

Sequence Note

Genetic and Phylogenetic Analysis of *env* Subtypes G and H in Central Africa

W. JANSSENS,¹ L. HEYNDRIKX,¹ K. FRANSEN,¹ J. MOTTE,¹ M. PEETERS,¹ J.N. NKENGASONG,¹
P.M. NDUMBE,² E. DELAPORTE,³ J.-L. PERRET,⁴ C. ATENDE,⁵ P. PIOT,¹ and G. VAN DER GROEN¹

PHYLOGENETIC ANALYSIS of HIV-1 strains has resulted in a classification in which different subtypes, or *clades*, are distinguished. Determining which variants are circulating in a given area at a given time is important when considering vaccination trials. To classify new strains phylogenetically, they are compared at the genetic level with a set of "reference" or "probe" strains that are representative for the respective subtypes. To date, there is evidence for at least six *env* subtypes.¹ Although HIV-1 variants of certain subtypes may be prevalent in certain geographic areas, it does not exclude circulation of less prevalent subtypes. Some variants have been reported as "uncertain" or "pending" with respect to their phylogenetic classification.² This may be due to a suboptimal sequence length or lack of related reference strains.¹

Here we report on HIV-1 strains that may be considered as representative of two new *env* subtypes G and H. Four HIV-1 isolates have been studied: two from Gabon (LBV21-7 and VI525), one from Cameroon (CA13), and one from Zaire (VI557). LBV21-7, VI525, and VI557 were chosen for study on the basis of their phylogenetic classification in *gag* subtypes G (LBV21-7) and H (VI557, VI525).^{1,3} This *gag* subtype classification replaced the classification earlier reported by Louwagie et al.³ whereby sequences formerly identified as "gag F" and "gag G" are called "gag G" and "gag H" respectively.¹ CA13 was part of a study on characterization of Cameroonian subtypes.⁴ The full-length HIV-1 *env* gene of VI525 and LBV21-7 was amplified from cultured peripheral blood mononuclear cells; a 900-bp part of the *env* gene encoding V3, V4, V5, and the beginning of gp41 was amplified for VI557 and CA13. Virus culture, polymerase chain reaction, cloning, and sequencing of the 900-bp fragment were performed as described previously.⁵ The nucleotide sequence data were deposited in the EBML, GenBank, and DDBJ Nucleotide Sequence Databases under the following accession numbers: U09664-U09667.

The 4 newly obtained HIV *env* sequences were aligned with 21 known sequences of HIV-1 isolates belonging to different subtypes, and the HIV-1 related chimpanzee isolate SIV_{cpz-gab} on the basis of primary structure. Starting from the alignment, a distance matrix corrected for multiple mutations per site, according to Jukes and Cantor, was constructed. Tree topologies were inferred by neighbor joining. Confidence values for individual branches of the resulting tree were determined by a bootstrap analysis in which 1000 bootstrap trees were generated from resampled data. Distance calculation, tree construction, and bootstrap analysis were realized with the software package TREECON.⁶ A phylogenetic tree based on the common 900-bp *env* fragments was generated. LBV21-7 and VI525, on the one hand, and VI557 and CA13, on the other hand, clustered significantly in two separate subgroups different from those described before (Fig. 1). On the basis of their genetic distances from the other reported subtypes they were assigned as new subtypes. LBV21-7, formerly described as subtype G on the basis of *gag* sequences, is now classified as *env* subtype G together with VI525. Remarkably, VI525 was earlier classified as *gag* subtype H, indicating that this virus is an *env* subtype G-gag subtype H recombinant. VI557, formerly described as *gag* subtype H, is classified as *env* subtype H together with CA13.

The predicted amino acid sequence of the V3 loop region for these strains is presented in Fig. 2. The GPGQ tetrameric tip of the V3 loop in the "global" consensus was also present in LBV21-7 and VI557 of subtypes G and H, respectively. The GTGR motif in VI525 (G) and the GRGQ tip in CA13 are rather unusual. The replacement of proline by threonine or arginine in the respective tetrapeptides disturbs the β turn and may result in an altered epitope in the principal neutralizing domain. Both subtype G isolates have a deletion of one amino acid proximal to the last consensus glycosylation site, which is also ab-

¹Department of Infection and Immunity, Institute of Tropical Medicine, Antwerp, Belgium.

²Virus-Immunology Unit, CUSS, Yaounde, Cameroon.

³INSERM U13/IHEA, Hôpital Claude Bernard, Paris, France.

⁴Centre Hospitalier de Libreville, Gabon.

⁵PNLS, Libreville, Gabon.



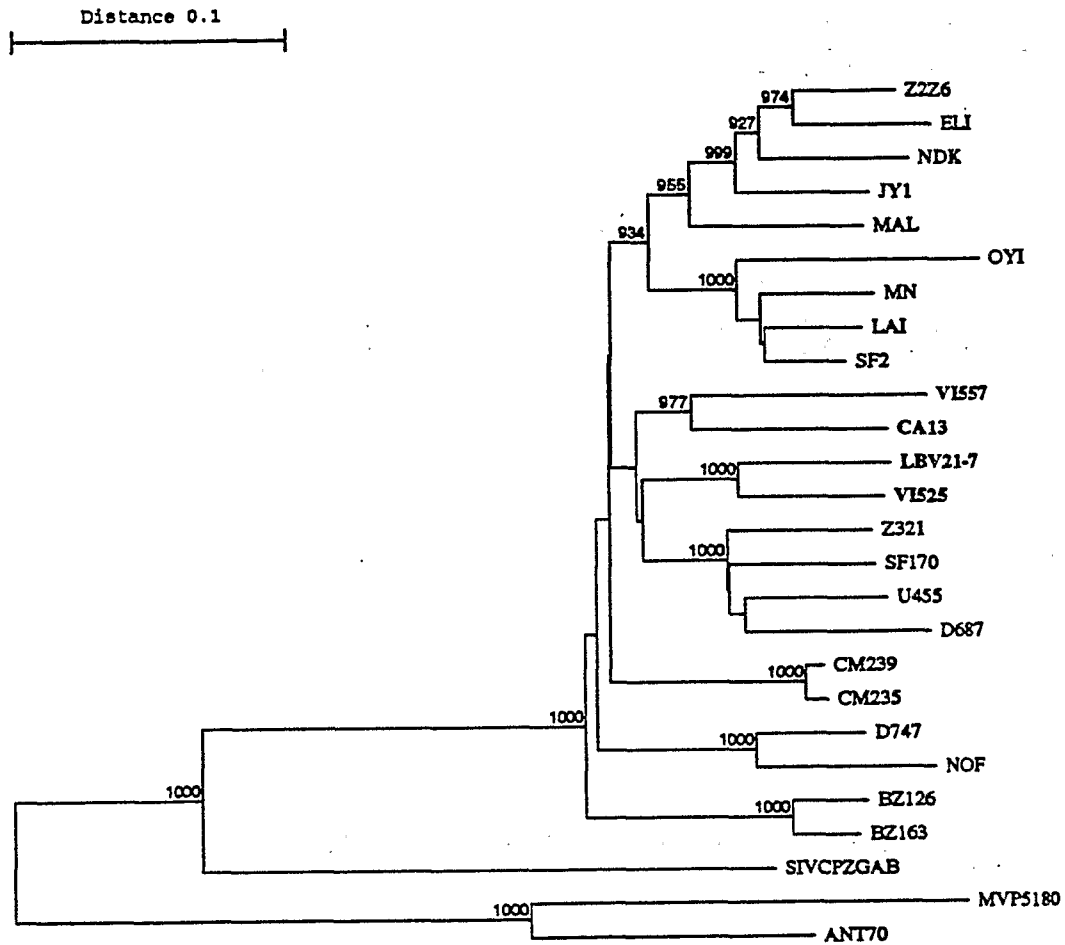


FIG. 1. Phylogenetic tree based on 872 unambiguously aligned positions of HIV-1 *env* gene sequences from 25 different subjects and the simian immunodeficiency virus SIV_{cpz-gab}. The sequences determined in this study are indicated in bold. The tree is unrooted. The distance between two sequences is obtained by summing the lengths of the connecting horizontal branches, using the scale at the top. The number of bootstrap trees out of 1000 replications supporting a particular phylogenetic group in more than 85% is placed alongside the node considered.

sent from the subtype A consensus. The subtype H isolates do not have the third consensus potential N-linked glycosylation site, which is also absent from the subtype A consensus.

To verify the presence of subtype G or H isolates in the Central African Republic, C2V3 sequence data of hereto un-

classified strains from an earlier reported study by Murphy *et al.*² were aligned with the corresponding fragments of all isolates in this study and analyzed phylogenetically. Isolate U4056 was classified with high confidence in subtype H (data not shown).

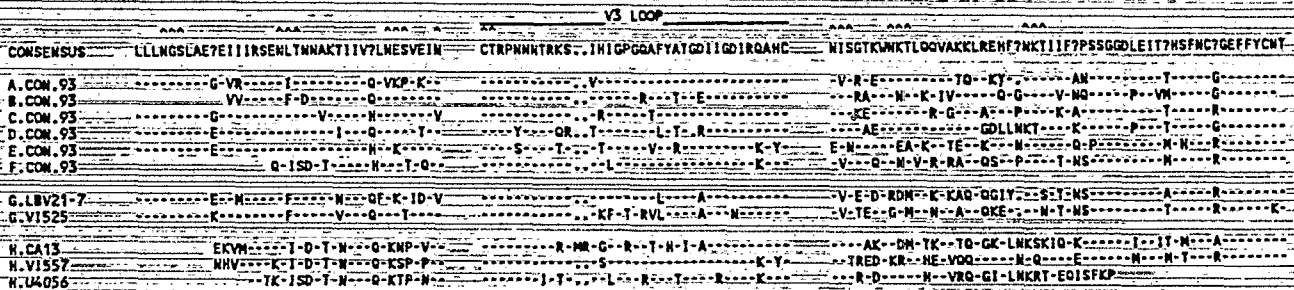


FIG. 2. Amino acid sequence alignment of *env* V3 regions of the "global" consensus and subtype consensus sequences A, B, C, D, E, and F,¹ as compared to the *env* subtype G and H sequences. Amino acid identity between sequences is represented by dashes; points are introduced to align the sequences. The positions of potential N-linked glycosylation sites are shown by carets.

Globally, eight HIV-1 *env* subtypes have now been documented, besides the HIV-1 outliers of clade O, that may eventually be representative for new subtypes.⁷ When more sequences become available, additional subtypes may arise.

ACKNOWLEDGMENTS

We thank Annemie Leonaers for expert technical assistance. The help of Gerald Myers in the phylogenetic analysis is greatly appreciated. This work was supported by Grants Nos. 3-3025-91 and F5/5/25-RN-N24 from the National Fonds voor Wetenschappelijk Onderzoek (Brussels).

REFERENCES

1. Myers G, Korber B, Wain-Hobson S, Smith RF, and Pavlakis GN: *Human Retroviruses and AIDS*. Los Alamos National Laboratory, Los Alamos, New Mexico, 1993.
2. Murphy E, Korber B, Georges-Courbot MC, You B, Pinter A, Cook D, Kiény MP, Georges A, Mathiot C, Barré-Sinoussi F, and Girard M: Diversity of V3 region sequences of human immunodeficiency viruses type 1 from the Central African Republic. *AIDS Res Hum Retroviruses* 1993;9:997-1006.
3. Louwagie J, McCutchan FE, Peeters M, Brennan TP, Sanders-Buell E, Eddy GA, van der Groen G, Fransen K, Gershy-Damet G-M, Deleys R, and Burke DS: Phylogenetic analysis of *gag* genes from seventy international HIV-1 isolates provides evidence for multiple genotypes. *AIDS* 1993;7:769-780.
4. Nkengasong JN, Janssens W, Heyndrickx W, Fransen K, Ndumbe PM, Motte J, Leonaers A, Ngolle M, Ayuk J, Piot P, and van der Groen G: Genotypic subtypes of human immunodeficiency virus type 1 (HIV-1) in Cameroon. *AIDS* 1994 (submitted).
5. Janssens W, Heyndrickx L, Van de Peer Y, Bouckaert A, Fransen K, Motte J, Gershy-Damet G-M, Peeters M, Piot P, and van der Groen G: Molecular phylogeny of part of the *env*-gene of human immunodeficiency virus type 1 strains isolated in Côte d'Ivoire. *AIDS* 1994;8:21-26.
6. Van de Peer Y, and De Wachter R: TREECON: A software package for the construction and drawing of evolutionary trees. *Comput Appl Biosci* 1993;9:177-182.
7. Janssens W, Nkengasong JN, Heyndrickx L, Fransen K, Ndumbe PM, Delaporte E, Peeters M, Perret J-L, Ndoumou A, Atende C, Piot P, and van der Groen G: Further evidence of the presence of genetically very aberrant HIV-1 strains in Cameroon and Gabon. *AIDS* 1994;8:1012-1013.

Address reprint requests to:

Wouter Janssens
 Department of Infection and Immunity
 Division of Microbiology
 Institute of Tropical Medicine
 Nationalestraat 155
 2000 Antwerp, Belgium

Year	Q1	Q2	Q3	Q4	Total
2011	100	100	100	100	400
2012	100	100	100	100	400
2013	100	100	100	100	400
2014	100	100	100	100	400
2015	100	100	100	100	400
2016	100	100	100	100	400
2017	100	100	100	100	400
2018	100	100	100	100	400
2019	100	100	100	100	400
2020	100	100	100	100	400
2021	100	100	100	100	400
2022	100	100	100	100	400
2023	100	100	100	100	400
2024	100	100	100	100	400
2025	100	100	100	100	400
2026	100	100	100	100	400
2027	100	100	100	100	400
2028	100	100	100	100	400
2029	100	100	100	100	400
2030	100	100	100	100	400
2031	100	100	100	100	400
2032	100	100	100	100	400
2033	100	100	100	100	400
2034	100	100	100	100	400
2035	100	100	100	100	400
2036	100	100	100	100	400
2037	100	100	100	100	400
2038	100	100	100	100	400
2039	100	100	100	100	400
2040	100	100	100	100	400
2041	100	100	100	100	400
2042	100	100	100	100	400
2043	100	100	100	100	400
2044	100	100	100	100	400
2045	100	100	100	100	400
2046	100	100	100	100	400
2047	100	100	100	100	400
2048	100	100	100	100	400
2049	100	100	100	100	400
2050	100	100	100	100	400
2051	100	100	100	100	400
2052	100	100	100	100	400
2053	100	100	100	100	400
2054	100	100	100	100	400
2055	100	100	100	100	400
2056	100	100	100	100	400
2057	100	100	100	100	400
2058	100	100	100	100	400
2059	100	100	100	100	400
2060	100	100	100	100	400
2061	100	100	100	100	400
2062	100	100	100	100	400
2063	100	100	100	100	400
2064	100	100	100	100	400
2065	100	100	100	100	400
2066	100	100	100	100	400
2067	100	100	100	100	400
2068	100	100	100	100	400
2069	100	100	100	100	400
2070	100	100	100	100	400
2071	100	100	100	100	400
2072	100	100	100	100	400
2073	100	100	100	100	400
2074	100	100	100	100	400
2075	100	100	100	100	400
2076	100	100	100	100	400
2077	100	100	100	100	400
2078	100	100	100	100	400
2079	100	100	100	100	400
2080	100	100	100	100	400
2081	100	100	100	100	400
2082	100	100	100	100	400
2083	100	100	100	100	400
2084	100	100	100	100	400
2085	100	100	100	100	400
2086	100	100	100	100	400
2087	100	100	100	100	400
2088	100	100	100	100	400
2089	100	100	100	100	400
2090	100	100	100	100	400
2091	100	100	100	100	400
2092	100	100	100	100	400
2093	100	100	100	100	400
2094	100	100	100	100	400
2095	100	100	100	100	400
2096	100	100	100	100	400
2097	100	100	100	100	400
2098	100	100	100	100	400
2099	100	100	100	100	400
2100	100	100	100	100	400