDWQNYTHGPGIRYPKFFGWLW 175 LYVT	8 9 KLUPPUTEEQVGANEG**DNNCLIH KLUPVUNSDEAQ ***EGENTETHC.VH KLUPVUNSDEAQ ***EGENTETHC.VH KLUPVUNSDEAQ ***EGENTETHC.VH KLUPVUNSDEAQ ***EGENTETHC.VH KLUPVONY SCONTENCE 200	10 PUSLHGMDPE****REVLEWRFDS PICQHGMBDPGE****REVLEWRFDS PAQTSQMDDPMGEVLAWRFDPTLAY PAQTSRFDDPGFTLVWRFDPLLAY PAQTSRFDDPGFTLVWRFDPLLAY 225	RLAFHHVARELHPEYF****KNC\$ RLALHHIAREQHPEYY****KD\$ TYEAVVRYPEEFGKSGLSEEEVRA DYTAFVVRYPEEFGHKSGLPEEPKA DYTAFVRYPEEFGHKSGLPEEPKA CSU KQLYQA-K K-E-TYQA-K N	RLTARGLLNMADKKET RLKARGIPFS RLKARGIPFSSTLR 264 KCR KCR CR
HIV-1 LAI SIVCPZGAB SD MM 239 HIV-2ROD PORT-CON 223 281-1 120 MP-1 956 741 1227 268 1570	GYPPDSQNYTPGPGVRYPLJFGMCY GPPPDWQNYTTGPGTRFPLCFGWCF GIIPDWQDYTSGPGIRYPKFFGWLW GIIPDWQDYTGGPGIRYPKFFGWLW GIIPDWQNYTHGPGIRYPKFFGWLW 175 L	8 8 9 KLUPVEDEXL BANKG*+EYS H KLUPVEDEQUE ANEG**DANC H KLUPVUNSDRAG KLUPVUNSTRAG 	10 PICULAND DES	RLAFHHVARELHPEYF****KNCS RLALRHIAREQHPEYY****KDS TYBAYUNYPEFOSKSGLSEEVRA DYTAFUNYPEFOSKSGLSEEVRA DYTAFUNYPEFOSKSGLPEEVRA DYTAFUNYPEFOSKSGLPEEVRA K.ETYM NYK N SKH SKH NK N	RLTARGLLNMADKKET RLKARGIPFSSTLR 264 KCR
HIV-1 LAI SIVCP2GAB SD MM 239 HIV-2ROD PORT-CON 223 281-1 120 MP-1 956 741 1227 268 1570 1215	GYPEDSQNYTPGPGVRYPLJFGWCY GPPPDWQNYTGGPGTRPLCFGWCP GIIPDWQDYTGGPGTRPPKTFGWLW GIIPDWQNYTHGPGIRYPRFGWLW GIIPDWQNYTHGPGIRYPKFGWLW 	8 9 KLUPPUTEEQVMANEG**DINCLIH KLUPPUTEEQVMANEG**DINCLIH KLUPPUTEEQVMANEG**DINCLIH KLUPPUTP**EGITEHCLIH KLUPUTVP**EGITEHCLIH COO LPOREN-AN-LI CRE-ED-AN-LI	10 PUSLHGMDTPE****REVLEWRFDS PTCQHGMBDED****REVLEWRFDS PAQTSQMDDWGEVLAWRFDPTLAY PAQTSRFDDTPELAY PAQTSRFDDTPELTUWRFDPLAY PAQTSRFDDTPETUWRFDPLAY 225	RLAFHHVARELHPEYF****KNC\$ RLALRHIAREQHPEYY****KD\$ TYBAYURYPEFGSKSGLSEEVKR DYTAFURYPEFGHKSGLPEEWKA DYTAFURYPEFGHKSGLPEEWKA K.ETYM NYK NNK SKHNK SKHNK SKHNK VNLYKR VNLYKR VNLYKR	RLTARGLLNMADKKET RLKARGIPFSSTLR 264 KCR CR CR CR
HIV-1 LAI SIVCP2GAB SD MM 239 HIV-2ROD PORT-CON 223 281-1 120 MP-1 1227 268 1570 12215 1320 1225	GYPPDSQNYTPGPGVRYPL/TFGWCY GFPPDWQNYTTGPGTRPL/CFGWCP GIIPDWQDYTGPGIRYPKTFGWLW GIIPDWQNYTHGPCNIPWPFFGWLW GIIPDWQNYTHGPCNIPWPFFGWLW GIIPDWQNYTHGPCNIPWPFGWLW L -Y -Y -A T -A T A M A M A M A M A M A M YWM	8 9 KLUPPUTEEQVE MANEG** DANKC.LH KLUPVUNSDEAQ KLUPVUNSDEAQ MAREG** DANKC.LH KLUPVUNSDEAQ KLUPVUNSDEAQ MARG** EGSTFETHCL.VH KLUPVUNSDEAQ R SOO	10 PUSLHGMDTPF****REVLEWRPDS PICQHGM2020 PAQTSQMDTPMGETU.WRFDPLAY PAQTSRPDTMGETUWRFDPLAY PAQTSRPLAY	RLAFHHVARELHPEYP****KNC\$ RLALRHIAREQHPEYY****KD\$ TYEAVVRYPEFGSKSGLSEEVRA DYTAFVVRYPEEFGHKSGLPEEPKA DYTAFVRYPEEFGHKSGLPEEPKA SPAPITAPEFGHKSGLPEEPKA K-ETYM	RLTARGLLAMADKKET RLKARGIPFSSTLR 264 KCR KCR
HIV-1 LAI SIVCPZGAB SD MM 239 HIV-2ROD PORT-CON 223 281-1 120 MP-1 956 5741 1227 268 1570 1215 1320 1544 483	GYPPDSQNYTPGPGVRYPLJFGWCY GPPPDWQNYTTGPGIRPPLCFGWCF GIIPDWQDYTSGPGIRYPKFFGWLW GIIPDWQNYTHGPGIRYPKFFGWLW GIIPDWQNYTHGPGIRYPKFFGWLW L	8 8 9 KLUPYEDEQUEQANEG**DANC KLUPYEDEQUEQANEG**DANC KLUPYUNSDRAQ KLUPYUNSDRAQ ***CetherHC.WH KLUPYUDYUPY**ECtHERHC.WH KLUPYUDYUPY**ECtHERHC.WH KLUPYUDYUPY**ECtHERHC.WH KLUPYUDYUPY**ECTHER 200 	10 PICQHOMS DED**** REVLEWREPS PICQHOMS DED**** KEVLEWREPS PAQTSKOPHOSELLWREPDFLAY PAQTSKPDPHOSELWREPDFLAY PAQTSKPLANE PAQTSKPLANE PAQTSKPLANE PAQTSKPLANE PAQTSKPLANE PAQTSKPLANE PAQTSKPLANE	RLAFHHVARELHPEYF****KNCS RLALRHIAREQHPEYY****KDS TYBAYUNYPEFOSKSGLSEEVRA SYBAFIRYPEFOSKSGLSEEVRA DYTAFUNYPEFOSKSGLPEEVRA LSS KQLYM	RLTARGLLNMADKKET RLKARGIPFSSTLR 264 KCR
HIV-1 LAI SIVCPZGAB SD MM 239 HIV-2ROD PORT-CON 220 223 281-1 120 MP-1 956 741 1227 268 1570 1215 1520 1220 1244 483 293	GYPEDSQNYTPGPGVRYELTFGWCY GPPPDWQNYTTGPGTRFPLCFGWCF GIIPDWQDYTSGPGIRYEKFFGWLW GIIPDWQNYTHGPGIRYEKFFGWLW GIIPDWQNYTHGPGIRYEKFFGWLW 	8 9 KLUPPUTEEQVM_MARGG**DANC *** KLUPPUTEEQVM_MARGG**DANC *** KLUPPUTEEQVM_MARGG**DANC *** KLUPPUTPUTPUTEQ***ECL *** 200 200 LPQ	10 PICQHOM DEPE**** REVLEMREPS PICQHOM DEVES*** REVLEMREPS PAQTSK0MDEWGEVLAMREPDFLAY PAQTSREPDEPHGETLUWREPDFLAY 225	RLAFHHVARELHPEYF****KNC\$ RLALHHIAREQHPEYY***KD\$ TYBAYUNYPEFOSKSGLSEEVKR DYTAFUNYPEFOSKSGLSEEVKR DYTAFUNYPEFOSKSGLPEEWKA NKQLYQA-K 250 KYQX-K NYQK N	RLTARGLLNMADKKET RLKARGIPFSSTLR 264 KCR CR CR R * N* N* KE* KE* KE* KN*
HIV-1 LAI SIVCPZGAB SD MM 239 HIV-2ROD PORT-CON 223 281-1 956 741 1227 268 1570 1215 1320 1544 483 293 1147-1	GYPPDSQNYTPGPGVRYPL/FGWCY GFPPDWQNYTTGPGVRYPL/FGWCW GIIDWQDYTGPGVRYPFFGULW GIIDWQNYTHGPGVRYPFFGWLW GIIDWQNYTHGPGVRYPFFGWLW GIIDWQNYTHGPGVRYPFFGWLW LYVT	8 9 KLVPVEDEQVENARGG**DANKG**DANKCI.H KLVPVUNSDEAQ ***EENTPIKK KLVPVUNSDEAQ ***EGSTFEHKCI.H KLVPVUNSDEAQ ***EGSTFEHKCI.H KLVPVUNSDEAQ ***EGSTFEHKCI.H KLVPVDVPV**EGSTFEHKCI.H ***EGSTFEHKCI.H	10 PVSLHAMDTPF****REVLEWRPDS PICQHAMEDED****REVLEWRPDS PAQTSQMDDPMGEVLAWREDPTLAY PAQTSRPDPHGETLUWRPDPLAY PAQTSRPDPMGETUWRPDPLAY 225	RLAFHHVARELHPEYP****KD\$ RLALRHIAREQHPEYY****KD\$ TYEAVVRYPEEFGKSGLSEEEVRA SYBAFIRYPEFGHKSGLPEEBWKA DYTAFVRYPEEFGHKSGLPEEBWKA DYTAFVRYPEEFGHKSGLPEEBWKA SE K.ETY	RLTARGLLAMADKKET RLKARGIPFSSTLR 264 KCR CR CR
HIV-1 LAI SIVCP2GAB SO MM 239 HIV-2ROD PORT-COM 223 281-1 120 MP-1 956 741 1227 268 1570 1225 1320 1544 483 293 1147-1 1139-1	GYPPDSQNYTPGPGVRYPLJFGWCY GPPPDWQNYTPGPGTRPLCFGWCP GIIPDWQDYTSGPGIRYPK7FGULW GIIPDWQNYTHGFGRIPPFFGULW GIIPDWQNYTHGFGRIPPFFGULW 175 L	8 8 9 8 8 8 8 8 8 8 8 9 9 9 9 9 9 9 9 9	10 PICQHOMS DED**** REVLEWGPDS PICQHOMS DED**** REVLEWGPDS PAQTSKOM DWGOVLANKEDPTLAN PAQTSKP DHGSTLWWEDPLAN PAQTSKP DHGSTLWEDPLAN	RLAFHHVARELHPEYF****KNCS RLALRHIAREQHPEY****KDS TYBAYUNYPEFOSKSGLSEEVRR SYBAFIRYPEFOHKSGLPEEWRA DYTAFUNYPEFOHKSGLPEEWRA LSD KQLYQA-K N N	RLTARGLLNMADKKET RLKARGIPFSSTLR 264 KCR
HIV-1 LAI SIVCPZGAB SD MM 239 HIV-2ROD PORT-CON 223 281-1 120 MP-1 9566 741 1227 268 1570 1215 1320 1215 1320 1215 1320 1147-1 1139-1 1139-1 PE L 794	GYPPDSQNYTPGPGVRYPLJTFGWCY GPPPDWQNYTTGPGVRPPLCFGWCF GIIDWQNYTGPGVRPPFGWLW GIIDWQNYTGPGVRPPFGWLW GIIDWQNYTGPGVRPPFGWLW GIIDWQNYTGPGVRPPFGWLW GIIDWQNYTGPGVRPPFGWLW GIIDWQNYTGPGVRPPFGWLW GIIDWQNYTGPGVRPFGWLW GIIDWQNYTGPGVRPPFGWLW GIIDWQNYTGPGVRPFGWLW GIIDWQNYTGPGVRPFGWLW GIIDWGWLW GIIDWGWLW GIIDWQUWLW GIIDWGWLW GIIDWGWLW GIIDWULW GIIDWULW GIIDWULWUWLW GIIDWULWULWUWLW	8 9 KLUPPUTEEQVM_ANEG**DANC ***EKUPVEDEQVM_ANEG**DANC KLUPVUNSDEAQ ***EKUPVUNSDEAQ KLUPVUNSDEAQ ***EKUPVUNSDEAQ CO ****EKUPVUNSDEAQ	10 PICQHGMED20**** REVLEMEPDS PICQHGMED2**** REVLEMEPDS PAQTSQMDPMGEVLAMKPDFLAY PAQTSRPDPHGETLVWRPDFLAY Q	RLAFHHVARELHPEYF****KNC\$ RLALRHIAREQHPEYY****KD\$ TYBAYUNYPEFOSKSGLSEEVKR SYBAFIRYPEFOSKSGLPEEVKA DYTAFVNYPEFOSKSGLPEEVKA MC KE-TYM NN N SKH	RLTARGLLNMADKKET RLKARGIPFSSTLR 264 KCR
HIV-1 LAI SIVCPZGAB SD MM 239 HIV-2ROD FORT-CON 223 281-1 120 MP-1 956 741 1227 2688 1570 1215 1320 1544 483 293 1147-1 1139-1 EL 794 423	GYPEDSQNYTPGPGVRYPLJFGWCY GPPPDWQNYTGGPGTRPLCFGWCP GIIPDWQDYTGGPGTRPPFCPGWLW GIIPDWQNYTGGPGTRPPFFGWLW GIIPDWQNYTHGPGTRPPFFGWLW J175 LYVT	8 9 KLUPVEDEQMEANEGC*DANC *** KLUPVENCEQ *** KLUPVENCEQ *** KLUPVENCEQ *** KLUPVENCEQ *** KLUPVENCEQ *** KLUPVENCE *** COO COO	10 PYSLHAK DPE**** REVLEWRPDS PICQHOK MEDE**** REVLEWRPDS PAQTSKØMDPMGSULAWRPDFLLAY PAQTSKPDPHGSTLVWRPDPLLAY PAQTSKPDPHGSTLVWRPDPLLAY PAQTSKPDPHGSTLVWRPDPLAY 225	RLAFHHVARELHPEYF****KNC\$ RLALHHIAR2QHPEYY****KD\$ TYBAYUNYPEFOSKSGLSEEUKR DYTAFURYPEFOSKSGLSEEUKR DYTAFURYPEFOSHSGLPEEWKA STAFIRYPEYPESHSGLPEEWKA STAFIRYPEYPESHSGLPEEWKA NUNNUNUNUNUNUNUNUNUNUNUNUNUNUNUNUNUNUN	RLTARGLLNMADKKET RLKARGIPFSSTLR 264 KCR CR CR CR R* N* N* KE* KE* K* K* K*
HIV-1 LAI SIVCP2GAB SID MM 239 HIV-2ROD PORT-CON 223 281-1 120 MP-1 956 741 1227 268 1570 1215 1320 1544 483 293 1147-1 1139-1 1139-1 1139-3 794 423 379 428	GYPPDSQNYTPGPGVRYPLTFGWCY GPPPDWQNYTPGPGTRPPLCFGWLW GIIPDWQNYTSGPGIRYPKTPGNLW GIIPDWQNYTHGFGIRYPKPFGNLW GIIPDWQNYTHGFGIRYPKPFGNL L	9 8 8 8 8 9 8 8 9 8 8 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9	10 PICQHEMC DEPENT ** REVLEMBEDS PACTORN RED**** PACTORN REVEAL PACTORN R	RLAFHHVARELHPEYF****KNCS RLALAFHIAREQHPEYY****KDS TYEAYUNYPEEFOSKSGLSEEEVRR SYBAFI RIY PEEFORKSGLPEENRA DYTAFUNYPEFORKSGLPEENRA LSO KQLYQA-K N N	RLTARGLLNMADKKET RLKARGIPFSSTLR 264 KCR
HIV-1 LAI SIVCP2GAB SD MM 239 PORT-CON 223 281-1 120 MP-1 956 741 1227 268 1570 1225 1320 1544 483 293 1147-1 1139-1 BE 794 423 379 1428 293 2142	GYPPDSQNYTPGPGVRYPLJFG0WCY GPPPDWQNYTTGPGTRPPLCFGWCF GIIPDWQDYTGGPGIRYPRFGULW GIIPDWQNYTHGPGIRYPRFGULW GIIPDWQNYTHGPGIRYPRFGULW L	8 9 KLUPYEDEQUEANEG**DANCG**DANC ***EKUPYEDEQUEANEG**DANC KLUPYEDEQUEACE ***EKUPYENEK KLUPYEDEQUEACE ***EKUPYENEK COO ***CO COO ***CO COO ***CO COO ***CO COO ***CO COO ***CO	10 PICQHGMEDE2**** REVLEWREPS PICQHGMEDE2**** REVLEWREPS PAQTSK0MDFWGEVLAWREPDFLAY PAQTSKPDHGETLWREPDFLAY PAQTSKPDHGETLWREPDFLAY PAQTSKPDHFGETLWREPDFLAY PAQTSKPDHFGETLWREPDFLAY PAQTSKPDHFGETLWREPDFLAY PAQTSKPDHFGETLWREPDFLAY PAQTSKPDHFGETLWREPDFLAY PAQTSKPDHFGETLWREPDFLAY PAQTSKPHENEX Q	RLAFHHVARELHPEYF****KNČS RLALRHIAREQHPEYY****KDŠ TYBAYUNYPEFOSKSGLSEEVRA DYTAFUNYPEFOSKSGLSEEVRA DYTAFUNYPEFOSKSGLSEEVRA 250 KQLYQA-K NNK NNK SNNK SNNK SN SN	RLTARGLLNMADKKET RLKARGIPFSSTLR 264 KCR CR CR * N* * N* KE* K
HIV-1 LAI SIVCPZGAB SD MM 239 HIV-2ROD PORT-CON 223 281-1 120 MP-1 9566 1570 1215 1320 1544 483 293 1147-1 1139-1 1139-1 2154 483 293 1147-1 1139-1 215 124 483 293 1147-1 1139-1 11139-1 11139-1 1139-1 1139-1 11139-1 111	GYPEDSQNYTPGROUNYPLTFGWCY GPPPDWQNYTTGGOTRPPLCFGWCP GIIPDWQDYTGGOTRPPRCFWLW GIIPDWQNYTGGOTRPPKPGWLW GIIPDWQNYTGGOTRPPKPGWLW GIIPDWQNYTGGOTRPKPFGWLW GIIPDWQNYTGGOTRPKPFGWLW GIIPDWQNYTGGOTRPKFFGWLW GIIPDWGWLW GIIPDWLW GIIPDWLW GIIPDWLW GIIPMFFGWLW	8 9 KLUPYEDEQYMANEG**DANC ***EENTENC KLUPYEDEQYMANEG**DANC ***EENTENC KLUPYEDTEQY**EGT ***EENTENC KLUPYEDTYP***EGT ***EENTENC KLUPYEDTYP***EGT ***EENTENC ***EGT ***EICT ***EGT ***EICT <tr< td=""><td>10 PICUHAM DEP:**** REVLEMREDS PICUHAM DEP:**** REVLEMREDS PAQTSK0MDEWGEVLAMKEDPTLAY PAQTSKPDHGETLWREDPLLAY PAQTSKPDHGETLWREDPLAY PAQTSKPDHGETLWREDPLAY PAQTSKPDHOPHGETLWREDPLAY PAQTSKPDHAM PAQTSKPDHOPHGETLWREDPLAY PAQTSKPLAY PAQTSKPLAY PAQTSKPHOPHOPHANK PAQTSKPLAY PAQTSKPLAY PAQTSKPLAY PAQTSKPANK PAQTSKPLAY PAQTSKPLAY</td><td>RLAFHHVARELHPEYP****KD\$ RLALHHIAREQHPEYY****KD\$ TYBAYURYPEEKGKSGLSEEUKR DYTAFURYPEFCHKSGLPEEBKKA DYTAFURYPEFCHKSGLPEEBKKA SK=D-TYM</td><td>RLTARGLLNMADKKET RLKARGIPFSSTLR 264 KCR CR CR </td></tr<>	10 PICUHAM DEP:**** REVLEMREDS PICUHAM DEP:**** REVLEMREDS PAQTSK0MDEWGEVLAMKEDPTLAY PAQTSKPDHGETLWREDPLLAY PAQTSKPDHGETLWREDPLAY PAQTSKPDHGETLWREDPLAY PAQTSKPDHOPHGETLWREDPLAY PAQTSKPDHAM PAQTSKPDHOPHGETLWREDPLAY PAQTSKPLAY PAQTSKPLAY PAQTSKPHOPHOPHANK PAQTSKPLAY PAQTSKPLAY PAQTSKPLAY PAQTSKPANK PAQTSKPLAY PAQTSKPLAY	RLAFHHVARELHPEYP****KD\$ RLALHHIAREQHPEYY****KD\$ TYBAYURYPEEKGKSGLSEEUKR DYTAFURYPEFCHKSGLPEEBKKA DYTAFURYPEFCHKSGLPEEBKKA SK=D-TYM	RLTARGLLNMADKKET RLKARGIPFSSTLR 264 KCR CR CR
HIV-1 LAI SIVCP2GAB SID MM 239 HIV-2ROD PORT-CON 223 281-1 120 MP-1 956 741 1227 268 1570 1215 1320 1544 483 293 1147-1 1139-1 1139-1 BL 822 293 1147-1 1139-1 1139-1 1227 94 423 379 944 222	GYPPDSQNYTPGPCVRYPLTFGWCY GPPPDWQNYTPGPCTRPLCFGWCB GIIPDWQNYTGPCRYPRPGWLW GIIPDWQNYTHGPCRYPRPFGWLW GIIPDWQNYTHGPCRYPRPFGWLW GIIPDWQNYTHGPCRYPRPFGWLW -L Y -A T -A I -M -A I -M -A -M	8 9 KLUPPUTEEQVEMANEGC**DANCLIH KLUPPUTEEQVEMANEGC**DANCLIH KLUPPUTEEQVEXECTERTICLIH KLUPPUTYP***EGTTERTICLIH KLUPUTYP***EGTTERTICLIH KLUPUTYP****EGTTERTICLIH KLUPUTYP****EGTTERTICLIH KLUPUTYP****EGTTERTICLIH KLUPUTYP****EGTTERTICLIH KLUPUTYP****EGTTERTICLIH KLUPUTYP****EGTTERTICLIH KLUPUTYP**********************************	10 PICQHEMCTPE**** REVLEMBTES PICQHEMCTED**** REVLEMBTES PAQTSKP DPHOSTLAMPEDVLAM PAQTSKP DPHOSTLAMPEDVL	RLAFHHVARELHPEYF****KNCS RLALAHIAREQHPEY****KDS TYEAVUNYPEFOSKSGLSEEVRR SYBAT RIY PEFOHKSGLPEEWKA DYTAFVNYPEFOHKSGLPEEWKA SKE-T-N-Y-WA KQLY-WAK NK N S	RLTARGLLAMADKKET RLKARGIPFSSTLR 264 KCR CR CR CR
HIV-1 LAI SIVCP2GAB SD MM 239 HIV-2ROD PORT-CON 223 281-1 120 MP-1 956 741 1227 268 1570 1225 1320 1544 483 293 1147-1 1139-1 BL 794 423 379 1428 293 293 1147-1 84 483 293 293 1147-1 85 1139-1 BL 794 483 379 984 423 379 125 125 125 125 125 125 125 125 125 125	GYPPDSQNYTPGPGVRYPLJFGWCY GPPPDWQNYTPGPGIRPPLCFGWCF GIIPDWQDYT9GPGIRPPFCFWLW GIIPDWQNYTHGFGVIPHFFGWLW GIIPDWQNYTHGFGVIPHFFGWLW L	8 9 KLUPYEDEQUEQANEG**DANK 9 KLUPYEDEQUEQANEG**DANK ***EENY BAH KLUPYEDEQUEQANEG**DANK ***EENY BAH KLUPYEDEQUEQANEG**EGT FEHYLCLWH 200 LPQREN-AN-L 200 CS	10 PICQHOMS DED**** REVLEWREPS PICQHOMS DED**** REVLEWREPS PAQTSKOP DWEVLAMKEDPTLAN PAQTSKED DHIGGTLWREPDTLAN PAQTSKED DHIGGTLWREPTLAN PAQTSKED DHIGGTLWREPTLAN PAQTSKED DHIGGTLWREPTLAN PAQTSKED DHIGGTLWREPTLAN PAQTANCH PAQTANCH PAQTANCH PAQTANCH PAQTANC	RLAFHHVARELHPEYF****KNČS RLALRHIAREQHPEYY****KDŠ TYPAZVNYPEFOSKSGLSEEVRA DYTAFUNYPEFOSKSGLSEEVRA DYTAFUNYPEFOSKSGLSEEVRA SYAPITNYPEFOSKSGLSEEVRA KQLYQA-K NYQ N	RLTARGLLNMADKKET RLKARGIPFSSTLR 264 KCR
HIV-1 LAI SIVCPZGAB SD MM 239 HIV-2ROD FORT-CON 223 281-1 120 MP-1 9566 741 1227 268 1320 1215 1225 1225 1225 1225 1225 1225 12	GYPEDSONYTPGROUNYPLTFGWCY GPPPDWQNYTTGBOINYPRFGWCW GIIDDWQNYTGBOINYPRFFGWLW GIIDWQNYTGBOINYPRFFGWLW GIIDWQNYTGBOINYPRFFGWLW GIIDWQNYTGBOINYPRFFGWLW GIIDWQNYTGBOINYPRFFGWLW GIIDWGNYTGBONN GIIDWGNYTGBONN GIIDWGNYTGBONN GIIDWGNYTGBONN GIIDWGNYTGBONN GIIDWGNY GIIDWGNY <t< td=""><td>8 9 KLUPPUTEEQVM_ANEGC*DANC ***EENTENCLIH KLUPPUTEEQVM_ANEGC*DANC ***EENTENCLIH KLUPPUTPC**ECTTENCLIN ** KLUPPUTPC**ECTTENCLIN ** KLUPPUTPC**ECTTENCLIN ** COO 200 LPQREN-AN-L 200 </td><td>10 PICQHGMED2**** REVLEMEPDS PICQHGMED2**** REVLEMEPDS PAQTSQMDPMGEVLAMKPDFLAY PAQTSRPDPHGETLVMRPDFLAY PAQTSRPLAY PAQTSRPLAY PAQTSRPLAY PAQTSRPLAY PAQTSRPLAY PAQTSRPLAY PAQTSRPLAY PAQTSRPLAY PAQTSRPLAY <</td><td>RLAFHHVARELHPEYF****KNC\$ RLALHHIAREQHPEYY***KD\$ TYBAYURYPEFOSKSGLSEEVKR DYTAFVRYPEFOSKSGLSEEVKR DYTAFVRYPEFOSKSGLPEEVKA K.E-TYM</td><td>RLTARGLINMADKKET RLKARGIPFSSTLR 264 KCR </td></t<>	8 9 KLUPPUTEEQVM_ANEGC*DANC ***EENTENCLIH KLUPPUTEEQVM_ANEGC*DANC ***EENTENCLIH KLUPPUTPC**ECTTENCLIN ** KLUPPUTPC**ECTTENCLIN ** KLUPPUTPC**ECTTENCLIN ** COO 200 LPQREN-AN-L 200	10 PICQHGMED2**** REVLEMEPDS PICQHGMED2**** REVLEMEPDS PAQTSQMDPMGEVLAMKPDFLAY PAQTSRPDPHGETLVMRPDFLAY PAQTSRPLAY PAQTSRPLAY PAQTSRPLAY PAQTSRPLAY PAQTSRPLAY PAQTSRPLAY PAQTSRPLAY PAQTSRPLAY PAQTSRPLAY <	RLAFHHVARELHPEYF****KNC\$ RLALHHIAREQHPEYY***KD\$ TYBAYURYPEFOSKSGLSEEVKR DYTAFVRYPEFOSKSGLSEEVKR DYTAFVRYPEFOSKSGLPEEVKA K.E-TYM	RLTARGLINMADKKET RLKARGIPFSSTLR 264 KCR
HIV-1 LAI SIVCP2GAB SID MM 239 HIV-2ROD PORT-CON 223 281-1 120 MP-1 956 741 1227 268 1570 1215 1320 1544 483 293 1147-1 1139-1 1139-1 1139-1 1227 984 423 379 1422 292 292 1096 984 1541 1541 423 379 1428 293 1147-1 1138 294 295 295 295 295 295 295 295 295 295 295	GYPEDSQNTTPGPGVRYPLJTFGWCY GPPPDWQNTTGFGVRYPLCFGWCP GIIPDWQDYTGFGVRYPFFGWLW GIIPDWQNTTGFGVRYPFFGWLW GIIPDWQNTTGFGVRYPFFGWLW GIIPDWQNTTGFGVRYPFFGWLW GIIPDWQNTTGFGVRYPFFGWLW GIIPDWQNTTGFGVRYPFFGWLW GIIPDWQNTTGFGVRYPFFGWLW GIIPDWQNTTGFGVRYPFGWLW GIIPDWQNTTGFGVRYPFGWLW GIIPDWQNTTGFGVRYPFGWLW N	8 9 KLUPVEDEQWEANEGC*DANC ***EENTENC KLUPVENDERQO ***EENTENC KLUPVEDT **ECTENC KLUPVEDT **ECTENC KLUPVEDT **ECTENC VA 200 LPQ	10 P120LH0MD/DE5**** REVLEWRPDS P1CQH0MB/RED**** REVLEWRPDF PAQTSK/DMD/MGEVLAWRPDFLLAY PAQTSR/DMD/MGETLWREPDFLLAY PAQTSR/DMD/MGETLWREPDFLAY PAQTSR/DMD/MGETLWREPTLAY PAQTSR/DMD/MGETLWREPTLAY PAQTSR/DMD/MGETLWREPTLAY PAQTSR/DMD/MGETLWREPTLAY PAQTSR/DMD/MGETLWREPTLAY PAQTSR/DMD/MGETLWREPTLAY PAQTSR/DMD/MGETLWREPTLAY PAQTSR/DMD/MGETLWREPTLAY P	RLAFHHVARELHPEYF****KNC\$ RLAFHHVARELHPEYF***KD\$ TYBAYUNYPEFOSKSGLSEEVKR DYTAFUVRYPEFOSKSGLSEEVKR SYBAFIRYPEFOSHSGLPEEWKA DYTAFURYPEFOSHSGLPEEWKA SYBAFIRYPEFOSHSGLPEWKA SYBAFIRYPEFOSHSGLPEWKA SYBAFIRYPEFOSHSGLPEWKA SYBAFIRYPEFOSHSGLPEWKA	RLTARGLINMADKKET RLKARGIPFSSTLR 264 KCR CR CR R* N* KN* K*
HIV-1 LAI SIVCP2GAB SD MM 239 HIV-2ROD PORT-COM 223 241-1 120 MP-1 956 741 1227 268 1570 1215 1320 1544 483 293 1147-1 1139-1 BL 423 379 1442 293 1147-1 1139-1 BL 423 379 1428 293 1147-1 1139-1 BL 544 423 379 1428 293 1147-1 1139-1 BL 545 11428 295 11428 205 215 11428 215 215 215 215 215 215 215 215 215 215	GYPPDSQNYTPGPGURYPLJTFGWCY GFPPDWQNYTTGPGIRPLCFGWCF GIIPDWQDYTGGPGIRPPFGULW GIIPDWQNYTHGGURYPPFGULW GIIPDWQNYTHGGURYPFGULW T LYT	8 9 KLUPYEDEQUEQANEG**DANK F***EENSIE KLUPYEDEQUEQANEG**DANK F***EENSIE KLUPYEDEQUEQANEG**EGU FEHY LMH KLUPYUDYP**EGU FEHY LMH KLUPYUDYUPY**EGU FEHY LMH	10 PICQHOMS DED**** REVLEMBEDS PICQHOMS DED**** REVLEMBEDS PAQTSKOP DWGSTLWKEDPTLAN PAQTSKED DHGGTLWREDPTLAN PAQTSKED DHGTLWREDPTLAN PAQTSKED DHGTLWREDPTLAN PAQTSKED DHGTLWREDPTLAN PAQTSKED DHGTLWREDPTLAN PAQTSKED TANKEDPTLWREDPTLAN PAQTSKED TANKEDPTLAN PAQTANE PAQTANE PAQTANKEDPTLAN PAQTANE </td <td>RLAFHHVARELHPEYF****KNČS RLALRHIAREQHPEY****KDŠ TYBAYUNYPEFOSKSGLSEEVKR SYBAFIXPEPEOHKSGLPEEWKA DYTAFUNYPEFOHKSGLPEEWKA LSD KQLYQA-K N N</td> <td>RLTARGLLNMADKKET RLKARGIPFSSTLR 264 KCR </td>	RLAFHHVARELHPEYF****KNČS RLALRHIAREQHPEY****KDŠ TYBAYUNYPEFOSKSGLSEEVKR SYBAFIXPEPEOHKSGLPEEWKA DYTAFUNYPEFOHKSGLPEEWKA LSD KQLYQA-K N N	RLTARGLLNMADKKET RLKARGIPFSSTLR 264 KCR
HIV-1 LAI SIVCP2GAB SD MM 22GAB SD MM 22GAB PORT-CON 223 281-1 12C0 MP-1 227 268 741 1227 268 1570 1215 11227 1215 1124 433 293 1147-1 1139-1 82 292 1096 984 423 379 1428 292 1096 984 1543 1543 1543 1543 1543 1543 1543 154	GYPPDSQNYTPGPGURYPLJTFGWCY GFPPDWQNYTTGPGIRPPLCFGWCF GIIPDWQNYTGPGIRPPFCFWLW GIIPDWQNYTHGGURYPKFFGWLW GIIPDWQNYTHGGURYPKFFGWLW GIIPDWQNYTHGGURYPKFFGWLW LYVT	8 9 KLUPPUTEQVM_ANRG**DANC ****EMX KLUPPUTEQVM_ANRG**DANC ****EMX KLUPPUTPUTP***EGT ****EMX KLUPPUTP***EGT ****EMX KLUPPUTP***EGT ****EMX ****EGT ******EMX KLUPPUTP***EGT ****EMX ****EGT *****EMX ****EGT ************************************	10 PICQHGMEDED**** REVLEMEPDS PICQHGMEDED**** REVLEMEPDS PICQHGMEDED**** REVLEMEPDS PAQTSK0MDFWGEVLAMKEDPTLAY PAQTSKPDHGETLWREPDHLAY PAQTSKPDHGETLWREPDHLAY PAQTSKPDHFGETLWREPDHLAY PAQTSKPDHFGETLWREPDHLAY PAQTSKPDHFGETLWREPDHLAY PAQTSKPHENKER Q M-K Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q PAQTSKPARD Q Q Q H H V K H K Q K Q V K Q K	RLAFHHVARELHPEYF****KNČS RLALRHIAREQHPEYF****KS TYBAYUNYPEFOSKSGLSEEVRA DYTAFVNYPEFOSKSGLSEEVRA DYTAFVNYPEFOSKSGLSEEVRA STAFITNYPA KQLYQA-K NN N S	RLTARGLLNMADKKET RLKARGIPFSSTLR 264 KCR CR
HIV-1 LAI SIVCPZGAB SD MM 239 HIV-2ROD PORT-CON 223 281-1 120 MP-1 9556 1320 1215 1320 1215 1320 1544 483 293 1147-1 1139-1 1139-1 1139-1 2158 484 293 1147-1 1139-1 1139-1 1139-1 1139-1 1139-1 1138 EL E E E E 2056 2057 2057 2057 2057 2057 2057 2057 2057	GYPEDSQNYTPGROUNYPLTFGWCY GPPPDWQNYTTGGOTNPPLCFGWCP GIIPDWQDYTGGOTNPPRFWLW GIIPDWQNYTGGOTNPPKPGULW GIIPDWQNYTGGOTNPPKPGULW GIIPDWQNYTGGOTNPKPFGULW GIIPDWQNYTGGOTNPKPFGULW GIIPDWQNYTGGOTNPKFFGULW GIIPDWQNYTGGOTNFFFGULW GIIPDWQNYTGGOTNFFFGULW GIIPDWQNYTGGOTNFFFGULW GIIPDWQNYTGGOTNFFFGULW GIIPDWQNYTGGOTNFFFGULW GIIPDWQNYTGGOTNFFFGULW GIIPDWQNYTGGOTNFFFGULW GIIPDWQNYTGGOTNFFFGULW GIIPMTGGULW GIIPMGULW GIIPMGULW GIIPMGULW GIIPMGULW GIIPMGULW GIIPMGULW GIIPMGULW GIIPMGULW	8 9 KLUPYEDEQMEANEG**DANC ***EENTENENE KLUPYEDEQMEANEG**DANC ***EENTENE KLUPYEDTEQMEANEG**ECHTENENC ** KLUPYEDTYP**ECHTENE ** COO ***ECHTENENC KLUPYEDTYP**ECHTENE ** COO ***ECHTENENC KLUPYEDTYP***ECHTENE ** COO ***ECHTENE ***ECHTENE ** COO ***ECHTENE ***ECHTENE ** ***ECHTENE ** ***ECHTENENE ** ****ECHTENE ** ************************************	10 PJULHUMDPDF*****REVLEWRPDS PICQHOM BRED*****REVLEWRPDF PAQTSKMDPWGEVLAWRPDFLLAY PAQTSRPDPHGETLUWRPDFLAY PAQTSRPDPHGETLWRPDFLAY PAQTSRPDHENER PAQTSRPDHENER PAQTSRPLANCK PAQTSRPLANCK PAQTSRPLANCK PAQTSRPLANCK PAQTSRPLANCK PAQTSRPLANCK PAQTSRPLANCK PAQTSRPLANCK PAQTSRPLANCK PAQTSRPLANCK <td>RLAFHHVARELHPEYP****KNC\$ RLAFHHVARELHPEYP****KD\$ TYBAYVRYPEFOSKSGLSEEVKR DYTAFVRYPEFOSKSGLSEEVKR DYTAFVRYPEFOSKSGLSEEVKR 250 KQLYQA-K 250 N</td> <td>RLTARGLLNMADKKET RLKARGIPFSSTLR 264 KCR CR </td>	RLAFHHVARELHPEYP****KNC\$ RLAFHHVARELHPEYP****KD\$ TYBAYVRYPEFOSKSGLSEEVKR DYTAFVRYPEFOSKSGLSEEVKR DYTAFVRYPEFOSKSGLSEEVKR 250 KQLYQA-K 250 N	RLTARGLLNMADKKET RLKARGIPFSSTLR 264 KCR CR
HIV-1 LAI SIVCPZGAB SD MM 239 HIV-2ROD PORT-CON 223 221-1 120 MP-1 956 741 1227 268 1570 1215 1320 1544 483 293 1139-1 1139-1 1139-1 1139-1 1139-1 2125 1390 1244 483 293 1147-1 1139-11	GYPPDSQNYTPGPGURYPLJTFGWCY GFPPDWQNYTTGPGIRPLCFGWCP GIIPDWQDYTSGPGIRPPCFGULW GIIPDWQNYTHGGURYPPFGULW GIIPDWQNYTHGGURYPPFGULW GIIPDWQNYTHGGURYPPFGULW T LYVT	8 9 KLUPPUEDEQUEQANEG**DANK HKLUPLTEEQUEQANEG**DANK KLUPUDTEEQUEQANEG**DANK HKLUPUDTEQ CONSTRUCT HKLUPUDTEQ	10 PICQHOMS DED**** REVLEMBEDS PICQHOMS DED**** REVLEMBEDS PAQTSKOP DWGEVLAMKEDPTLAN PAQTSKP DHGETL/WREDPHLAN H V H VK U VK U K VK K VK VK VK VK VK	RLAFHHVARELHPEYF****KNČS RLALRHIAREQHPEY****KDŠ TYBAYUNYPEFOSKSGLŠEEVRA DYTAFUNYPEFOSKSGLŠEEVRA DYTAFUNYPEFOSKSGLŠEEVRA SYAPI IN PEFOSKSGLŠEEVRA KQLYQA-K NNK NNK SNNK SNNK SNNK SNNK N-VNL<	RLTARGLLNMADKKET RLKARGIPFSSTLR 264 KCR
HIV-1 LAI SIVCP2GAB SD MM 239 HIV-2ROD PORT-CON 223 281-1 120 MP-1 956 741 1227 268 1570 1225 1320 1544 483 293 1147-1 1139-1 BL 794 482 379 1428 293 1147-1 1139-1 BL 794 483 379 1428 293 1147-1 1139-1 BL 794 483 540 546 546 546 546 5546 5546 5546 5546	GYPPDSQNYTPGPGURYPLJTFGWCY GIFPDWQNYTGPGURYPRJCGWCF GIIDWQNYTGPGURYPRFGULW GIIDWQNYTHGPGURYPRFGULW A T M	8 9 KLUPYEDEQUE ANKG**DINC LH KLUPYEDEQUE ANKG**DINC LH KLUPYEDEQUE ** EGL FEHYL LM KLUPYDTYP**EGL FEHCL LM KLUPYDTYP**EGL FEHCL LM KLUPYDTYP**EGL FEHCL LM COO LPQREM-AN-L	10 PICQHOMS DED**** REVLEWGPDS PICQHOMS DED**** REVLEWGPDS PAQTSKOPDWGEVLANKEDPTLAN PAQTSKPDPHOETLUK PAQTSKPDHOETLUK PAQTSKPLANK PAQTSKPLANK PAQTSKPLANK PAQTON PAQTSKPLANK PAQTSKPLANK PAQTON PAQTSKPLANK PAQTON PAQTSKPLANK PAQTON <	RLAFHHVARELHPEYF****KNČS RLALRHIAREQHPEYT****KDŠ TYBAVUNYPEFOKSGLSEEVRA DYTAFVNYPEFOKSGLPEEPKKA DYTAFVNYPEFOKSGLPEEPKKA STAFITYPAV KQL N N N N N N N N N N N SN VT SN VT SN -VNL -VNL -VRL -VRL -VRL -VRL -VRL -VRL -V-RL -V-RL -V-RL -V-RL -V-RL -V-RL -V-RL -V-NL -V-RL -V-RL -V-RL -V-RL -V-RL -V-RL -V-NL -V-V-R	RLTARGLLNMADKKET RLKARGIPFSSTLR 264 KCR CR CR *

Fig. 2. Alignment of the *nef* gene amino acid sequence of 37 HIV-2 infected individuals in the Portuguese cohort related to HIV-1, SIV_{CPZ}, SIV_{MAC}, HIV-2_{ROD} and the Portuguese HIV-2 *nef* consensus sequence. Sites important for *nef* function from the N- to the C-termini are highlighted for the different groups as follows: 1, myristoylation site; 2, N-terminal helix and putative MHC-I downregulation site; 3, tyrosine-based adaptor protein (AP) recruitment motif; 4, CD4-binding site (as characterized for HIV-1); 5, acidic cluster (MHC-I downregulation); 6, proline-based repeat sequence (MHC-I and SH3 binding); 7, PAK binding; 8, COP-1 recruitment; 9, di-leucine-based AP recruitment (HIV-1 *nef*); and 10, V-ATPase and Raf-1 binding.

methionine, with L^{198} conserved throughout. Interestingly, patients 280, 120, 281, MP and 956 grouped into an EANCLL configuration, three of which (281, MP and 956) had tetra-proline (PxxP)₃ motifs. Overall, 15/37 (40.5%) individuals had the $L^{198}L^{199}$ motif, with seven (19%) substituting the leucine at position 199 for valine and 15 (40.5%) for methionine. The number of LL motifs was distributed evenly between asymptomatic and symptomatic groups at eight and seven, respectively, forming an asymptomatic consensus of LM and symptomatic consensus of LL. Interestingly, the region immediately preceding this was relatively heterogeneous across all

individuals (aa 185–195, Fig. 2), covering a putative site for β -COP recruitment. However, there was no apparent relationship between this variability and disease status.

(iii) Putative SH3-binding domains characterized by proline-rich (PxxP) motifs. The proline-rich motif (P¹⁰⁴xxPLR) was assessed for variation in more detail. Within this core region, $R^{105} \rightarrow K$ was observed frequently among these Portuguese sequences to form the sequence P¹⁰⁴KVPLR, with the arginine residue at position 109. This adds to the observation that this region, which mediates interaction with SH3 ligands, is evolutionarily conserved among primate lentiviruses and is present as a minus-orientation binding consensus. Moreover, HIV-2 nef sequences generated in this study exhibited a broad range of PxxP configurations (Fig. 2, Table 2). An HIV-1like (PxxP)₃ configuration was observed for six individuals (16.2%), three asymptomatic (MP, 511 and 293) and three symptomatic (281, 956 and 379). Conversely, disrupted minimal PxxP motifs were associated exclusively with an asymptomatic profile (11/26, 42.3%), where the P¹⁰⁷ residue was mutated to glutamine in nine individuals (PL, 1567, 1215, 1570, 1147, 794, 1268, 741 and 1139) and the P¹⁰⁴ residue to serine in two (120 and 292), the only individuals not preserving the proline at position 104. Disruption of the core P¹⁰⁴xxPLR region, $P^{107} \rightarrow Q$ (n=9) and $P^{104} \rightarrow S$ (n=2) was associated strongly with an asymptomatic phenotype, although patient PL who had been infected for 24 years had a relatively low CD4 count in 1997 (127 cells mm⁻³). The tyrosine (Y^{145}) forming part of a putative hydrophobic pocket $(Y^{145}LEKE^{149}E^{150})$ distal to the PxxP region implicated in SH3 binding in HIV-1 nef (Collette et al., 2000) was absolutely conserved, with minimal changes to other residues associated with this motif, with E¹⁴⁹ mutated to either a glutamine or a lysine in patients 1543 and 1567 and E¹⁵⁰ mutated to lysine in patient 223; all three individuals were asymptomatic. In addition, there was an $I^{144} \rightarrow L$ substitution in six individuals (281, MP, 120, 984, 1428 and 1096) immediately prior to Y145 and similar mutations in three other individuals: $I^{144} \rightarrow T$ (n=2) and $I^{144} \rightarrow M$ (*n*=1), although the significance of these changes is not known. Other sequences associated with the $(PxxP)_3$ region in HIV-1 is a phenylalanine residue (F^{122}) . In HIV-2, this appeared in four individuals (281, 120, MP and 741), although the consensus of the comparable residue is a valine. No additional PxxP motifs in other regions of *nef* were identified in any of these HIV-2 sequences. Variability in the central PxxP region was investigated further for associations with disease status and virus load.

Relationship between PxxP sequence and plasma virus load

The spectrum of proline configurations observed in these patients was related to HIV-2 viral RNA levels determined for 27 different individuals at the first time-point (Table 2). Overall, HIV-2 virus load was low, reflecting

the high proportion of asymptomatic persons, with 77.7% (21/27) having ≤ 500 copies ml⁻¹ and 59.3% (16/27) having ≤ 200 copies ml⁻¹. Only 3 of 21 individuals with $\leqslant\!500$ copies plasma RNA ${\rm ml}^{-1}$ were symptomatic (patients 423, 546 and 268). Of these, patient 423 had a CD4 count of 996 cells mm⁻³ and patients 546 and 268 had counts of 297 and 340 cells mm^{-3} , respectively. There was a non-significant trend for increasing levels of HIV-2 RNA with additional prolines, although the number of individuals with high virus load was small, with only three individuals having relatively high levels of HIV-2 RNA (patients 379, 1069 and 984), with 62 950, 34 750 and 8830 copies ml^{-1} , respectively; all three were symptomatic. Of four individuals who possessed the tetra-proline motif and for whom virus load data were available, two (patients 379 and 117-2) had high virus loads; the other two individuals (511 and MP) had low/undetectable virus loads and were asymptomatic, with CD4 counts of 473 and 610 cells mm^{-3} , respectively.

Of the 20 individuals with an intact $P^{104}xxPLR$ motif but not a tetra-proline configuration, eight (117-1, 423, 280, 268, 984, 138, 1069 and 546) had evidence of symptomatic infection/AIDS. The second sample taken from patient 117 at 5 years later indicated an $A^{107} \rightarrow P$ change, forming a (PxxP)₃ motif, and an $A^{106} \rightarrow V$ change. The Nef protein of all 20 would seem to be capable of interacting with SH3 in this region. Further subdivisions of this group indicate that 11 individuals have only the minimal $P^{104}xxPLR$ sequence (*n*=3 symptomatic: 268, 984 and 138; *n*=8 asymptomatic: 1378, 223, EP, 1096, 1544, 1320, BL and LF); eight (*n*=4 symptomatic: 280, 1069, 423 and 117; *n*=4 asymptomatic: 1543, 1428, 483 and 1227) had an additional P^{101} ($P^{101}xxPxxPLR$) and one symptomatic patient (546) had a PxxPLRP¹¹⁰ configuration.

Finally, 11 individuals had a disruption in the central region of the PxxP motif, mostly as a $P^{107} \rightarrow Q$ mutation (n=9), all of which were asymptomatic; this was statistically significant (Fisher's exact test, P=0.026). There was also a non-significant trend for lower virus loads with fewer prolines present (Table 2). Levels of viral RNA correlated more closely to clinical status, with a significantly higher plasma virus load in symptomatic than asymptomatic individuals (mean HIV-2 RNA levels were 17051 and 817 copies ml^{-1} , respectively; unpaired Student *t*-test, P=0.0026), whereby all symptomatic individuals possessed an intact core P¹⁰⁴xxPLR motif. Conversely, disruptions to the P¹⁰⁴xxPLR core were associated with lower virus loads, which only exceeded 5000 copies ml⁻¹ in one asymptomatic individual (PL) who had been infected for 24 years and who had a low (127 cells mm^{-3}) CD4 count.

Longitudinal changes in *nef* and cloned *nef* sequences

A second sample obtained from each of six individuals, two symptomatic (281-2 and 117-2) and four asymptomatic (LF-2, MP-2, 1147-2 and 1139-2), allowed comparison with *nef* sequences from earlier time-points. The close genetic

Table 2. Variation of the PxxP region of HIV-2 *nef* in the 37 HIV-2-infected individuals related to levels of plasma virus load (27 individuals) and CD4 lymphocyte counts (25 individuals)

Mean virus load and CD4 counts are indicated for three groups categorized according to the configuration of proline residues ranging from a tetra-proline configuration to a disrupted minimal core $P^{104}xxP^{107}$ motif. NA, Not available.

Reference	ference HIV-2 <i>nef</i> sequence variation of PxxP motif									Viral RNA load (copies ml ⁻¹)	CD4 count (cells mm ⁻³)	Clinical status	
C	D 101	N/	т	D	р	17	D	т	р	A 110			
Totro prolino	Р	v	1	P	ĸ	v	P	L	ĸ	A			
configuration													
(DyyD)													
(FXXF) ₃	D	V	т	D	V	V	D	т	D	р	NTA	16	Symptomatic
201-1	г	V	т	r D	R D	v	r D	L	к D	r D	NA 62.050	40 366	Symptomatic
379	r D	v	T	r n	К D	v	r n	L	K D	r D	62 950	220	Symptomatic
950 MD 1	P	V	I T	r D	K V	V	P D	L	K D	P D	NA < 200	220	Symptomatic
MP-1 202	P	V	I T	r D	N D	V T	P D	L	K	P	< 200	510	Asymptomatic
293	P	V	I T	P D	K D	1 V	P D	L	K D	P	NA	519	Asymptomatic
511 M (()	Р	v	1	Р	К	v	Р	L	К	Р	500	4/3	Asymptomatic
Mean $(n=6)$											$21156\ (n=3)$	3/2 (n=6)	
Intact PxxP ²⁰⁰													
sequence		3.7	T	ъ	ъ	17	n		ъ				o .
138	Н	V	T	P	R	V	P	V	R	A	NA	NA	Symptomatic
984	S	V	Т	P	R	V	P	L	R	A	8 830	195	Symptomatic
268	S	V	Т	Р	R	V	Р	L	R	E	200	340	Symptomatic
BL	S	V	Т	Р	K	Т	Р	L	R	E	<200	288	Asymptomatic
EP	А	V	Т	Р	K	V	Р	L	R	Т	<200	490	Asymptomatic
LF-1	S	V	Т	Р	K	V	Р	L	R	А	400	437	Asymptomatic
223	S	V	Т	Р	R	V	Р	L	R	А	NA	161	Asymptomatic
1096	S	V	Т	Р	R	V	Р	L	R	А	<200	902	Asymptomatic
1320	S	V	Т	Р	R	V	Р	L	R	G	200	444	Asymptomatic
1378	S	V	R	Р	R	V	Р	L	R	А	NA	NA	Asymptomatic
1544	Η	V	Κ	Р	R	V	Р	L	R	Е	NA	730	Asymptomatic
117-1	Р	V	Т	Р	Κ	А	Р	L	R	А	NA	199	Symptomatic
280	S	V	Т	Р	R	V	Р	L	R	Т	NA	NA	Symptomatic
423	Р	V	Т	Р	Κ	V	Р	L	R	А	< 200	966	Symptomatic
1069	Р	V	R	Р	R	V	Р	L	R	Т	34 750	NA	Symptomatic
483	Р	V	Т	Р	R	V	Р	L	R	А	200	NA	Asymptomatic
1227	Р	V	Т	Р	R	V	Р	L	R	Е	< 200	NA	Asymptomatic
1428	Р	V	S	Р	R	V	Р	L	R	Е	< 200	NA	Asymptomatic
1543	Р	V	Т	Р	Q	V	Р	L	R	Q	< 200	993	Asymptomatic
546	S	V	Т	Р	R	V	Р	L	R	Р	< 200	297	Symptomatic
Mean $(n=20)$											3195(n=14)	495 (<i>n</i> =13)	
Disrupted minin	nal												
motif P ¹⁰⁴ /P ¹⁰	7												
PL	S	V	V	Р	R	V	Q	L	R	Ι	6 920	127	Asymptomatic
741	Р	V	V	Р	R	V	Ō	L	R	Е	<200	NA	Asymptomatic
794	Р	V	Т	Р	R	Т	ò	L	R	Е	<200	NA	Asymptomatic
1139-1	S	V	R	Р	Κ	V	ò	L	R	А	< 200	NA	Asymptomatic
1147-1	A	V	Т	Р	Κ	V	ò	L	R	А	2 2 5 9	407	Asymptomatic
1215	S	V	Т	Р	R	V	ò	L	R	G	< 200	430	Asymptomatic
1268	S	V	Т	P	R	V	ò	L	R	A	3 7 3 0	529	Asymptomatic
1567	P	v	Ť	P	R	v	õ	Ľ	R	E	3 4 9 1	1 346	Asymptomatic
1570	S	v	т	P	R	v	õ	T	R	G	< 200	427	Asymptomatic
120	P	v	Ť	s	R	v	P	L.	R	p	NA	NA	Asymptomatic
292	s	v	Ť	S	R	v	p	T	R	A	< 200	NA	Asymptomatic
Mean $(n-11)$	0	v	1	0	л	v	1	г	1	11	(n-10)	544(n-6)	isymptomatic
(n-11)											1000 (n-10)	(n-0)	

relationship between the two sequences (Fig. 1) indicated no difference in *nef* sequence measured at two distinct timepoints. However, some evidence for increased nucleotide distances was observed, relating to sampling times that ranged from 0.6 to 1.5% for 6 months to 1 year (sequences 1147-1/1147-2, 1139-1/1139-2 and 281-1/281-2), 1.6-2.4%at 3 years (LF-1/LF-2 and MP-1/MP-2) and 3.4% for a 5 year break between sampling (117-1/117-2). These findings (approximately 0.5-1% per year) are similar to values estimated previously (Shankarappa *et al.*, 1999; Pieniazek *et al.*, 1999).

Clones isolated from three patients with disparate *nef* sequences were also sequenced: one AIDS patient (sample 117-2), one asymptomatic individual (sample MP-2) and one asymptomatic individual (1268) with disruptions in both proline-rich and di-leucine regions. No major variation in *nef* sequence, either synonymous or non-synonymous, was identified in individual clones compared to bulk sequence derived from a 'swarm' of proviruses and overall, the cloned data confirmed the observed changes in *nef* sequence to be present as majority sequence (Fig. 3).

DISCUSSION

The natural variability of *nef* and its relation to putative nef function in HIV-2-infected individuals, where there is a higher proportion of long-term survivors compared to HIV-1, is not well documented. One previous study of HIV-2 nef sequences reported a higher frequency (10–15%) of disrupted nef alleles in asymptomatic HIV-2-infected individuals (Switzer et al., 1998), implying a role for nef in the pathogenesis of HIV-2 infection. In this cross-sectional analysis of HIV-2-infected individuals living in Portugal, we found little evidence of major nef disruption, akin to the situation in HIV-1, whereas more discrete changes in nef sequence likely to impact on specific aspects of *nef* function were identified. All HIV-2 nef sequences were of a single subtype (A), reflecting the demographical links with Guinea Bissau (Grassly et al., 1998) and contrasting with the wide spectrum of nef sequences from HIV-1 group M (Jubier-Maurin et al., 1999). Interestingly, two of the six disrupted sequences described by Switzer et al. (1998) were HIV-2 subtype B infections, not subtype A, originating from the Ivory Coast where HIV-2 subtype B is prevalent (Pieniazek et al., 1999). Hence, truncation as a feature of nef disruption per se seems unlikely to account for the higher proportion of attenuated HIV-2 phenotypes observed frequently, particularly if the infecting virus belongs to subtype A. Similarly, LTR structures containing core promoter and enhancer elements important in gene regulation and virus expression (Clark *et al.*, 1995; Leiden *et al.*, 1992; Markovitz *et al.*, 1992) were also intact with a subtype A identity and conserved in vivo (Berry et al., 2001).

Studies of SIV_{MAC} have indicated that subtle differences in *nef* can alter dramatically the *in vivo* characteristics of virus replication (Whatmore *et al.*, 1995) and a broad body of data exists relating the structure of HIV-1 *nef* with functional



Fig. 3. Phylogenetic tree of *nef* sequences derived from 20 individual clones (CL) obtained from three HIV-2-infected individuals (117, 1268 and MP), five reference HIV-2 subtype A sequences (HIV-2_{BEN}, HIV-2_{ALI}, HIV-2_{ISY}, HIV-2_{ROD} and HIV-2_{ST}) and three reference sequences for subtype B (HIV-2_{EHO}, HIV-2_{ALT/D205} and HIV-2_{UC1}). SIV_{MM239} was used as the outgroup sequence. Data obtained from bulk sequence analysis (117-1, 281-1, 1268 and MP-1). The suffix 'dil' indicates sequences derived from an original 1/10 dilution of template DNA. The majority of clones from patients 117 and MP were obtained from the second time-point (117-2 and MP-2). Bootstrap values above 80%, based on 100 replicates, are indicated at the branching nodes. Evolutionary distances were estimated using the Kimura two-parameter model and genetic relationships were determined using the neighbourjoining method. Bar, 10% sequence divergence.

characteristics (Geyer *et al.*, 2001; Piguet & Trono, 1999). Sequence-specific motifs involved in key functions of *nef*, such as enhancement of virus infectivity, downregulation of CD4 and interaction with T-cell signalling pathways, have the potential to invoke a variety of cellular pathways that can have an impact on pathogenesis. Hence, we embarked on a more systematic study of the HIV-2 *nef* sequence based on known properties of HIV-1 or SIV *nef*.

Key features of *nef* function, such as the myristoylation signal, were absolutely conserved among HIV-2 sequences, re-enforcing the notion that *N*-myristoylation is critical for the biological activity of all *nef* proteins (Harris, 1995).

Analysis of motifs in HIV-2 *nef* indicated no difference in N-terminal tyrosine-based sorting signals, with no additional tyrosine residues, such as $RQ \rightarrow YE$ (YE *nef*), which, under certain circumstances, have been linked to acute lymphocyte activation and pathogenicity of some strains of SIV involving SH2 interactions (Luo & Peterlin, 1997). Interestingly, interactions of SIV *nef* with human Lck or Hck do not seem to be mediated via the consensus proline motif, as is the case for HIV-1 *nef* (Collette *et al.*, 1996), as SIV *nef* demonstrates the ability to bind Lck and Hck SH2 domains, suggesting the multiple mechanisms by which *nef* is able to bind to and regulate Src kinases (Greenway *et al.*, 1999). The role of Src family kinases and the interaction with either SH2 or SH3 remains to be clarified for HIV-2.

In this respect, the natural variability of proline sequences that could mediate putative SH3 interactions was an important finding in this study. This variability was wider than that recognized previously and appeared to be related, at least in part, to disease status. There was a significant correlation between a $P^{107} \rightarrow Q$ mutation at the centre of the $P^{104}xxPLR$ core and an asymptomatic phenotype, which formed one end of a wide spectrum of proline configurations. Though numbers were relatively small, an increased number of proline residues was accompanied by increased virus load. However, tetra-proline (PxxP)₃ sequences in three asymptomatic individuals were also identified and RNA levels in 5/6 patients with $(PxxP)_2$ motifs were below the level of detection. Moreover, the wide spectrum of PxxP configurations encompassed all theoretical combinations ranging from (PxxP)₃ to P¹⁰⁴xxQ. This situation contrasts markedly with HIV-1 where, even in studies of long-term survivors, the consensus tetra-proline motif remains preserved, with no difference in proline motif configurations between rapid progressors and long-term survivors (Kirchhoff et al., 1999).

Interestingly, in studies of SIV-infected macaques, there appears to be a strong selection pressure for a functional SH3-binding ligand in vivo, associated with kinase interactions (Khan et al., 1998), although other studies cast doubt on the functional significance of the PxxP motif in leading to disease induction during the acute phase (Lang *et al.*, 1997). However, the latter study seems unlikely to be representative of the broader picture of disease development in simian AIDS. Moreover, studies in the TgCD4 mouse model for HIV-1 nef have identified the PxxP region to be a crucial determinant of pathogenicity (Hanna et al., 2001). The natural variability of proline-rich sequences in these HIV-2infected individuals is compelling, given the potential for this region to interact with a variety of cellular partners and to modulate virus replication at the level of the T-cell receptor signalling environment (Fackler et al., 2001). The fact that a $P^{107} \rightarrow Q$ mutation was found only in asymptomatic individuals who also had low levels of peripheral viral RNA suggests a role for PxxP-SH3 interactions in influencing pathogenesis. Hence, one could envisage a situation where the secondary structure of the polyproline type II helix is stabilized by the additional proline residues,

with the critical P¹⁰⁴xxP¹⁰⁷ providing the important hydrophobic contacts with binding of the SH3 domain. The trend for increasing plasma virus load with increasing numbers of proline residues may reflect this and disruption of the binding consensus would have a major impact on putative SH3 interactions.

Immunological control also seems more effective for HIV-2 than HIV-1 (Whittle et al., 1998), with lower levels of immune activation and reduced levels of apoptosis (Michel et al., 2000) reflecting the low virus steady-state of many HIV-2 infections (Andersson et al., 2000; Ariyoshi et al., 2000; Berry et al., 1998, 2002; Popper et al., 1999). Recent data have indicated a role for *nef* in apoptotic regulatory pathways (Geleziunas et al., 2001; Wolf et al., 2001; Fackler & Baur, 2002) and aspects of cell signalling that may have a bearing on virulence (Simmons et al., 2001). Functional data are required to determine which cellular signalling pathways may be involved and define a more precise role for HIV-2 nef PxxP-SH3 interactions. Whether changes in a single region or gene such as *nef* might be expected to influence pathogenicity directly in the face of multiple virus-host interactions, including immunological considerations, is an interesting question. Certainly, the natural variation observed in the PxxP region was unexpected and would suggest that this region may be one of the more relevant regions for HIV-2 nef function in vivo. Molecular clones derived from this study are currently being taken forward to address some of these issues in order to provide a fuller understanding of the role of *nef* in the pathogenesis of HIV-2 infections.

ACKNOWLEDGEMENTS

We would like to thank clinical colleagues Antonio Silva-Graca, Luis Caldeira, Manuel Pinheiro and Jorge Cardoso for their collaboration in this study.

REFERENCES

Almond, N., Jenkins, A., Slade, A., Heath, A., Cranage, M. & Kitchin, P. (1992). Population sequence analysis of a simian immunodeficiency virus (32H reisolate of SIV_{mac251}): a virus stock used for international vaccine studies. *AIDS Res Hum Retroviruses* 8, 77–88.

Andersson, S., Norrgren, H., da Silva, Z., Biague, A., Bamba, S., Kwok, S., Christopherson, C., Biberfeld, G. & Albert, J. (2000). Plasma viral load in HIV-1 and HIV-2 singly and dually infected individuals in Guinea-Bissau, West Africa: significantly lower plasma virus set point in HIV-2 infection than in HIV-1 infection. *Arch Intern Med* **160**, 3286–3293.

Ariyoshi, K., Jaffar, S., Alabi, A. & 7 other authors (2000). Plasma RNA viral load predicts the rate of CD4 T cell decline and death in HIV-2-infected patients in West Africa. *AIDS* 14, 339–344.

Bell, I., Ashman, C., Maughan, J., Hooker, E., Cook, F. & Reinhart, T. A. (1998). Association of simian immunodeficiency virus Nef with the T-cell receptor (TCR) ζ chain leads to TCR down-modulation. *J Gen Virol* **79**, 2717–2727.

Berry, N., Ariyoshi, K., Jobe, O., N'Gom, P. T., Corrah, T., Wilkins, A., Whittle, H. & Tedder, R. (1994). HIV type 2 proviral load measured

by quantitative polymerase chain reaction correlates with CD4⁺ lymphopenia in HIV type 2-infected individuals. *AIDS Res Hum Retroviruses* **10**, 1031–1037.

Berry, N. J., Ariyoshi, K., Jaffar, S., Sabally, S., Corrah, T., Tedder, R. & Whittle, H. (1998). Low peripheral blood viral HIV-2 RNA in individuals with high CD4 percentage differentiates HIV-2 from HIV-1 infection. *J Hum Virol* 1, 457–468.

Berry, N. J., Ariyoshi, K., Balfe, P., Tedder, R. & Whittle, H. (2001). Sequence specificity of the human immunodeficiency virus type 2 (HIV-2) long terminal repeat U3 region *in vivo* allows subtyping of the principal HIV-2 viral subtypes A and B. *AIDS Res Hum Retroviruses* 17, 263–267.

Berry, N. J., Jaffar, S., Schim van der Loeff, M. & 9 other authors (2002). Low level viremia and high CD4% predict normal survival in a cohort of HIV type 2-infected villagers. *AIDS Res Hum Retroviruses* 18, 1167–1173.

Bock, P. J. & Markovitz, D. M. (2001). Infection with HIV-2. *AIDS* 15 (suppl. 5), S35–S45.

Bonfield, J. K., Smith, K. F. & Staden, R. A. (1995). A new DNA sequence assembly program. *Nucleic Acids Res* 23, 4992–4999.

Boom, R., Sol, C. J., Salimans, M. M., Jansen, C. L., Wertheim-van Dillen, P. M. & van der Noordaa, J. (1990). Rapid and simple methods for purification of nucleic acids. *J Clin Microbiol* 28, 495–503.

Bresnahan, P. A., Yonemoto, W. & Greene, W. C. (1999). SIV Nef protein utilizes both leucine- and tyrosine-based protein sorting pathways for down-regulation of CD4. *J Immunol* 163, 2977–2981.

Carl, S., Iafrate, A. J., Lang, S., Stolte, N., Stahl-Hennig, C., Matz-Rensing, K., Fuchs, D., Skowronski, J. & Kirchhoff, F. (2000a). Simian immunodeficiency virus containing mutations in N-terminal tyrosine residues and in the PxxP motif in Nef replicates efficiently in rhesus macaques. *J Virol* 74, 4155–4164.

Carl, S., Daniels, R., lafrate, A. J., Easterbrook, P., Greenough, T. C., Skowronski, J. & Kirchhoff, F. (2000b). Partial 'repair' of defective *nef* genes in a long-term nonprogressor with human immunodeficiency type 1 infection. *J Infect Dis* 181, 132–140.

Chen, Z., Luckay, A., Sodora, D. L. & 9 other authors (1997). Human immunodeficiency virus type 2 (HIV-2) seroprevalence and characterization of a distinct HIV-2 genetic subtype from the natural range of simian immunodeficiency virus-infected sooty mangabeys. *J Virol* **71**, 3953–3960.

Cheng, H., Hoxie, J. & Parks, W. P. (1999). The conserved core of human immunodeficiency virus type 1 Nef is essential for association with Lck and for enhanced viral replication in T-lymphocytes. *Virology* **264**, 5–15.

Clark, N. M., Hannibal, M. C. & Markovitz, D. M. (1995). The peri- κ B site mediates human immunodeficiency virus type 2 enhancer activation in monocytes but not in T cells. *J Virol* **69**, 4854–4862.

Clavel, F., Guetard, D., Brun-Vezinet, F. & 9 other authors (1986). Isolation of a new human retrovirus from West African patients with AIDS. *Science* **233**, 343–346.

Collette, **Y. (1997)**. Towards a consensus for a role of Nef in both viral replication and immunomodulation. *Res Virol* **148**, 23–30.

Collette, Y., Dutartre, H., Benziane, A., Ramos-Morales, F., Benarous, R., Harris, M. & Olive, D. (1996). Physical and functional interaction of Nef with Lck. HIV-1 Nef-induced T-cell signaling defects *J Biol Chem* 71, 6333–6341.

Collette, Y., Arold, S., Picard, C., Janvier, K., Benichou, S., Benarous, R., Olive, D. & Dumas, C. (2000). HIV-2 and SIV Nef proteins target different Src family SH3 domains than does HIV-1 **Craig, H. M., Pandori, M. W. & Guatelli, J. C. (1998).** Interaction of HIV-1 Nef with the cellular dileucine-based sorting pathway is required for CD4 down-regulation and optimal viral infectivity. *Proc Natl Acad Sci U S A* **95**, 11229–11234.

Deacon, N. J., Tsykin, A., Solomon, A. & other authors (1995). Genomic structure of an attenuated quasi species of HIV-1 from a blood transfusion donor and recipients. *Science* **270**, 988–991.

Echarri, A., Gonzalez, M. E. & Carrasco, L. (1997). The N-terminal Arg-rich region of human immunodeficiency virus types 1 and 2 and simian immunodeficiency virus Nef is involved in RNA binding. *Eur J Biochem* 246, 38–44.

Fackler, O. T. & Baur, A. S. (2002). Live and let die: Nef functions beyond HIV replication. *Immunity* 16, 493–497.

Fackler, O. T., Wolf, D., Weber, H. O. & 9 other authors (2001). A natural variability in the proline-rich motif of Nef modulates HIV-1 replication in primary T cells. *Curr Biol* **11**, 1294–1299.

Felsenstein, J. (1989). Phylogenies from molecular sequences: inference and reliability. *Annu Rev Genet* 22, 521–565.

Foti, M., Mangasarian, A., Piguet, V., Lew, D. P., Krause, K.-H., Trono, D. & Carpentier, J.-L. (1997). Nef-mediated clathrin-coated pit formation. *J Cell Biol* 139, 37–47.

Gao, F., Yue, L., Robertson, D. L. & 9 other authors (1994). Genetic diversity of human immunodeficiency virus type 2: evidence for distinct sequence subtypes with differences in virus biology. *J Virol* 68, 7433–7447.

Geffin, R., Wolf, D., Muller, R., Hill, M. D., Stellwag, E., Freitag, M., Sass, G., Scott, G. B. & Baur, A. S. (2000). Functional and structural defects in HIV type 1 *nef* genes derived from pediatric long-term survivors. *AIDS Res Hum Retroviruses* 16, 1855–1868.

Geleziunas, R., Xu, W., Takeda, K., Ichijo, H. & Greene, W. C. (2001). HIV-1 Nef inhibits ASK1-dependent death signalling providing a potential mechanism for protecting the infected host cell. *Nature* **410**, 834–838.

Geyer, M., Fackler, O. T. & Peterlin, B. M. (2001). Structure-function relationships in HIV-1 Nef. *EMBO Rep* 21, 580–585.

Grassly, N. C., Xiang, Z., Ariyoshi, K., Aaby, P., Jensen, H., van der Loeff, M. S., Dias, F., Whittle, H. & Breuer, J. (1998). Mortality among human immunodeficiency virus type 2-positive villagers in rural Guinea-Bissau is correlated with viral genotype. *J Virol* 72, 7895–7899.

Greenberg, M. E., Bronson, S., Lock, M., Neumann, M., Pavlakis, G. N. & Skowronski, J. (1997). Co-localization of HIV-1 Nef with the AP-2 adaptor protein complex correlates with Nef-induced CD4 down-regulation. *EMBO J* **16**, 6964–6976.

Greenway, A. L., Dutartre, H., Allen, K., McPhee, D. A., Olive, D. & Collette, Y. (1999). Simian immunodeficiency virus and human immunodeficiency virus type 1 Nef protein show distinct patterns and mechanisms of Src kinase activation. *J Virol* 73, 6152–6158.

Hanna, Z., Kay, D. G., Rebai, N., Guimond, A., Jothy, S. & Jolicoeur, P. (1998). Nef harbours a major determinant of pathogenicity for an AIDS-like disease induced by HIV-1 in transgenic mice. *Cell* 95, 163–175.

Hanna, Z., Weng, X., Kay, D. G., Poudrier, J., Lowell, C. & Jolicoeur, P. (2001). The pathogenicity of human immunodeficiency virus (HIV) type 1 Nef in CD4C/HIV transgenic mice is abolished by mutation of its SH3-binding domain, and disease development is delayed in the absence of Hck. *J Virol* 75, 9378–9392.

Harris, M. (1995). The role of myristoylation in the interactions between the human immunodeficiency virus type 1 Nef and cellular proteins. *Biochem Soc Trans* 23, 557–561.

Jubier-Maurin, V., Saragosti, S., Perret, J.-L. & 7 other authors (1999). 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. *AIDS Res Hum Retroviruses* 15, 23–32.

Karn, T., Hock, B., Holtrich, U., Adamski, M., Strebhardt, K. & Rubsamen-Waigmann, H. (1998). Nef proteins of distinct HIV-1 or -2 isolates differ in their binding properties for HCK: isolation of a novel Nef binding factor with characteristics of an adaptor protein. *Virology* 246, 45–52.

Kestler, H. W., III Ringler, D. J., Mori, K., Panicali, D. L., Sehgai, P. K., Daniel, M. D. & Desrosiers, R. C. (1991). Importance of the *nef* gene for maintenance of high virus loads and for development of AIDS. *Cell* **65**, 651–662.

Khan, I. H., Sawai, E. T., Antonio, E., Weber, C. J., Mandell, C. P., Montbriand, P. & Luciw, P. A. (1998). Role of the SH3-ligand domain of simian immunodeficiency virus Nef in interaction with Nefassociated kinase and simian AIDS in rhesus macaques. *J Virol* 72, 5820–5830.

Kirchhoff, F., Easterbrook, P. J., Douglas, N., Troop, M., Greenough, T. C., Weber, J., Carl, S., Sullivan, J. L. & Daniels, R. S. (1999). Sequence variations in human immunodeficiency virus type 1 Nef are associated with different stages of disease. *J Virol* 73, 5497–5508.

Lang, S. M., lafrate, A. J., Stahl-Hennig, C. & 7 other authors (1997). Association of simian immunodeficiency virus Nef with cellular serine/threonine kinases is dispensable for the development of AIDS in rhesus macaques. *Nat Med* **3**, 860–865.

Learmont, J. C., Geczy, A. F., Mills, J. & 9 other authors (1999). Immunologic and virologic status after 14 to 18 years of infection with an attenuated strain of HIV-1. *N Engl J Med* 340, 1715–1722.

Lee, C.-H., Leung, B., Lemmon, M. A., Zheng, J., Cowburn, D., Kuriyan, J. & Saksela, K. (1995). A single amino acid in the SH3 domain of Hck determines its high affinity and specificity in binding to HIV-1 Nef protein. *EMBO J* 14, 5006–5015.

Leiden, J. R., Wang, C.-Y., Petryniak, B., Markovitz, D. M., Nabel, G. J. & Thompson, C. B. (1992). A novel Ets-related transcription factor, Elf-1, binds to human immunodeficiency virus type 2 regulatory elements that are required for inducible trans activation in T cells. *J Virol* 66, 5890–5897.

Lock, M., Greenberg, M. E., lafrate, A. J., Swigut, T., Muench, J., Kirchhoff, F., Shohdy, N. & Skowronski, J. (1999). Two elements target SIV Nef to the AP-2 clathrin adaptor complex, but only one is required for the induction of CD4 endocytosis. *EMBO J* 18, 2722–2733.

Luo, W. & Peterlin, M. (1997). Activation of the T-cell receptor signaling pathway by Nef from an aggressive strain of simian immunodeficiency virus. *J Virol* 71, 9531–9537.

Mangasarian, A., Piguet, V., Wang, J.-K., Chen, Y.-L. & Trono, D. (1999). Nef-induced CD4 and major histocompatibility complex class I (MHC-I) down-regulation are governed by distinct determinants: N-terminal α helix and proline repeat of Nef selectively regulate MHC-I trafficking. *J Virol* **73**, 1964–1973.

Manninen, A., Hiipakka, M., Vihinen, M., Lu, W., Mayer, B. J. & Saksela, K. (1998). SH3-domain binding function of HIV-1 Nef is required for association with a PAK-related kinase. *Virology* 250, 273–282.

Markovitz, D., Smith, M., Hilfinger, J., Hannibal, M. C., Petryniak, B. & Nabel, G. J. (1992). Activation of the human immunodeficiency virus type 2 enhancer is dependent on purine box and κ B regulatory elements. *J Virol* 66, 5479–5484.

Marlink, R., Kanki, P., Thior, I. & other authors (1994). Reduced rate of disease development after HIV-2 infection as compared to HIV-1. *Science* 265, 1587–1590.

Michael, N. L., Chang, G., d'Arcy, L., Ehrenberg, P. K., Mariani, R., Busch, M. P., Birx, D. L. & Schwartz, D. H. (1995). Defective accessory genes in a human immunodeficiency virus type-1-infected long-term survivor lacking recoverable virus. J Virol 69, 4228–4236.

Michel, P., Balde, A. T., Roussilhon, C., Aribot, G., Sarthou, J.-L. & Gougeon, M.-L. (2000). Reduced immune activation and T cell apoptosis in human immunodeficiency virus type 2 compared with type 1: correlation of T cell apoptosis with β_2 microglobulin concentration and disease evolution. *J Infect Dis* **181**, 64–75.

Mourich, D. V., Lee, S., Reyes-Teran, G., Mackewicz, C. E. & Levy, J. A. (1999). Lack of differences in *nef* alleles among HIV-infected asymptomatic long-term survivors and those who progressed to disease. *AIDS Res Hum Retroviruses* 15, 1573–1575.

National AIDS Commission (2002). HIV/AIDS infection, situation in Portugal at 30 June 2002, doc. 128. National Institute of Health, Lisbon, Portugal: CVEDT.

Pieniazek, D., Ellenberger, D., Janini, L. M. & 11 other authors (1999). Predominance of human immunodeficiency virus type 2 subtype B in Abidjan, Ivory Coast. *AIDS Res Hum Retroviruses* 15, 603–608.

Piguet, V. & Trono, D. (1999). The Nef protein of primate lentiviruses. *Rev Med Virol* 9, 111–120.

Piguet, V., Chen, Y.-L., Mangasarian, A., Foti, M., Carpentier, J.-L. & Trono, D. (1998). Mechanisms of Nef-induced CD4 endocytosis: Nef connects CD4 with the mu chain of adaptor complexes. *EMBO J* 17, 2472–2481.

Piguet, V., Gu, F., Foti, M., Demaurex, N., Gruenberg, J., Carpentier, J.-L. & Trono, D. (1999). Nef-induced CD4 degradation: a diacidicbased motif in Nef functions as a lysosomal targeting signal through the binding of β-COP in endosomes. *Cell* **97**, 63–73.

Popper, S. J., Sarr, A. D., Travers, K. U., Gueye-Ndiaye, A., Mboup, S., Essex, M. E. & Kanki, P. J. (1999). Lower human immunodeficiency (HIV) type 2 viral load reflects the difference in pathogenicity of HIV-1 and HIV-2. *J Infect Dis* 180, 1116–1121.

Poulsen, A. G., Aaby, P., Larsen, O., Jensen, H., Naucler, A., Lisse, I. M., Christiansen, C. B., Dias, F. & Melbye, M. (1997). 9-year HIV-2-associated mortality in an urban community in Bissau, west Africa. *Lancet* 349, 911–914.

Reeves, J. D. & Doms, R. W. (2002). Human immunodeficiency virus type 2. J Gen Virol 83, 1253–1265.

Saksela, K., Cheng, G. & Baltimore, D. (1995). 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* **14**, 484–491.

Schim van der Loeff, M. F. & Aaby, P. (1999). Towards a better understanding of the epidemiology of HIV-2. *AIDS* 13 (suppl. A), S69–S84.

Schulz, T. F., Whitby, D., Hoad, J. G., Corrah, T., Whittle, H. & Weiss, R. A. (1990). Biological and molecular variability of human immunodeficiency virus type 2 isolates from The Gambia. *J Virol* 64, 5177–5182.

Shankarappa, R., Margolick, J. B., Gange, S. J. & 9 other authors (1999). Consistence viral evolutionary changes associated with the progression of human immunodeficiency virus type 1 infection. *J Virol* 73, 10489–10502.

Simmons, A., Aluvihare, V. & McMichael, A. (2001). Nef triggers a transcriptional program in T cells imitating single-signal T cell activation and inducing HIV virulence mediators. *Immunity* 14, 763–777.

Simon, F., Matheron, S., Tamalet, C. & 9 other authors (1993). Cellular and plasma viral load in patients infected with HIV-2. *AIDS* 7, 1411–1417. **Soriano, V., Gomes, P., Heneine, W. & 10 other authors (2000).** Human immunodeficiency virus type 2 (HIV-2) in Portugal: clinical spectrum, circulating subtypes, virus isolation and plasma viral load. *J Med Virol* **61**, 111–116.

Switzer, W. M., Wiktor, S., Soriano, V. & 7 other authors (1998). Evidence of Nef truncation in human immunodeficiency virus type 2 infection. *J Infect Dis* 177, 65–71.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22, 4673–4680.

Whatmore, A. M., Cook, N., Hall, G. A., Sharpe, S., Rud, E. W. & Cranage, M. P. (1995). Repair and evolution of *nef in vivo* modulates simian immunodeficiency virus virulence. *J Virol* 69, 5117–5123.

Whittle, H. A., Morris, J., Todd, J., Corrah, T., Sabally, S., Bangali, J., N'Gom, P. T., Rolfe, M. & Wilkins, A. (1994). HIV-2-infected patients survive longer than HIV-1-infected patients. *AIDS* 8, 1617–1620.

Whittle, H. C., Ariyoshi, K. & Rowland-Jones, S. (1998). HIV-2 and T cell recognition. *Curr Opin Immunol* 10, 382–387.

Wolf, D., Witte, V., Laffert, B., Blume, K., Stromer, E., Trapp, S., d'Aloja, P., Schurmann, A. & Baur, A. S. (2001). HIV-1 Nef associated PAK and P13-kinases stimulate Akt-independent Bad-phosphorylation to induce anti-apoptotic signals. *Nat Med* 7, 1217–1224.

Yamaguchi, J., Devare, S. G. & Brennan, C. (2000). Identification of a new HIV-2 subtype based on phylogenetic analysis of full-length genomic sequence. *AIDS Res Hum Retroviruses* 9, 925–930.