JOURNAL OF VIROLOGY, Mar. 1996, p. 1651–1667 Vol. 70, No. 3
0022-538X/96/$04.00+0
Copyright © 1996, American Society for Microbiology

Molecular Cloning and Analysis of Functional Envelope Genes from Human Immunodeficiency Virus Type 1 Sequence Subtypes A through G

FENG GAO,[1] SANDRA G. MORRISON,[1] DAVID L. ROBERTSON,[2] CHARLOTTE L. THORNTON,[1] STEVENSON CRAIG,[1] GUNILLA KARLSSON,[3] JOSEPH SODROSKI,[3] MARIZA MORGADO,[4] BERNARDO GALVAO-CASTRO,[5] HAGEN VON BRIESEN,[6] SIMON BEDDOWS,[7] JONATHAN WEBER,[7] PAUL M. SHARP,[2] GEORGE M. SHAW,[1] BEATRICE H. HAHN,[1]* AND THE WHO† AND NIAID‡ NETWORKS FOR HIV ISOLATION AND CHARACTERIZATION

Departments of Medicine and Microbiology, University of Alabama at Birmingham, Birmingham, Alabama 35294[1]; Department of Genetics, University of Nottingham, Queens Medical Centre, Nottingham,[2] and Department of Communicable Diseases, Imperial College, St. Mary's Hospital Medical School, London,[7] United Kingdom; Department of Pathology, Dana-Farber Cancer Institute, Boston, Massachusetts 02115[3]; Department of Immunology, Instituto Oswaldo Cruz, Fiocruz, Rio de Janeiro,[4] and Advanced Laboratory of Public Health, Centro de Pesquisa Goncalo Moniz, Fiocruz, Salvador,[5] Brazil; and Chemotherapeutisches Forschungsinstitut, Georg Speyer Haus, Frankfurt am Main, Germany[6]

Received 26 September 1995/Accepted 28 November 1995

Present knowledge of human immunodeficiency virus type 1 (HIV-1) envelope immunobiology has been derived almost exclusively from analyses of subtype B viruses, yet such viruses represent only a minority of strains currently spreading worldwide. To generate a more representative panel of genetically diverse envelope genes, we PCR amplified, cloned, and sequenced complete gp160 coding regions of 35 primary (peripheral blood mononuclear cell-propagated) HIV-1 isolates collected at major epicenters of the current AIDS pandemic. Analysis of their deduced amino acid sequences revealed several important differences from prototypic subtype B strains, including changes in the number and distribution of cysteine residues, substantial length differences in hypervariable regions, and premature truncations in the gp41 domain. Moreover, transiently expressed glycoprotein precursor molecules varied considerably in both size and carbohydrate content. Phylogenetic analyses of full-length *env* sequences indicated that the panel included members of all major sequence subtypes of HIV-1 group M (clades A to G), as well as an intersubtype recombinant (F/B) from an infected individual in Brazil. In addition, all subtype E and three subtype G viruses initially classified on the basis of partial *env* sequences were found to cluster in subtype A in the 3′ half of their gp41 coding region, suggesting that they are also recombinant. The biological activity of PCR-derived *env* genes was examined in a single-round virus infectivity assay. This analysis identified 20 clones, including 1 from each subtype (or recombinant), which expressed fully functional envelope glycoproteins. One of these, derived from a patient with rapid CD4 cell decline, contained an amino acid substitution in a highly conserved endocytosis signal (Y721C), as well as a premature truncation of its gp41 domain, which lacked 17 amino acids. Several other *env* constructs mediated virus entry with very poor efficiency, although they did not contain sequence changes predicted to alter protein function. These results indicate that the *env* genes of primary HIV-1 isolates collected worldwide can vary considerably in their genetic, phylogenetic, and biological properties. The panel of *env* constructs described here should prove valuable for future structure-function studies of naturally occurring envelope glycoproteins as well as AIDS vaccine development efforts targeted against a broader spectrum of viruses.

The failure of natural human immunodeficiency virus type 1 (HIV-1) infection to elicit a protective immune response, the complexities of viral replication and persistence, and the propensity of the virus for genetic change have suggested that the development of an effective AIDS vaccine will be difficult (for reviews, see references 6, 21, 24, and 62). Genetic diversity, in particular, is believed to pose a formidable challenge, since it accumulates rapidly (about 1% per year) and affects the structure, function, and immunogenicity of viral gene products (6, 24, 37). To evaluate the extent of global HIV-1 variation, sequences of virus strains originating from numerous countries have been compared (1, 3–5, 7, 8, 10, 12, 19, 20, 22, 23, 29–35, 41, 43, 47, 52–54, 58, 59, 67, 69, 74, 76, 81, 87, 100). These studies have shown that HIV-1 can be classified into two major groups, designated M and O, which are defined as distinct clusters on phylogenetic trees (52, 70, 90, 91). Group M comprises the great majority of HIV-1 isolates and can be further subdivided into at least nine sequence subtypes or clades, des-

* Corresponding author. Mailing address: Departments of Medicine and Microbiology, University of Alabama at Birmingham, 701 S. 19th St., LHRB 613, Birmingham, AL 35294. Phone: (205) 934-0412. Fax: (205) 934-1580.

† Participants of the WHO Network for HIV Isolation and Characterization who contributed to this study include: S. Osmanov, W. L. Heyward, and J. Esparza (Global Programme on AIDS, World Health Organization, Geneva, Switzerland); B. Galvao-Castro (Fiocruz, Salvador, Brazil); P. van de Perre (National AIDS Control Program, Kigali, Rwanda); E. Karita (Institute of Tropical Medicine, Antwerp, Belgium); S. Sempala, B. Tugume, and B. Biryahwaho (Uganda Virus Virus Research Institute, Entebbe, Uganda); C. Wasi (Siriraj Hospital, Mahidol University, Bangkok, Thailand); H. Rübsamen-Waigmann and H. von Briesen (Georg-Speyer-Haus Chemotherapeutisches For-

Continued on following page

Downloaded from http://jvi.asm.org/ on October 16, 2019 by guest

ignated A to I (41, 53, 54, 70, 90). Group O has been discovered only recently and thus far includes only a small number of viruses from Cameroon and Gabon (8, 23, 52, 97). Both groups are characterized by considerable sequence diversity. For example, Env amino acid sequence variation within group M ranges from 3 to 23% among members of the same subtype and from 25 to 35% among members of different subtypes. Group O viruses exhibit a similar level of diversity but have not been classified into subtypes because of the small number of representatives characterized (52). The two groups, M and O, differ by up to 47% in their Env protein sequences. Finally, recent data indicate that individuals can become coinfected with viruses belonging to different subtypes of group M and that this can result in the production of recombinants (82, 83, 87). Thus, intersubtype recombination can generate yet more viral diversity, with potential consequences for vaccine development.

Given this degree of diversity, it is widely believed that a vaccine based on a single strain or subtype of HIV-1 will be unsuccessful against the larger spectrum of globally circulating HIV-1 variants (62). Most vaccine formulations currently in clinical trials contain immunogens derived exclusively from subtype B viruses, the predominant genotype in the United States and Europe (21). By contrast, little emphasis has been placed on the development of candidate vaccines against non-subtype B viruses, although these cause the vast majority of HIV-1 infections in developing countries, where they continue to spread extremely rapidly (e.g., subtype A in central Africa [100], subtype C in India [22, 30], and subtype E in Thailand [44, 99]). One reason for the emphasis on subtype B vaccines has been the lack of well-characterized virologic reagents corresponding to the other clades, many of which have been discovered only recently (e.g., subtypes F, G, H, and I [12, 32, 41]). Another reason is that even for subtype B viruses, the correlates of protective immunity remain unknown. In the absence of these data, it is impossible to predict which combinations of viral antigens from which genetic subtypes are likely to produce the broadest immunity.

Given the increasing importance of genetically diverse viral strains in the global AIDS pandemic, it is clear that a comprehensive vaccine strategy must include non-subtype B viruses. In particular, clones encoding functional *env* gene products are required for the generation of subtype-specific immunological reagents, recombinant vectors for protein expression, and DNA- or protein-based subunit vaccines. Most full-length envelope sequences in the HIV database are derived from viruses adapted to grow in immortalized T-cell lines (71), raising the possibility that their structure-function relationships are not biologically or clinically germane (66). Moreover, most envelopes available from peripheral blood mononuclear cell (PBMC)-propagated viruses have not been functionally characterized (71), again raising concerns about the relevance of available clones and sequences for certain biological applications. Because of the growing need for functionally active envelope clones, particularly for non-subtype B viruses, we have generated a comprehensive panel of genetically diverse *env* genes from 35 primary HIV-1 isolates (PBMC cocultures). In this paper, we describe their genetic, phylogenetic, and biological characterization and confirm their function in the context of infectious HIV-1 virions. This panel of recombinant envelope constructs thus represents a unique collection of reagents, which should be useful for investigators interested in HIV-1 envelope biology and vaccine development.

MATERIALS AND METHODS

Virus isolates. The majority of HIV-1 isolates used in this study were obtained from the World Health Organization (WHO)- and National Institutes of Allergy and Infectious Diseases (NIAID)-sponsored Networks for Virus isolation and Characterization and represent viruses collected at potential vaccine evaluation sites (Table 1). Additional isolates were obtained from two acute (symptomatic) seroconvertors (91US005 and 91US006) evaluated at the University of Alabama at Birmingham (UAB), a Zairian patient with end-stage AIDS (93ZR001), an individual (92US657) enrolled in the Multi AIDS Center Study cohort (MACS), and two patients (93UG975 and 92RU131) with known subtype G infections studied at St. Mary's Hospital Medical School in London (SMHMS). Informed consent was obtained from all study subjects prior to blood collection. Epidemiological and clinical information relevant to this study, as well as references describing the various cohorts or individual cases in greater detail, is summarized in Table 1. For consistency, isolates were labelled according to WHO nomenclature (40); some isolates have been reported previously under different names, and these are listed in parentheses. All isolates were established and propagated by cocultivation with normal donor PBMCs (27, 85). For most cultures, the number of in vitro passages was less than five. High-molecular-weight DNA was extracted from expanded PBMC cultures as described previously (57). Preliminary subtype designation of isolates was made on the basis of partial *env* gene sequences (WHO, UAB and SMHMS isolates [4, 19, 34, 100]) or by heteroduplex mobility analysis of PCR-amplified gp120 (V1 to V5) fragments (WHO, NIAID, and MACS isolates [9]).

PCR amplification of full-length gp160 genes. PCR amplification was carried out as described previously (19) with minor modifications. All *env* (gp160) genes were amplified from cultured PBMC DNA. Most isolates required two rounds of PCR to yield sufficient quantities of amplification products for further subcloning (*env* A [5′-GGCTTAGGCATCTCCTATGGCAGGAAGAA-3′] and *env* N [5′-CTGCCAATCAGGGAAGTAGCCTTGTGT-3′] in the first round; *env* B [5′-AGAAAGAGCAGAAGACAGTGGCA**ATGA**-3′] and *env* M, [5′-TAGCCCTTCCAGTCCCCCCTTTTCTTTTA-3′] in the second round). Isolates 92UG037, 92UG975, 93BR019, 93BR020, and 93BR029 required only a single round, with primer pairs *env* B and *env* N. Finally, an entire *tat-rev-vpu-env* cassette was amplified from isolate 92US657 with primers S1 (5′-ATATCTATGAAACTTATGGGGATAC-3′) and *env* N in the first round and S2 (5′-GGCAGGAGTGGAAGCCATAATAAGAA-3′) and *env* M in the second round.

PCRs were carried out with a total volume of 100 μl, containing 10 mM Tris HCl (pH 8.0), 50 mM KCl, 1.5 mM $MgCl_2$, 0.01% (wt/vol) gelatin, 200 μM deoxynucleoside triphosphates, 20 pmol of each primer, and 2.5 U of *Taq* polymerase. Samples were subjected to 35 cycles of a denaturing step at 94°C for 1 min, an annealing step at 55°C for 1 min, and an extension step at 72°C for 6 min (in some instances, annealing and extension temperatures were varied to achieve

Continued from preceeding page
schungsinstitut, Frankfurt, Germany); H. Holmes, A. Newberry, S. Ranjbar, and P. Tomlinson (National Institutes of Biological Standards and Control, London, United Kingdom); J. Bradac (National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Md.); J. I. Mullins (University of Washington, Seattle, Wash.); E. L. Delwart (Aaron Diamond AIDS Research Center, New York, N.Y.); F. Gao and B. H. Hahn (University of Alabama at Birmingham, Birmingham, Ala.); J. Weber, R. Cheingsong-Popov, and P. Kaleebu (St. Mary's Hospital Medical School, London, United Kingdom); G. Myers and B. T. M. Korber (Los Alamos National Laboratory, Los Alamos, N.Mex.).

‡ Participants of the NIAID Network for HIV Isolation and Characterization who contributed to this study include: J. Chiphangwi and T. Taha (Malawi College of Medicine, Blantyre, Malawi); J. Desormeaux (Centres pour le Developpement et la Sante, Port-au-Prince, Haiti); S. Eiumtrakul (The Royal Thai Army Hospital, Chiang Mai, Thailand); C. Natpratan (Royal Thai Ministry of Public Health, Chiang Mai, Thailand); C. Khamboonruang (Chiang Mai University, Chiang Mai, Thailand); P. Miotti, N. A. Halsey, D. Vlahov,and K. E. Nelson (School of Hygiene and Public Health, Johns Hopkins University, Baltimore, Md.); J. Phair (Northwestern University, Chicago, Ill.); Y. Cao, J. P. Moore, and D. D. Ho (Aaron Diamond AIDS Research Center, New York, N.Y.); F. Gao, and B. H. Hahn (University of Alabama at Birmingham, Birmingham, Ala.); M. Matocha, A. Fowler, S. Dilworth, and O. Sharma (AIDS Research and Reference Reagent Program, National Institute of Allergy and Infectious Diseases, Bethesda, Md.); R. Brown, S. Dusing, J. Whitman, D. Hoekzema, and F. Vogel (AIDS Vaccine Reagent Program, National Institute of Allergy and Infectious Diseases, Bethesda, Md.); G. Myers and B. T. M. Korber (Los Alamos National Laboratory, Los Alamos, N.Mex).

Downloaded from http://jvi.asm.org/ on October 16, 2019 by guest

TABLE 1. Epidemiological and clinical information for study isolates

Sequence subtype	Isolate[a]	Sex of patient[b]	Age of patient (yr)	City	Country	Risk factor[c]	Time since seroconversion (mo)	Disease status (CD4 counts/mm^3)[d]	Antiviral therapy	Yr of isolation	Source	Biological phenotype (MT-2 assay)[e]	Reference(s)
A	92RW020	F	33	Kigali	Rwanda	Het	9.7	AS	No	1992	WHO	NA[f]	100
	92UG031	M	27	Nakiwugo	Uganda	Het	NA	AS	No	1992	WHO	NSI	100
	92UG037	F	31	Entebbe	Uganda	Het	NA	AS	No	1992	WHO	NA	100
B	92BR020	M	19	Belo Horizonte	Brazil	Ho	12.6	AS	No	1992	WHO	NSI	100
	92TH014	M	38	Bangkok	Thailand	IVDU	25.6	AS	No	1992	WHO	NSI	100
	92HT593 (301593)	F	23	Port of Prince	Haiti	Het	NA	AS	No	1992	NIAID	NSI	86
	92HT594 (301594)	F	25	Port of Prince	Haiti	Het	NA	AS	No	1992	NIAID	SI	86
	92HT596 (301596)	F	31	Port of Prince	Haiti	Het	NA	AS	No	1992	NIAID	SI	86
	92HT599 (301599)	F	23	Port of Prince	Haiti	Het	NA	AS	No	1992	NIAID	SI	86
	91HT651 (301651)	F	NA	Port of Prince	Haiti	Het	NA	AS	No	1991	NIAID	NA	86
	91HT652 (301652)	F	NA	Port of Prince	Haiti	Het	NA	AS	No	1991	NIAID	NA	86
	92US657 (301657)	M	33	Chicago	United States	Ho	15	AS (1,012)	No	1992	MACS	NA	unpub[g]
	92US711 (301711)	M	44	Baltimore	United States	IVDU	17	AS (853)	No	1992	NIAID	NSI	72
	91US712 (301712)	F	35	Baltimore	United States	IVDU	15	AS (537)	No	1991	NIAID	SI	72
	92US714 (301714)	M	28	Baltimore	United States	IVDU	12	AS (546)	No	1992	NIAID	NSI	72
	92US715 (301715)	M	36	Baltimore	United States	IVDU	20	AS (470)	AZT[h]	1992	NIAID	NSI	72
	92US716 (301716)	M	39	Baltimore	United States	IVDU	4	AS (787)	No	1992	NIAID	NSI	72
	91US006 (HOBR)	M	28	Birmingham	United States	Ho	1	PSI (794)	No	1991	UAB	NSI	77, 78
	91US005 (FASH)	F	17	Birmingham	United States	Het	1	PSI (262)	No	1991	UAB	NSI	77, 78
C	92BR025	M	23	Porto Alegre	Brazil	Hemo	>1.2	AS	No	1992	WHO	NSI	100
	93MW959 (301959)	F	30	Blantyre	Malawi	Het	12	AS	No	1993	NIAID	NSI	60
	93MW960 (301960)	F	28	Blantyre	Malawi	Het	12	AS	No	1993	NIAID	NSI	60
	93MW965 (301965)	F	29	Blantyre	Malawi	Het	12	AS	No	1993	NIAID	NSI	60
D	92UG021	F	36	Entebbe	Uganda	Het	NA	AS	No	1992	WHO	SI	100
	92UG024	F	23	Kireka	Uganda	Het	NA	AS	No	1992	WHO	SI	100
	93ZR001 (AMK)	M	40	NA	Zaire	Het	NA	AIDS (<5)	No	1993	UAB	NA	20
"E"	92TH022	M	23	Bangkok	Thailand	Het	4.5	AS	No	1992	WHO	NSI	100
	93TH966 (301966)	M	21	NA	Thailand	Het	<18	AS	No	1993	NIAID	NSI	73
	93TH975 (301975)	M	21	NA	Thailand	Het	<18	AS	No	1993	NIAID	NSI	73
	93TH976 (301976)	M	21	NA	Thailand	Het	<18	AS	No	1993	NIAID	NA	73
F	93BR020	M	52	Rio de Janeiro	Brazil	Bi	>1	AS (327)	NA	1993	WHO	NA	unpub
	93BR029	M	17	Sao Paulo	Brazil	NA	NA	AS	NA	1993	WHO	NA	unpub
"G"	92UG975 (JW3)	F	44	NA	Uganda	Het	NA	AIDS (20)	AZT	1992	SMHMS	NSI	34
	92RU131 (RU131)	F	3.5	Rostov-on-Don	Russia	Noso	36	AIDS	AZT	1992	SMHMS	NSI	4, 5
F/B[i]	93BR019	M	20	Rio de Janeiro	Brazil	Bi	11.5	AS (515)	NA	1993	WHO	NA	unpub

[a] Isolates were named according to WHO nomenclature (previous designations are listed in parentheses).
[b] M, male; F, female.
[c] Het, heterosexual contact; Ho, homosexual contact; Bi, bisexual contact; IVDU, intravenous drug use; Hemo, hemophiliac patient; Noso, nosocomial infection.
[d] AS, asymptomatic; PSI, primary symptomatic infection.
[e] NSI, non-syncytium inducing; SI, syncytium inducing.
[f] NA, information not available.
[g] unpub, unpublished.
[h] AZT, azidothymidine.
[i] 93BR019 was initially classified as subtype F but later confirmed to represent an F/B intersubtype recombinant. Subtype "E" and "G" viruses were also found to be mosaic (see Results Section).

optimal amplification). PCR products were visualized by agarose gel electrophoresis, purified from gel slices with Geneclean (Bio 101 Inc., La Jolla, Calif.), and subcloned into pCRII by T/A overhang (Invitrogen, San Diego, Calif.). Recombinant clones were identified by colony hybridization and mapped by restriction enzyme digestion of plasmid miniprep DNA.

PCR amplification of partial *env* sequences from uncultured PBMC DNA. A 1,300-bp fragment spanning the C terminus of gp41 was amplified from uncultured PBMC DNA from subject 93BR019 by nested PCR (first round, *env* B7 [5′-GAGCAGCAGGAAGCACTATGGGCGC-3′] and *env* N; second round, *env* B7 and *env* M; 30 cycles of 94°C for 1 min, 40°C for 1 min, and 72°C for 4 min). The amplification product was visualized on an agarose gel, purified with Geneclean, and sequenced directly without interim cloning.

Transient expression of HIV-1 envelope genes in mammalian cells. Full-length pCRII-gp160 clones were analyzed for their functional integrity in a T7-based transient-expression system (19). Briefly, HeLa cells (2×10^5) were seeded in six-well plates and grown to 80 to 90% confluency. Monolayers were infected at a multiplicity of infection of 10 with a recombinant vaccinia virus carrying the T7 RNA polymerase gene (16). Following a 90-min adsorption period, cells were washed with Hanks' balanced salt solution and then transfected with 1.5 μg of pCRII plasmid DNA containing full-length and properly oriented gp160 genes by using the Lipofectin method (Gibco BRL, Gaithersburg, Md.). After 5 h, the medium was changed and HeLa cells were grown in Dulbecco's minimum essential medium plus 10% fetal calf serum overnight. The HeLa cells were then washed twice with Hanks balanced salt solution, starved for 30 min in methionine- and cysteine-free Dulbecco's minimum essential medium (Gibco BRL), labelled with 100 μCi of [^{35}S]methionine-[^{35}S]cysteine (ICN Biomedicals, Inc., Irvine, Calif.) per ml for 30 min at 37°C, and chased with complete medium for 1 h. For gel electrophoresis, cells were lysed with 300 μl of RIPA lysing buffer (50 mM Tris HCl [pH 7.4], 1% Triton X-100, 0.15% sodium dodecyl sulfate [SDS], 150 mM NaCl, 20 mM EDTA) for 5 min at 4°C. Protein A-agarose was reacted with serum from an HIV-1-infected individual overnight at 4°C, washed three times with RIPA buffer, added to labelled cell lysates, and incubated overnight

Downloaded from http://jvi.asm.org/ on October 16, 2019 by guest

at 4°C. Protein A-agarose complexes were pelleted, washed, and resuspended in loading buffer (50 mM Tris-HCl [pH 6.8], 1% SDS, 0.7 M 2-mercaptoethanol, 0.1% bromphenol blue, 10% glycerol), boiled for 2 min, and loaded onto SDS–6 or 8% polyacrylamide gels. The gels were fixed, enhanced for 30 min, dried, and analyzed by autoradiography.

Inhibition of gp160 N-glycosylation by tunicamycin. HeLa cells were infected with recombinant vaccinia virus carrying the T7 gene and transfected with pCRII-gp160 constructs as described above, except for the addition of the glycosylation inhibitor tunicamycin to the culture supernatant (1 μg/ml) 5 h before radiolabelling and immunoprecipitation with serum from an HIV-1-positive patient.

In vitro transcription and translation of HIV-1 *env* genes. In vitro transcription and translation of PCR-derived gp160 genes were performed with the TNT coupled reticulocyte lysate system (Promega, Madison, Wis.) as specified by the manufacturer. Briefly, 1 μg of plasmid DNA (pCRII-gp160 constructs) was added to the TNT lysate (50-μl reaction volume) containing 40 μCi [^{35}S]methionine-[^{35}S]cysteine as well as the appropriate polymerase (T7 or SP6) depending on insert orientation. Following incubation for 1 h at 30°C, 5 μl of reaction mixture was added to 20 μl of RIPA loading buffer, boiled for 2 min, electrophoresed on an SDS-polyacrylamide gel, and analyzed by autoradiography.

Single-round virus infectivity assay. PCR-derived *env* genes (pCRII-gp160s) were cloned into plasmid pSVIII*env*(Kpn) (25) under the control of an HIV-1 long terminal repeat promoter (Fig. 1). This was done by digesting pSVIII *env*(Kpn) with *Kpn*I and by exchanging the HXB2 *env* coding region (except for 36 amino acid residues at the N terminus) with the corresponding *Kpn*I fragments of selected pCRII-gp160 constructs. Five pCRII-gp160 constructs lacked *Kpn*I cloning sites (92RW020.5, MA301959.18, TH301966.8, TH301975.15, and TH301976.17), thus requiring reamplification and introduction of *Kpn*I sites by PCR mutagenesis.

To determine whether the pSVIII-gp160 constructs expressed functionally active envelope genes, Cos-1 cells (2×10^5 per well) were cotransfected with these constructs (5 μg) and HXBH10ΔenvCAT (2.5 μg), an *env*-deficient HIV-1 provirus which contains a chloramphenicol acetyltransferase (CAT) gene in place of the *nef* gene (56). At 16 h following transfection, cells were washed and cultured in complete Dulbecco's minimum essential medium for 2 days at 37°C. Supernatant was then collected, and equivalent amounts of produced virions (based on reverse transcriptase activity) were used to infect 2×10^6 normal donor PBMCs. Following 2 days in culture, PBMCs were pelleted, washed with 10 mM phosphate-buffered saline (pH 7.4), and resuspended in 150 μl of 0.25 M Tris-HCl (pH 7.5). Cells were lysed by repeated freezing and thawing and centrifuged to remove cellular debris, and lysates were assayed for CAT activity with [^{14}C]chloramphenicol as previously described (11). The products were subsequently subjected to thin-layer chromatography, exposed to film, and quantitated with a phosphoimager (Molecular Dynamics) radioanalytical imaging system.

Sequence analysis. Sequence analysis was performed by cycle-sequencing and dye terminator methods with an automated DNA Sequenator (model 373A; Applied Biosystems, Inc.) as recommended by the manufacturer. Individual sequence fragments were assembled by using the Sequencher program (Gene Codes Corp., Ann Arbor, Mich.) and analyzed with Eugene (Baylor College of Medicine, Houston, Tex.), PIMA (92, 93), MASE (13), and DOTS (45). Sequences were determined for both strands of DNA (each base pair was sequenced at least three times). GenBank accession numbers are listed in Table 2.

Phylogenetic analysis. Phylogenetic relationships of the newly characterized viruses were estimated from comparisons of *env* nucleotide sequences with those of previously reported representatives of group M subtypes A to G (71); subtype H and I viruses were not included, since only partial *env* sequences are presently available. Nucleotide sequences were aligned by using CLUSTAL (26) with minor manual adjustments, bearing in mind their predicted protein sequences. Pairwise evolutionary distances were estimated by using Kimura's two-parameter method (36) to correct for superimposed substitutions; sites where there was a gap in any sequence and ambiguous areas within the alignment were excluded from all comparisons. Phylogenetic trees were constructed by the neighbor-joining method (88), and their reliability was estimated from 1,000 bootstrap replicates (14). These methods were implemented by using CLUSTAL W (96). Phylogenetic relationships were also analyzed by the maximum-parsimony method with repeated randomized input order, implemented by using the program DNAPARS from the PHYLIP package (15). The results were the same as those obtained from the neighbor-joining method in all essential aspects.

Recombination breakpoint analysis. All newly determined *env* sequences were analyzed for any evidence of mosaicism generated by intersubtype recombination (82, 83). Phylogenetic trees were constructed from regions (of 300 to 500 bp) within the *env* sequence alignment (available upon request) and examined for discrepant subtype assignments. Viruses exhibiting significantly discordant branching positions in different trees were identified as putative recombinants. To localize the recombination breakpoints within each of these viruses, we examined the linear distribution of phylogenetically informative sites in a four-sequence alignment, including the putative recombinant, a consensus sequence for each of the two subtypes inferred to have been involved in the recombination event, and a simian immunodeficiency virus outgroup ($SIV_{CPZ}GAB$). Breakpoints were inserted at each possible point between adjacent informative sites, and a 2×2 heterogeneity χ^2 value was calculated for the numbers of sites (to either side of the breakpoint) supporting the clustering of the putative recombinant with each of the consensus sequences. The likely breakpoint was identified as that which gave the maximal χ^2 value. To assess the probability of obtaining by chance as high a χ^2 value as that observed, 10,000 simulations were performed in which the informative sites were randomly shuffled.

TABLE 2. Biological activity of PCR-derived envelope clones

Sequence subtype	Clone	PCR-amplified region	GenBank accession no.	Transient expression/ in vitro translation[a]	Infectivity (CAT) assay[b]
A	92RW020.5	gp160	U08794	pos	+++
	92UG031.7	gp160	L34667	pos	ND
	92UG037.8	gp160	U09127	pos	+++
B	92BR020.4	gp160	U08797	pos	++
	92TH014.12	gp160	U08801	pos	++
	92HT593.1	gp160	U08444	pos	+++
	92HT594.10	gp160	U08445	pos	+
	92HT596.4	gp160	U08446	pos	(+)
	92HT599.24	gp160	U08447	pos	(+)
	91HT651.11	gp160	U08441	pos	+
	91HT651.1a	gp160	U08442	pos	ND
	91HT652.11	gp160	U08443	pos	(+)
	92US657.1	*tat*, *rev*, *vpu*, gp160	U04908	pos	ND
	92US711.14	gp160	U08448	pos	ND
	91US712.4	gp160	U08449	pos	+
	92US714.1	gp160	U08450	neg[c]	(+)[c]
	92US715.6	gp160	U08451	pos	+++
	92US716.6	gp160	U08452	pos	ND
	91US006.10	gp160	U27443	pos	+
	91US005.11	gp160	U27434	pos	+++
C	92BR025.9	gp160	U09126	pos	++
	93MW959.18	gp160	U08453	pos	(+)
	93MW960.3	gp160	U08454	pos	(+)
	93MW965.26	gp160	U08455	pos	+++
D	92UG021.16	gp160	U27399	pos	++
	92UG024.2	gp160	U43386	pos	++
	93ZR001.3	gp160	U27419	pos	ND
"E"	92TH022.4	gp160	U09131	pos	ND
	93TH966.8	gp160	U08456	pos	++
	93TH975.15	gp160	U08457	pos	(+)
	93TH976.17	gp160	U08458	pos	+++
F	93BR020.17	gp160	U27401	pos	(+)
	93BR029.2	gp160	U27413	pos	+
"G"	92UG975.10	gp160	U27426	pos	+++
	92RU131.16	gp160	U27445	pos	ND
	92RU131.9	gp160	U30312	pos	ND
F/B	93BR019.4	gp160	U27404	pos	+
	93BR019.10	gp160	U27408	pos	ND

[a] pos, positive; neg, negative.

[b] CAT assay results are derived from nine independent experiments, scoring the highest chloramphenicol conversion rate for each envelope construct. +++, strongly positive (>30% conversion of chloramphenicol to its acetylated forms); ++, positive (10 to 30% conversion); +, weakly positive (1 to 10% conversion); (+), borderline positive (0.3 to 1% conversion); ND, not done.

[c] The subcloning of the 92US714.1 *env* gene into pSVIII*env*(Kpn) eliminated the stop codon in its signal peptide sequence, thus allowing its biological characterization in the CAT complementation assay.

RESULTS

Selection of virus isolates for envelope characterization. A primary objective of this study was to generate and analyze a

Downloaded from http://jvi.asm.org/ on October 16, 2019 by guest

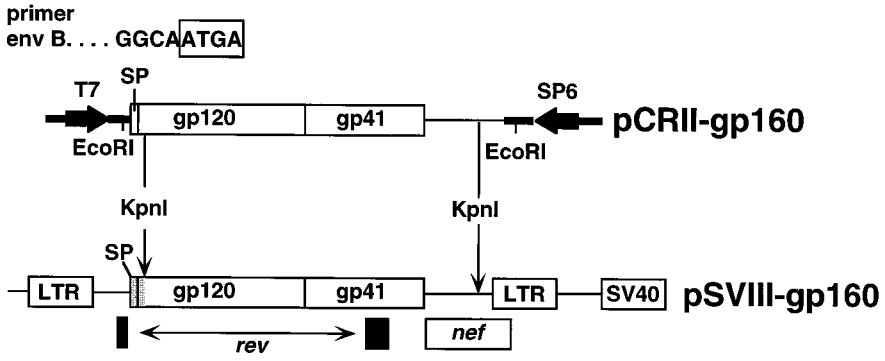


FIG. 1. Schematic representation of different envelope expression plasmids. PCR-derived full-length gp160 sequences were subcloned into pCRII by T/A overhang, generating the pCRII-gp160 constructs (the first four nucleotides of the amplification products are derived from primer *env* B). The positions of T7 and SP6 promoters, as well as the location of the plasmid polylinker (thick line) in relation to the inserted *env* genes, are indicated. Also shown are the positions of conserved *Kpn*I sites, which were used to replace the HXB2 *env* gene in plasmid pSVIIIenv(Kpn) with the corresponding pCRII-gp160 fragments. A 36-amino-acid stretch remaining from the HXB2 *env* gene product is highlighted (light shade). The relative positions of gp120 and gp41 coding regions, *nef*, *rev*, the HIV-1 long terminal repeat (LTR), and the simian virus 40 (SV40) origin of replication are indicated. The composition of the pSVIII*env*(Kpn) plasmid has been described previously (25). SP, signal peptide.

panel of functional HIV-1 envelope constructs that was representative of genetically diverse viruses circulating worldwide. For this purpose, we selected 35 isolates from individuals who acquired their HIV-1 infection relatively recently in the United States, Haiti, Brazil, Thailand, Russia, Rwanda, Uganda, Malawi, and Zaire. These viruses were also known to represent all major sequence subtypes of HIV-1 group M (clades A to G) on the basis of preliminary sequence or heteroduplex gel shift analysis (4, 9, 19, 34, 100). Table 1 summarizes available demographic and clinical information, as well as biological data concerning the isolate phenotype (SI/NSI). Only viruses grown in normal donor PBMCs were selected for analysis. Emphasis was placed on isolates from known seroconvertors (established within 2 years of primary infection), on viruses collected at international vaccine evaluation sites (sponsored by both WHO and NIAID), and on isolates from different risk groups (heterosexual, homosexual, intravenous drug users, etc.). Finally, isolates from Brazil, Thailand, Haiti, and the United States were selected to ensure a wide geographic representation of subtype B viruses. The envelope genes reported in this study are thus representative of naturally occurring HIV-1 infections worldwide and are not biased with respect to in vitro culture methods, mode of transmission, disease association, or geographic origin.

PCR amplification and transient expression of full-length gp160 genes. PCR amplification of complete gp160 genes was performed with primers designed according to HIV-1 group M consensus sequences. Amplification products contained the entire gp160 coding region (as well as adjacent *nef* sequences), except for the first 4 bases of the *env* gene (ATGA) which were derived from the PCR primer. All isolates, regardless of subtype classification, yielded PCR amplification products with the same set of primer pairs. Amplified *env* fragments were purified following gel electrophoresis (to remove smaller amplification products) and ligated into pCRII by T/A overhang. This vector was chosen because it does not require restriction enzyme cleavage for subcloning and contains a T7 promoter, which facilitates expression of properly oriented *env* genes in a transient T7-based expression system (pCRII also contains an SP6 promotor which allows in vitro translation of gp160 genes in the TNT system [Fig. 1]). Restriction enzyme sites in the plasmid polylinker allow excision of the complete *env* gene for additional genetic manipulation.

To identify clones that contained uninterrupted gp160 coding regions, HeLa cells were infected with a recombinant vaccinia virus carrying the bacteriophage T7 RNA polymerase and then transfected with pCRII-gp160 constructs containing full-length (based on restriction enzyme analysis) and properly oriented inserts. Expressed envelope glycoproteins were radiolabelled with [^{35}S]methionine-[^{35}S]cysteine, precipitated with HIV-1 antiserum, and analyzed by polyacrylamide gel electrophoresis. Figure 2 depicts such an analysis for representative pCRII-gp160 constructs (because of the experimental conditions used, only unprocessed gp160 precursor molecules are detected). Expression levels were comparable for all constructs, although the molecular mass of the various gp160 proteins differed somewhat (sizes ranged between 140 and 170 kDa). Using this approach, we identified expression-competent *env* constructs for all isolates, except for 92US714 (expression data are summarized in Table 2). Transient expression of PCR-derived envelope genes thus represents an efficient means to exclude defective gp160 clones prior to sequence analysis.

Nucleotide sequence analysis of pCRII-gp160 constructs.

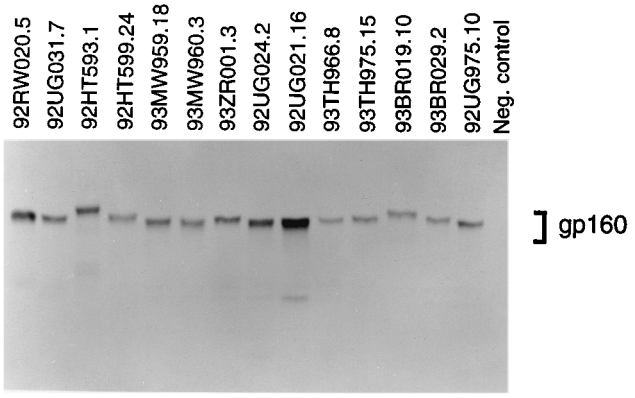


FIG. 2. Expression of recombinant HIV-1 envelope precursor glycoproteins. Plasmid constructs containing full-length and properly oriented *env* genes (pCRII-gp160) were analyzed in a T7-based transient-expression system as described in Materials and Methods. The size range of unprocessed gp160 precursor molecules is indicated (because of the experimental conditions used, only unprocessed gp160 precursor molecules are detectable). The blot includes representative clones from all subtypes.

Downloaded from http://jvi.asm.org/ on October 16, 2019 by guest

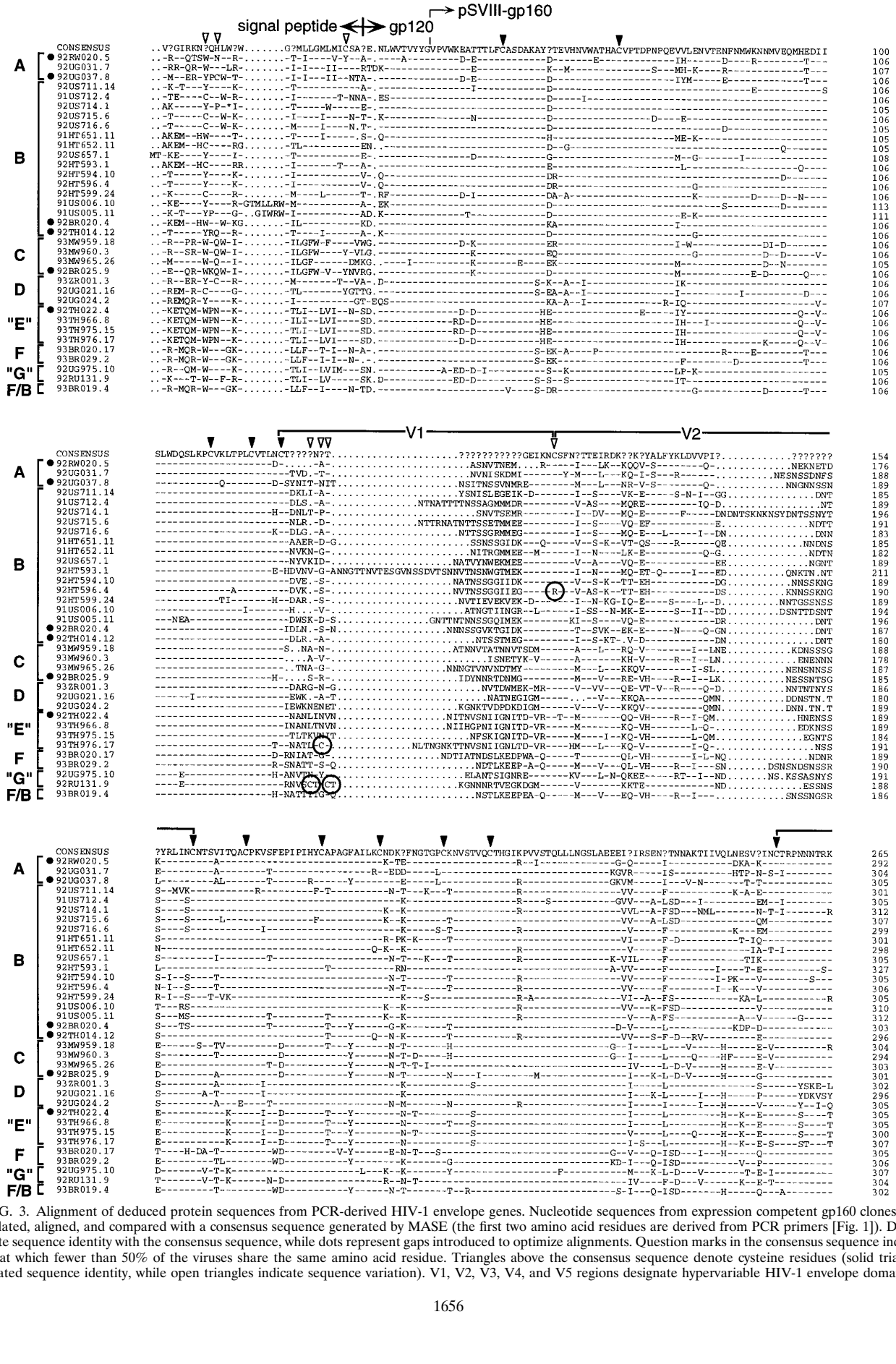
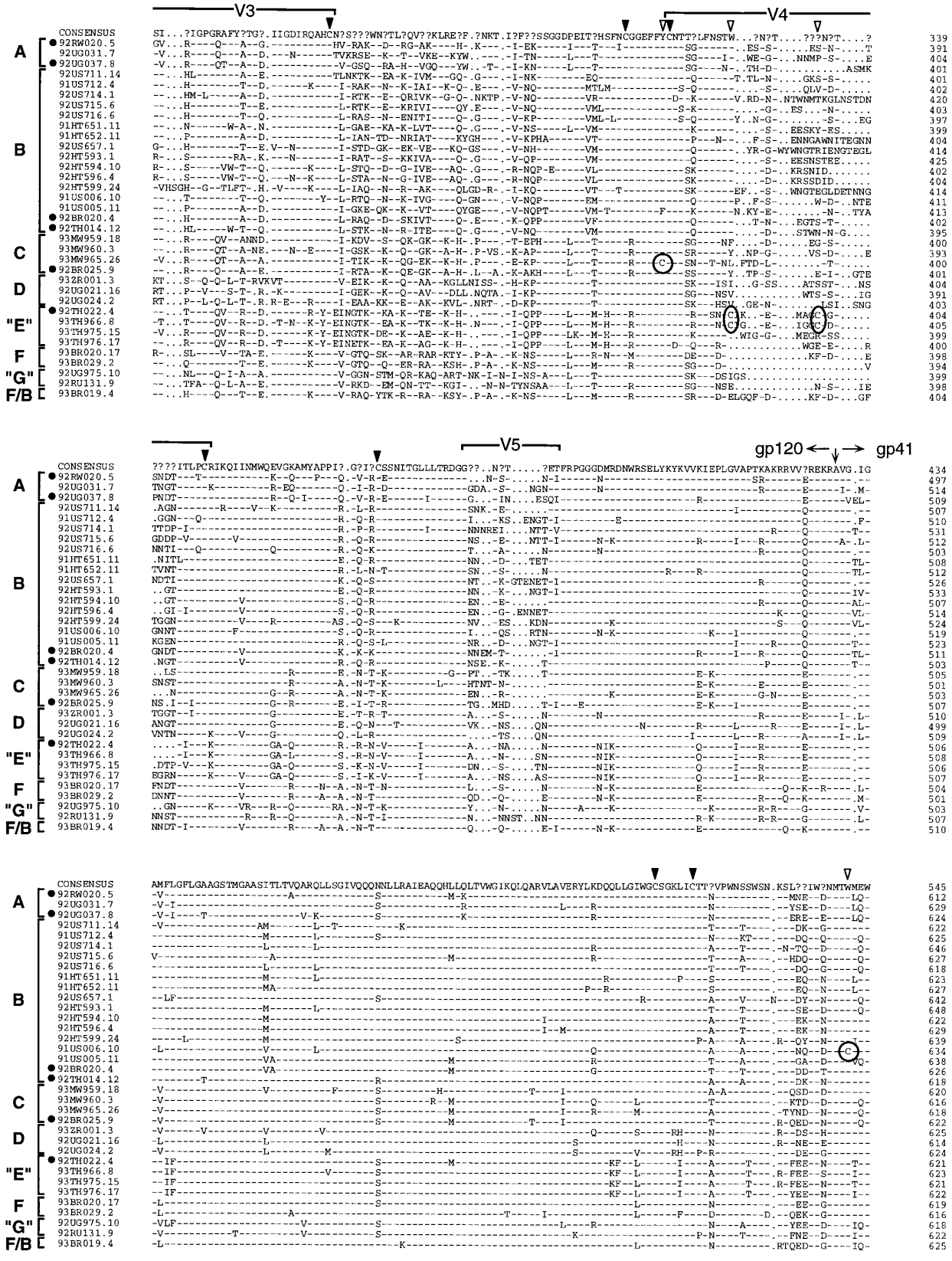


FIG. 3. Alignment of deduced protein sequences from PCR-derived HIV-1 envelope genes. Nucleotide sequences from expression competent gp160 clones were translated, aligned, and compared with a consensus sequence generated by MASE (the first two amino acid residues are derived from PCR primers [Fig. 1]). Dashes denote sequence identity with the consensus sequence, while dots represent gaps introduced to optimize alignments. Question marks in the consensus sequence indicate sites at which fewer than 50% of the viruses share the same amino acid residue. Triangles above the consensus sequence denote cysteine residues (solid triangles indicated sequence identity, while open triangles indicate sequence variation). V1, V2, V3, V4, and V5 regions designate hypervariable HIV-1 envelope domains as

Downloaded from http://jvi.asm.org/ on October 16, 2019 by guest



previously described (61). The signal peptide, the envelope precursor cleavage site, and the position where HXB2 coding sequences are replaced in pSVIII-gp160 constructs are indicated; msd denotes the membrane-spanning domain in gp41; asterisks mark in-frame stop codons; and pound signs indicate translational frameshifts. Black dots at the left of the clone designations indicate previously reported sequences included for comparison. Sequence subtypes are shown on the left margin. Quotation marks indicate the mosaic nature of subtype E and G envelope genes (see Results for details). Open circles highlight altered cysteine residues in the extracellular portions of gp120 and gp41, and a box denotes a mutated endocytosis signal (Y721C) in clone 91US005.11 (89).

Downloaded from http://jvi.asm.org/ on October 16, 2019 by guest

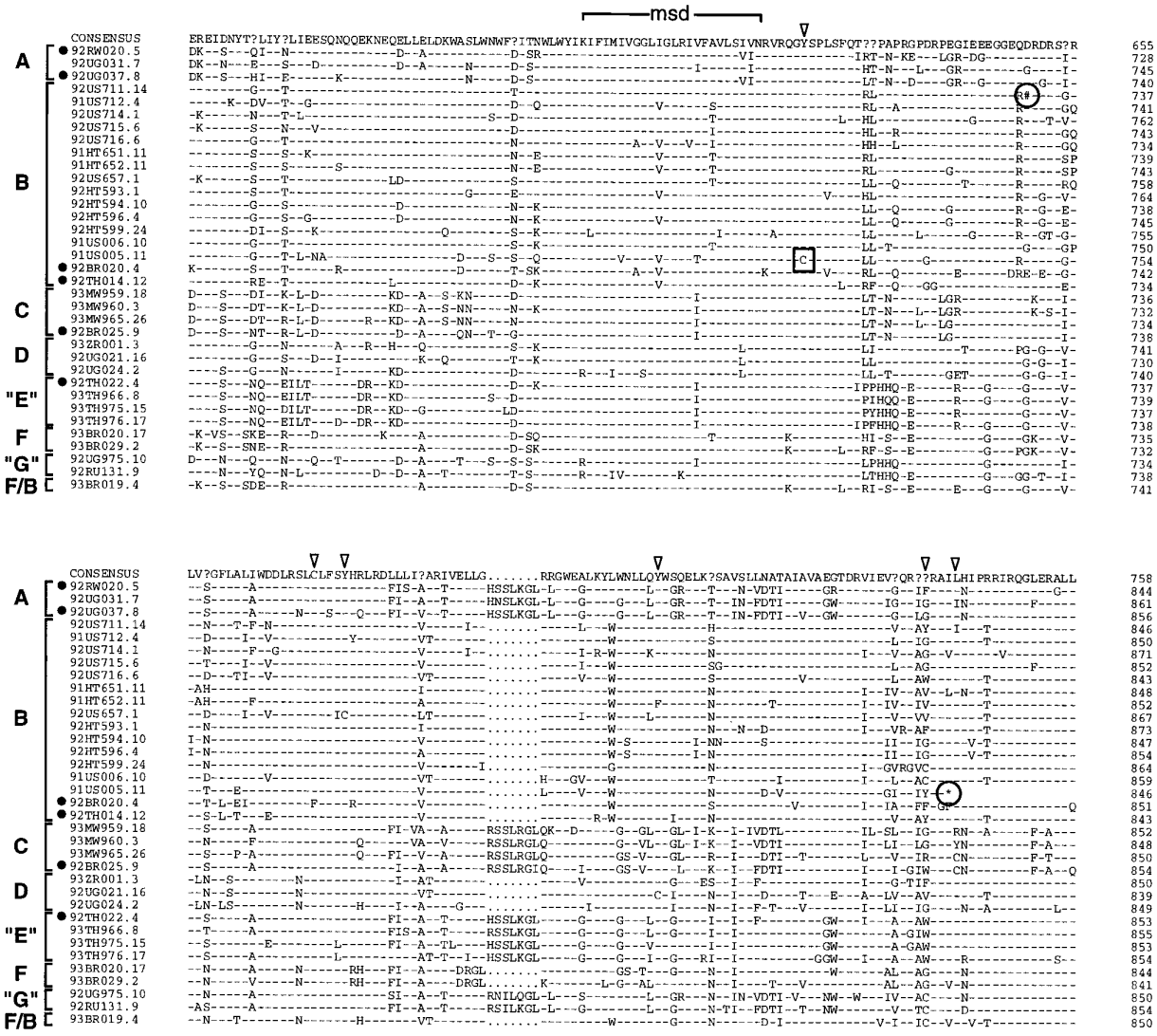


FIG. 3—*Continued.*

For most isolates, one expression-competent pCRII-gp160 construct (or in the case of 92US714, one construct containing a full-length insert) was selected for nucleotide sequence analysis. Two clones were sequenced for isolates 91HT651, 93BR019, and 92RU131 (Table 2). Figure 3 depicts an alignment of their deduced envelope amino acid sequences (solid circles indicate previously reported sequences, which are included for comparison [19]). As expected, all expression competent constructs contained uninterrupted *env* open reading frames. 92US714.1 contained an in-frame stop codon in the signal peptide, thus explaining its failure to yield an expression product. Two other clones, however, were also found to encode genetically altered gp160 genes. 91US005.11 contained a 34-bp deletion followed by an in-frame stop codon (at position 2541), and 92US711.14 contained a frameshift mutation (at position 2190), both resulting in a premature truncation of their gp41 proteins (Fig. 3). These truncations could not be recognized in the transient-expression assay because of the considerable size variation among the different (intact) gp160 precursor proteins (compare Fig. 2).

Sequence alignment also identified potentially important amino acid variations in some clones. Seven envelope sequences exhibited differences in the number and distribution of cysteine residues within their gp120 and extracellular gp41 coding regions (highlighted by circles in Fig. 3). Three of these contained an extra cysteine pair in a hypervariable region (V1 for clone 92RU131.9; V4 for clones 92TH022.4 and 93TH966.8). Four others contained single cysteine changes, including the loss of a highly conserved cysteine between V1 and V2. Since all cysteine residues in subtype B viruses are believed to form disulfide bonds (48), the finding of nonpaired cysteine changes suggests differences in glycoprotein three-dimensional structure and possibly altered biological function. Finally, three clones exhibited unusually long V1 (92HT593.1), V2 (92US714.1), or V3 (92HT599.24) regions, and one clone (93TH975.15) contained an amino acid substitution in a conserved domain (E364K) previously shown to be important for CD4 binding (75).

Phylogenetic analyses. To determine the subtype classification of the newly characterized isolates, phylogenetic trees were constructed from full-length gp160 nucleotide sequences (Fig. 4). All 35 viruses fell into the expected sequence subtypes

Downloaded from http://jvi.asm.org/ on October 16, 2019 by guest

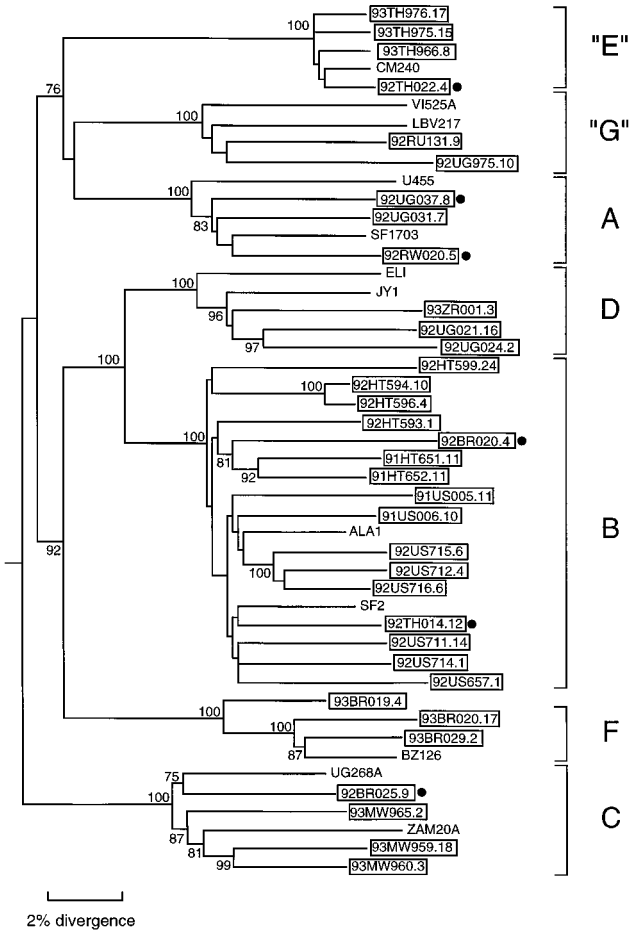


FIG. 4. Phylogenetic relationships of newly characterized, full-length *env* sequences (boxed) with representatives of subtypes A to G (see the text for a description of the method; quotation marks indicate the mosaic nature of subtype E and G envelope genes). Horizontal branch lengths are drawn to scale (the scale bar represents 0.02 nucleotide substitution per site), but vertical separation is for clarity only. Values at nodes indicate the percentage of bootstraps in which the cluster to the right was found; only values of 75% or greater are shown. Black dots indicate viruses whose phylogenetic position has been reported previously (19) but which are included for comparison. Brackets at the right indicate the major sequence subtypes.

consistent with their preliminary classification based on heteroduplex mobility analysis or partial sequence analysis. None of the newly identified viruses were closely related to each other, except for two subtype B Haitian isolates (92HT594 and 92HT596), whose *env* nucleotide sequences differed by only 2%. The other subtype B viruses were approximately equidistantly related and differed on average by 9%, which is typical for epidemiologically unrelated subtype B viruses. Considering the extent of sequence divergence that generally arises within a cultured isolate (the intraisolate variation in this study was <1.5%), it is unlikely that the close relationship of 92HT594 and 92HT596 represents a laboratory contamination (38). Instead, the similarity of the two Haitian isolates most probably indicates that they were derived from individuals with epidemiologically linked infections (definitive conclusions cannot be drawn because of lack of epidemiological information).

The phylogenetic positions of the newly characterized viruses were also examined with respect to their geographic origins. As expected (58, 76, 99), viruses from heterosexual transmission cases in Thailand clustered in subtype E, while a Thai intravenous drug user was infected with a virus clustering in subtype B. Viruses from Brazil clustered within subtypes B, C, and F. Heterosexually infected women from Malawi were infected with viruses belonging to subtype C. Viruses obtained from African nationals (originating from Uganda, Rwanda, and Zaire) clustered in subtypes A, D, or G. All Haitian and U.S. isolates grouped in subtype B. Finally, an isolate from Russia known to have resulted from a nosocomial infection (4, 5) was confirmed to belong to subtype G. Thus, our results are in good agreement with previously reported information concerning the molecular epidemiology of HIV-1 group M (19, 37, 70, 71, 90).

Intersubtype recombination. Because recombination has recently been reported to occur among members of different group M subtypes (82, 83, 87), we examined the PCR-derived *env* sequences more closely for evidence of mosaicism. Phylogenetic trees were constructed from regions within the *env* (and partial *nef*) sequence alignment and examined for discrepant subtype assignments. The results showed that most of the newly characterized strains clustered consistently with respect to their original subtype classification. However, three sets of clones exhibited discordant branching orders. The first was an isolate derived from a Brazilian individual (93BR019), which was a mosaic of subtypes B and F. Approximately 200 bp of sequence from the 3′ region of gp41 clustered in subtype B, but the surrounding sequences, i.e., the remainder of *env* and the 5′ region of *nef*, fell into subtype F. These discordant phylogenetic positions were observed for two different clones (93BR019.4 and 93BR019.10) from the same isolate and were supported by high bootstrap values (Fig. 5A).

To investigate further the recombinant nature of the 93BR019 virus, we obtained uncultured PBMCs from this subject and amplified viral sequences directly without interim culture. The sequence of a 1,217-bp fragment spanning the mosaic gp41 region was determined and aligned with 93BR019.4 and 93BR019.10. This analysis showed that sequences amplified from cultured and uncultured patient material were very closely related to each other (1% divergence) and that they exhibited the same discordant positions in phylogenetic trees (Fig. 5A). Importantly, PCR products were sequenced directly without prior subcloning (93BR019.dir) to characterize the most predominant proviral form present in vivo.

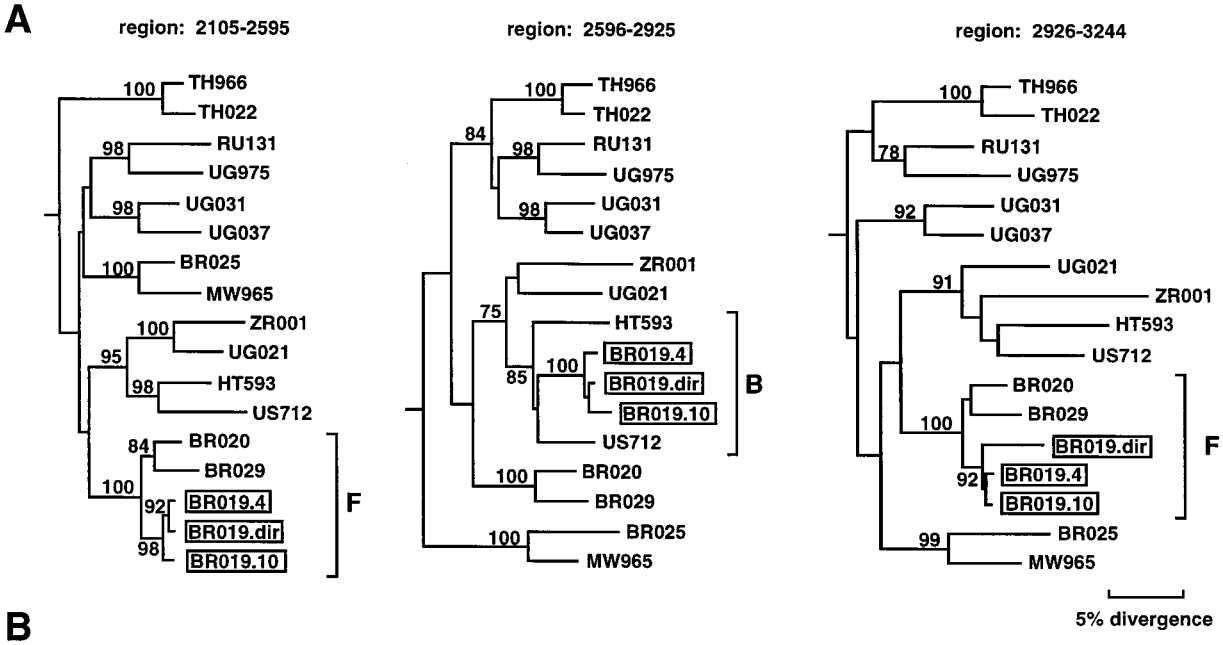
To map the recombination breakpoints more precisely, we examined the distribution of phylogenetically informative sites along the amplified gp41 fragment. As shown in Fig. 5B, sites supporting the grouping of 93BR019.dir with subtype B (designated 1 under the alignment) clustered in a 198-bp region immediately 5′ to the C terminus of gp41. Sites supporting the grouping of 93BR019.dir with subtype F (designated 2) clustered 5′ and 3′ to this region. Maximum χ^2 analysis indicated the likely breakpoints which were found to be statistically significant ($P < 0.0001$) in simulation studies. Similar analyses of *env* sequences from the cultured isolate (performed on clone 93BR019.4) placed breakpoints at the same positions (Fig. 5C). Taken together, these results indicate that subject 93BR019 was infected with an intersubtype recombinant composed of F and B sequences.

A second example of apparent mosaicism involved isolates 92UG975 and 92RU131. Sequences from these viruses clustered with isolates classified as subtype G (GA-VI525 and LBV21-7 [53, 54]) for almost the entire *env* coding region (nucleotides 1 to 2450 of the *env* consensus sequence) (Fig. 6A) but with subtype A at the 3′ end of gp41 (nucleotides 2451 to 2925) (Fig. 6B). Interestingly, similar discordant phylogenetic positions were observed for LBV21-7 (Fig. 6), which represents one of only two complete "subtype G" envelope

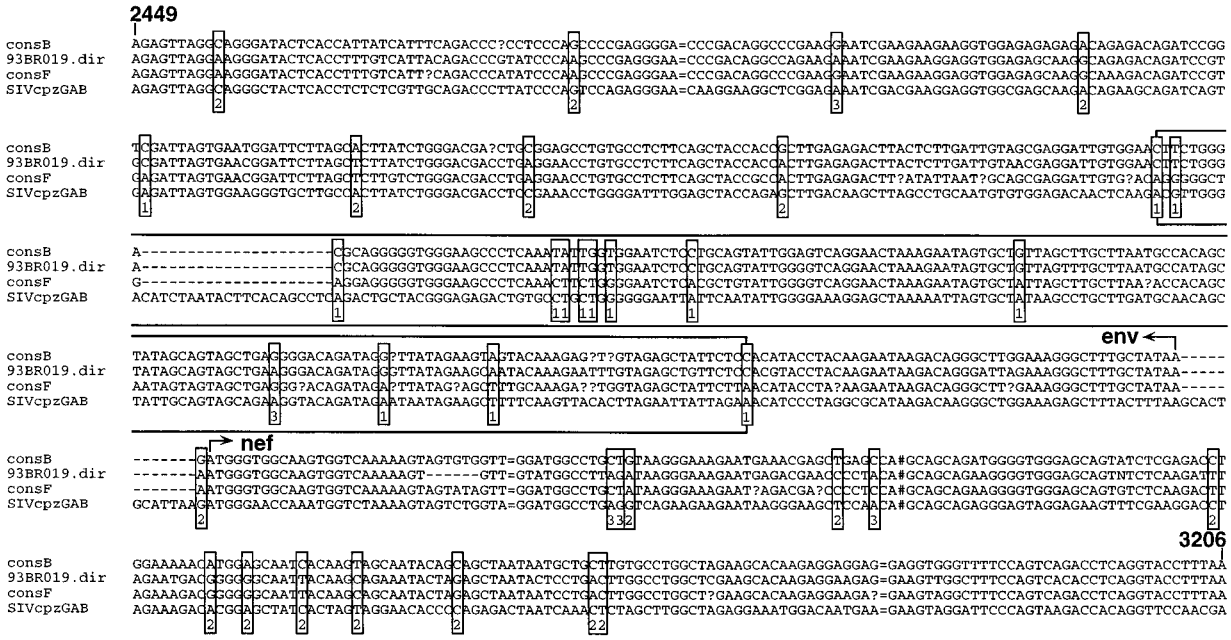
Downloaded from http://jvi.asm.org/ on October 16, 2019 by guest

Downloaded from http://jvi.asm.org/ on October 16, 2019 by guest

A



B



C

Isolate	Gene	Subtype	Region	Informative sites: subtype B 1	Informative sites: subtype F 2	Informative sites: outgroup SIVcpz	P value
93BR019.4 (cultured)	env	F	1-2642	11	77	11	< 0.0001
	env	B	2683-2878	14	0	2	< 0.0001
	nef	F	2938-3247	1	13	0	
93BR019.dir (uncultured)	env	F	1965-2642	4	21	4	< 0.0001
	env	B	2683-2878	13	0	1	< 0.0001
	nef	F	2938-3244	0	12	3	

FIG. 5. Mosaic nature of isolate 93BR019. (A) Phylogenetic relationships of 93BR019 sequences derived from cultured (.4 and .10) and uncultured (.dir) patient material to representatives of subtypes A to G, as derived from three different regions of the *env* alignment. Only subtypes in which the 93BR019 sequences fall are indicated. See Fig. 4 for subtype classifications of other sequences and for more details about the phylogenies. Both initial digits (denoting the year of isolation) and clone numbers have been dropped from each sequence name for clarity. (B) Four sequence alignment showing phylogenetically informative sites (boxed) supporting the clustering of 93BR019.dir with (1) the subtype B consensus sequence, (2) the subtype F consensus sequence, or (3) the outgroup ($SIV_{CPZ}GAB$). The alignment extends from positions 2449 to 3206 of the envelope consensus (complete alignment available upon request; the GenBank accession number for 93BR019.dir is U27444). Equals signs indicate sites with gaps in all four sequences shown (but bases in other sequences from other subtypes); pound signs indicate sites with gaps in three of the sequences shown but a long string of undefined bases in the subtype F consensus sequence. The end of the *env* gene and the start of the *nef* gene are depicted. The long box indicates the region defined as subtype B by the breakpoint analysis (C). (C) Recombination breakpoint analysis. Numbers of phylogenetically informative sites, in different regions of the alignment, supporting the clustering of 93BR019 sequences with each of the three other sequences in the alignment (B). The regions were chosen as those giving the maximum heterogeneity χ^2 value for sites of types 1 and 2 between adjacent regions. The *P* values indicate probabilities of finding this heterogeneity by chance, derived from 10,000 simulations.

sequences currently in the database (71). The other, GA-VI525, is also recombinant (as a whole), because its *env* sequences cluster in subtype G whereas its *gag* sequences cluster in subtype H (32, 83). In the 3′ region of *env*, GA-VI525 remained distinct from other subtypes (Fig. 6B). However, since there is no sequence for this region from any subtype H virus, we cannot be certain that this phylogenetic position of GA-VI525 reflects that its entire *env* gene is of subtype G origin (this 3′ region of *env* could, like *gag*, be of subtype H origin). Nevertheless, the former explanation appears to be much simpler, and we have thus performed breakpoint analysis (as above for 93BR019) for 92UG975 and 92RU131 by using GA-VI525 as the subtype G "consensus." These analyses identified a significant ($P < 0.0001$) breakpoint at about position 2260 in both 92UG975 and 92RU131 (i.e., 39 codons upstream of the region encoding the membrane-spanning domain in gp41), leading us to conclude that both are probably G/A recombinants. A similar analysis for LBV21-7 placed a (significant) breakpoint at around position 2380. However, we cannot exclude the possibility that the breakpoints for all three of these "subtype G" viruses are in fact the same, since there are only a small number of informative sites in this region and only a single sequence was used to construct the subtype G "consensus." If the breakpoints were the same, this would imply that the recombination event had occurred in a common ancestor of the three viruses.

Finally, discordant branching orders were also seen for all three new subtype E envelope sequences. These sequences clustered closely with CM240, a prototypic subtype E virus also from Thailand. As in the example above, these subtype E viruses formed a distinct cluster for most of their *env* coding region but grouped with subtype A at the 3′ end of gp41 (Fig. 6). All "subtype E" viruses from Thailand have been found to be closely related, presumably reflecting a founder event (70, 90). Also, for all those that have been examined, their *gag* sequences fall within subtype A, as if the founder virus was an A/E recombinant (70, 90). Our finding of subtype A-like sequences at the 3′ end of subtype E *env* genes is consistent with this hypothesis. However, we could not use the recombination breakpoint analysis described above for these viruses because

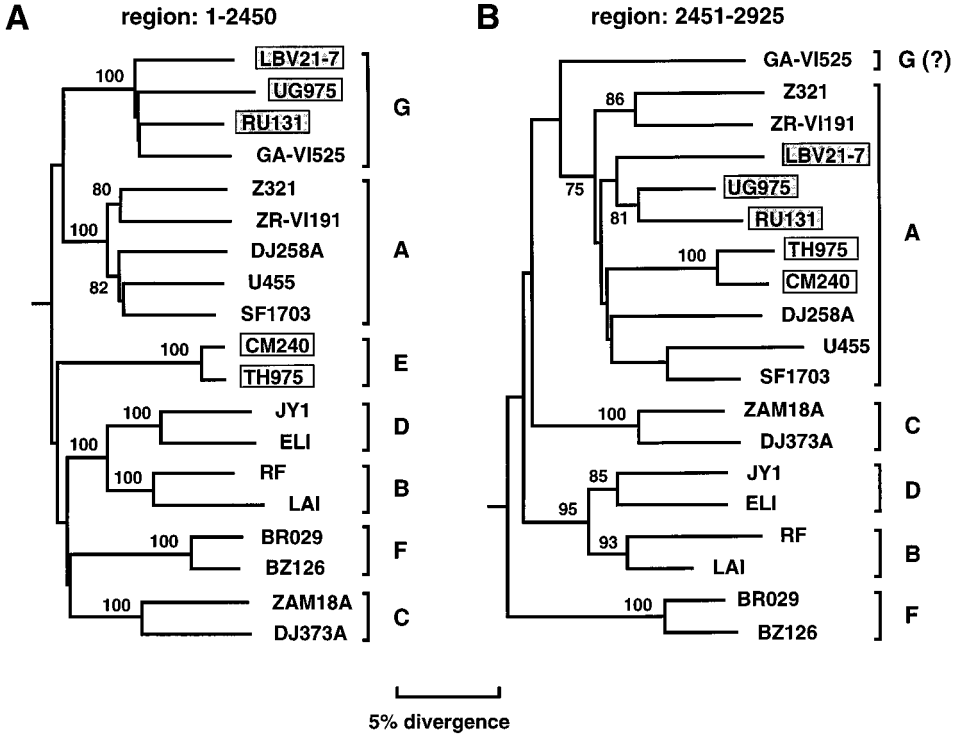


FIG. 6. Discordant phylogenetic relationships of the newly characterized "subtype E" and "subtype G" sequences derived from two different regions of the *env* alignment. (A) Region 1 to 2450. (B) Region 2451 to 2925. See Fig. 4 for additional details. The question mark next to subtype G in panel B indicates that the subtype classification of GA-VI525 in this region of the *env* gene remains uncertain.

Downloaded from http://jvi.asm.org/ on October 16, 2019 by guest

no sequence is yet available for any virus known to be a full-length representative of subtype E. Thus, as with the putative G/A recombinants 92UG975 and 92RU131, definitive analysis must await additional sequence information.

Functional characterization of PCR-derived envelope clones. To assess the biological activity of the PCR-derived *env* genes in the context of HIV-1 virions, we used a single-round infection assay previously described for structure-function analysis of mutant envelope clones (25). The assay is based on the coexpression of a candidate *env* glycoprotein (under the control of an HIV-1 long terminal repeat promoter) with an HIV-1 proviral clone which is defective in *env* but contains the CAT gene in place of the *nef* gene (25, 56). Infectivity of in *trans* complemented virions is measured by quantifying the amount of CAT activity present in newly infected target cells. This assay thus examines biosynthesis, transport, packaging, fusogenicity, and uncoating and scores positive only if the in *trans* complemented envelope glycoproteins perform all of these functions efficiently.

Using a highly conserved *Kpn*I site (located 12 codons downstream of the region encoding the signal sequence cleavage site), we subcloned the majority of the PCR-derived envelope genes into the vector pSVIII*env*(Kpn) (compare Fig. 1). The resulting constructs (pSVIII-gp160) were then individually cotransfected with pHXBH10ΔenvCAT (56), and equivalent amounts of progeny virions (based on reverse transcriptase activity) were used to infect phytohemagglutinin-stimulated normal donor lymphocytes. Clones were tested in nine independent experiments, with representative CAT activity data depicted in Fig. 7. The results identified 20 of 28 envelope constructs tested to yield positive CAT activity in all assays (Table 2). Eight of these constructs were strongly positive (>30% conversion of chloramphenicol to its acetylated forms), six were moderately positive (10 to 30% conversion), and six were weakly positive (1 to 10% conversion). The remaining eight *env* constructs scored weakly positive in some but not all experiments (designated borderline positive in Table 2). Since the CAT-based complementation assay depends on efficient cotransfection, an occasional negative result may indicate suboptimal transfection levels. However, the borderline-positive clones were tested in multiple independent experiments (Table 2) and thus appear to have partial biological activity. Nevertheless, their ability to facilitate virus entry into new target cells appears to be greatly diminished.

Expression of envelope glycoproteins in the presence of tunicamycin. Because the envelope glycoprotein of a subtype D virus (HIV-1_{NDK}) was recently reported to differ from subtype B *env* glycoproteins in size and carbohydrate content (103), we examined whether our panel of envelopes exhibited similar differences and whether such variations occurred in a subtype-specific manner. For this purpose, we first expressed the pCRII-gp160 constructs in the presence of tunicamycin. This compound is known to inhibit N-linked but not O-linked glycosylation. As shown in Fig. 8a, envelope precursor proteins varied in size but not according to their subtype classification. Similar results were obtained following in vitro translation (Fig. 8b), which prevents all forms of glycosylation. These results thus indicate that the size of the envelope precursor backbone as well as the extent of its glycosylation can differ considerably among genetically divergent HIV-1 strains. However, this variation does not appear to correlate with the sequence subtype.

PCR amplification of a *tat-rev-env* cassette from isolate 92US657. Because a number of vector systems with potential utility in vaccine development require Tat and Rev proteins for efficient envelope glycoprotein expression, we tested the possibility of amplifying complete *tat-rev-env* cassettes from a selected number of HIV-1 isolates. By using newly designed sense primers upstream of the first exon of *tat* (see Materials and Methods), amplification products were obtained for most viruses. However, subsequent subcloning into pCRII proved to be extremely difficult. On average, several hundred colonies had to be screened to identify one clone with a full-length insert. Moreover, most of these inserts were found to encode defective *env* genes when analyzed in the T7-based transient-expression system. Using amplification products from three different subtype B isolates and multiple rounds of ligation and transformation, we finally identified one clone from one isolate (92US657.1) to contain a complete *tat-rev-env* cassette with an intact envelope open reading frame (compare Table 2). Nucleotide sequence analysis confirmed the integrity of *env*, *tat*, and *rev* genes but also revealed a frameshift mutation in *vpu* (data not shown). These results indicate that PCR amplification of *tat-rev-env* cassettes from HIV-1 isolates is possible in

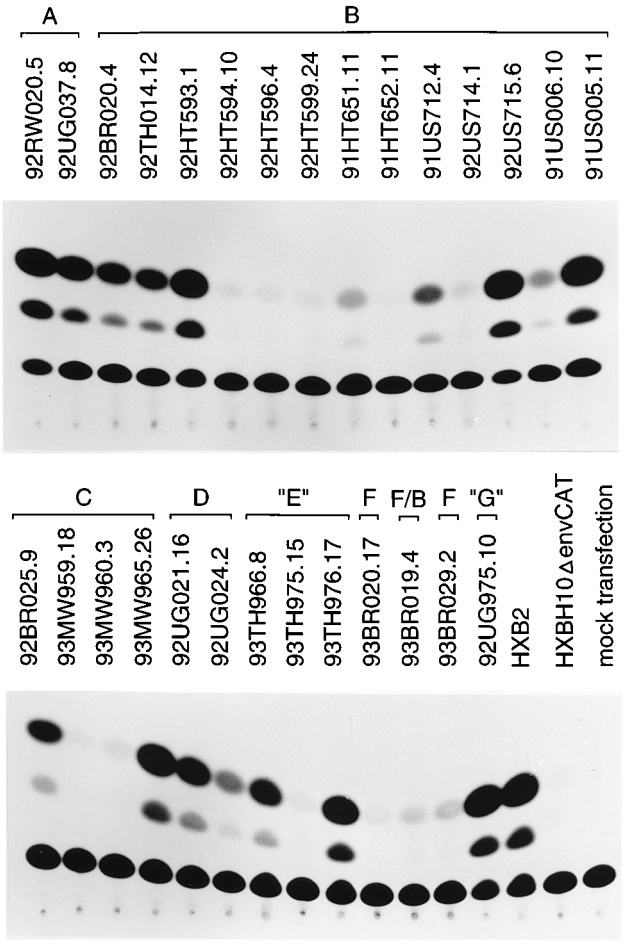


FIG. 7. Biological characterization of PCR-derived envelope genes. Cos-1 cells were cotransfected with pSVIII-gp160 constructs and HXBH10ΔenvCAT as described in Materials and Methods. Virus-containing supernatants were then used to infect phytohemagglutinin-stimulated normal donor PBMCs. Following 2 days in culture, PBMCs were pelleted and lysed by repeated freezing and thawing. Lysates were assayed for CAT activity with [^{14}C]chloramphenicol. Products were analyzed by thin-layer chromatography and quantified on a phosphoimager radioanalytical imaging system. HXB2, positive control; HXBH10ΔenvCAT and mock transfection, negative controls. Brackets at the top indicate sequence subtypes (quotation marks indicate the mosaic nature of subtype E and G envelope genes).

Downloaded from http://jvi.asm.org/ on October 16, 2019 by guest

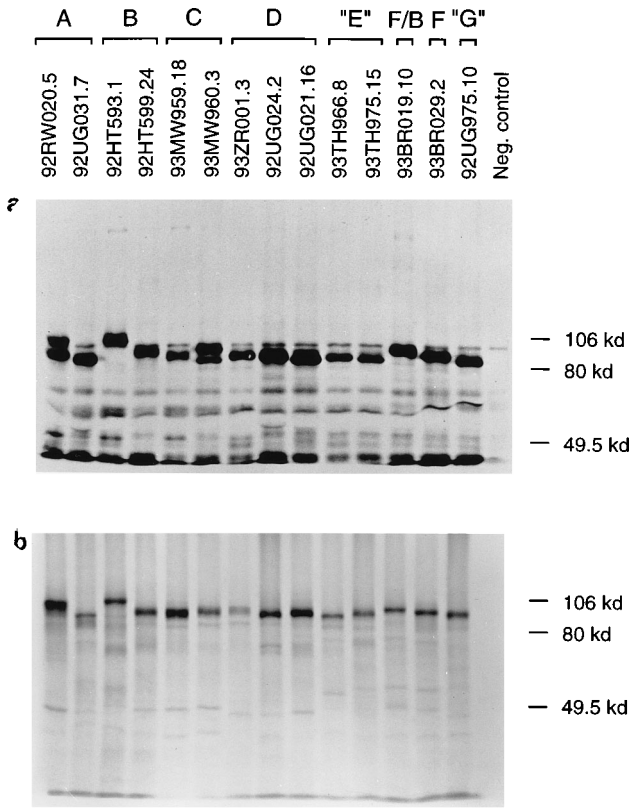


FIG. 8. Analysis of envelope precursor backbone size and glycosylation. pCRII-gp160 constructs were expressed in the T7 transient-expression system in the presence of tunicamycin (a) and in an in vitro transcription-translation reaction (b). Protein marker sizes are indicated in the margin in kilodaltons. Clone designations and subtype classifications are indicated (quotation marks indicate the mosaic nature of subtype E and G envelope genes).

principle. However, given the apparently strong selection pressures for inserts with defective open reading frames, this approach seems impractical, at least for large numbers of HIV-1 isolates.

DISCUSSION

A panel of functional envelope constructs representing HIV-1 sequence subtypes A through G. The HIV sequence database currently contains over 750 envelope sequences from individual HIV-1 strains, almost all of which represent group M viruses (71). The great majority of these comprise small subgenomic regions (generally less than 500 bp in length) which have been generated for subtype classification or studies of V3 loop amino acid variation (39). Most full-length *env* sequences have been generated for subtype B viruses, although there is generally very little information concerning the biological activity of their gene products. For non-subtype B viruses, 35 full-length *env* sequences have been deposited in the database (71). Only five of these are known to express functional glycoproteins, since they are derived from replication-competent proviruses (NDK, Z2Z6, SF170, ELI, and MAL). Two of these (NDK and Z2Z6) were cloned from T-cell-adapted isolates following prolonged in vitro culture (71, 94). A third (MAL) was shown to represent an intersubtype recombinant (49, 82, 83). Thus, there are currently only two full-length and functional *env* genes available from PBMC-propagated viruses which group in clades other than subtype B. These are SF170 from subtype A (71) and ELI from subtype D (2).

Because the products of the HIV-1 envelope gene are prime targets of ongoing vaccine development efforts, we generated a panel of genetically diverse *env* constructs from 35 primary HIV-1 isolates collected at major epicenters of the current AIDS pandemic. Sixteen isolates were derived from geographically diverse subtype B infections and included viruses from North and South America as well as Asia. The other 19 were selected from international vaccine evaluation sites and represented all major subtypes of HIV-1 group M (Table 1). Envelope genes were amplified by PCR from short-term PBMC cultures, transiently expressed to ensure integrity of their open reading frames, and sequenced in their entirety. Phylogenetic analysis was used to confirm their subtype classification and to examine full-length gp160 sequences for evidence of recombination. Most importantly, however, envelope genes were assessed for biological activity in the context of HIV-1 virions by using an in *trans* complementation assay that examines envelope glycoprotein function at all stages of the virus life cycle (11, 25, 56, 95). Using this approach, we identified 20 envelope genes, including at least one from each subtype (or recombinant), to encode fully functional glycoproteins (Table 2). These constructs thus represent a unique set of reagents that should be useful both for basic and applied studies of HIV-1 envelope glycoprotein structure, function, and immunogenicity.

Intersubtype recombination. We also examined whether any of the HIV-1 isolates analyzed in this study represented an intersubtype recombinant. Indeed, detailed phylogenetic analyses of full-length gp160 sequences revealed that several viruses contained mosaic *env* genes. The most clear-cut example was 93BR019, which represents a B/F recombinant (Fig. 5). In addition, subtype G and E viruses exhibited discordant branching orders, clustering in subtype A in the 3′ half of their gp41 coding regions (because the E/A and G/A recombination breakpoints remain tentative, we have not changed the nomenclature of subtype "E" and "G" viruses; however, we indicate their mosaic nature in quotation marks in figures and tables). These results thus confirm our previous findings of the relatively frequent occurrence of intersubtype recombination in geographic areas where multiple subtypes of HIV-1 cocirculate (82, 83). In a recent survey of all near-full-length *gag* or *env* sequences in the database, we found that approximately 10% originated from probable recombinant viruses (83). We have identified a similar fraction of mosaic sequences among the newly characterized HIV-1 isolates in this independent data set. Moreover, in contrast to our previous study, we could now confirm the presence of one of these recombinants in primary uncultured patient material. These results thus indicate that intersubtype recombinants are generated and can be expected to increase in frequency in populations that experience epidemic spread and intermixing of different sequence subtypes.

Two major clades of HIV-1 (subtypes B and E) are currently circulating in Thailand (44, 58, 76, 99). Members of one of these (subtype E) have long been known to branch inconsistently in phylogenetic trees derived from different parts of their genome. These viruses cluster as a distinct group in *env* but fall into subtype A in their *gag* region (58, 90). This observation, along with our recent finding that mosaic sequences are much more frequent than previously thought (83), led us to conclude that "subtype E" viruses most probably represent A/E recombinants. However, it has been suggested that the discordant branching orders may be an artifact of unequal evolutionary rates in *env* and *gag* gene sequences. Two different scenarios would be consistent with this assumption. One possibility would be that subtype E viruses are actually members of sub-

Downloaded from http://jvi.asm.org/ on October 16, 2019 by guest

type A but that they have evolved faster in their envelope coding region and thus branch as a distinct group in *env*-derived trees. However, to explain this possibility, one would have to assume a much increased rate of evolution, and it would not be sufficient that the changes be merely divergent, or else the viruses would simply fall within subtype A but have very long branches (as is seen, for example, for viruses that have been subject to G-to-A hypermutation). Another possibility would be that the *env* trees actually reflect the true phylogenetic position of these viruses but that their *gag* gene sequences have evolved in such a way that they now cluster in subtype A. Because the positions of subtype E viruses in *gag* and *env* trees are so different (71, 90), this would require an extraordinarily large fraction of *gag* substitutions to be convergent on subtype A. In either case, the extent of convergent change would be unprecedented. Given these considerations, it thus seems far more likely that subtype E viruses are indeed intersubtype (A/E) recombinants. Our present finding of subtype A-like sequences in the 3′ half of the subtype E gp41 coding region is consistent with this hypothesis.

The example of the Thai A/E viruses also raises the possibility that certain "subtypes" initially designated on the basis of partial sequences are composed entirely of recombinant viruses. All subtype E viruses so far analyzed appear to have arisen from a single initial A/E recombinant ancestor, presumably generated in central Africa, where viruses classified as A and E coexist. The viral lineage giving rise to subtype E must have existed separately for some time to accumulate the amount of divergence from other viruses now seen in subtype E *env* sequences. However, it is quite possible that the true subtype E lineage has become extinct and that the only evidence of this subtype remaining is in the descendants of the A/E recombinant. Our recent finding that a "subtype E" provirus from the Central African Republic exhibits a pattern of mosaicism indistinguishable from Thai A/E recombinants is consistent with this hypothesis (18).

Finally, it is important to note that both subtype G viruses characterized in this study contain mosaic *env* genes. Although most of their *env* coding region (including all of gp120 and the exterior portion of gp41) falls into subtype G, sequences in the 3′ half of their gp41 domain cluster in subtype A. This discordant branching order was significant for both 92UG975 and 92RU131, as well as for a third virus (LBV21-7) from the database (Fig. 6). Although we cannot define the recombination crossover points with certainty, primarily because of lack of sufficient subtype G sequences in the database, our preliminary data would suggest that they are near the membrane-spanning domain and possibly at the same location in all three viruses.

Biological characterization of PCR-derived envelope clones. Previous reports have relied on sequence analysis (53) and in some instances transient glycoprotein expression (19) as a means to identify potentially functional *env* clones. However, as shown in this study, inspection of the deduced *env* amino acid sequence, even after documentation of protein expression, is not sufficient to reliably predict glycoprotein biological activity. For example, clones 91HT652.11 and 93BR020.17 contained no alterations in the numbers or positions of cysteine residues, no unusual codon insertions or deletions, no unexpected changes in N-linked glycosylation sites, and no premature truncations. Nevertheless, these clones yielded only very weak activity in the CAT complementation assay. By contrast, clones 93MW965.26 and 93TH976.17 yielded CAT conversion rates comparable to that of the positive HXB2 control (Fig. 7), although both contained unpaired cysteine residues in their deduced amino acid sequence expected to give rise to an inappropriately folded glycoprotein (Fig. 3). Finally, in vitro translation and transient expression demonstrated that HIV-1 envelope glycoproteins can vary considerably in size and carbohydrate content (Fig. 8). Consequently, constructs with premature gp41 truncations can go unrecognized when analyzed only for glycoprotein precursor size in transient-expression studies. These results thus indicate that the biological activity of PCR-derived envelope constructs has to be confirmed in the context of infectious HIV-1 virions.

Primary HIV-1 isolates (in contrast to T-cell-adapted strains) are notoriously difficult to neutralize, which is most probably due to a relative inaccessibility of neutralization epitopes on the surface of their oligomeric envelope glycoprotein complex (65, 66, 101). It has thus been proposed that HIV-1 has evolved to evade the host humoral immune response by masking neutralization epitopes, although this may also entail the loss of some replicative capacity as a result of a concomitant reduction in CD4-binding affinity (63, 65, 95). Our finding of envelope constructs with low to borderline activity in the CAT complementation assay which lack amino acid changes predictive of altered glycoprotein function (e.g., 91HT652.11, 93BR020.17, 93MW959.18, and 93MW960.3) is interesting in this context, since it confirms earlier reports of a reduced infectious potential of the *env* gene products from some primary viruses (95). Using the same in *trans* complementation assay, Sullivan et al. found that the *env* gene of YU-2 resulted in only 4% CAT conversion (95), although it was derived from a fully replication competent provirus (50). By contrast, the envelope clone of another primary isolate, 89.6, yielded 77% CAT conversion (95). Given this degree of variability, it is possible that the low activity of some of our constructs represents a property intrinsic to the envelopes of primary viruses (suboptimal transfection efficiencies can be excluded, since all borderline-positive clones were analyzed in multiple independent experiments). Alternatively, the observed reduction in infectivity of some *env* complemented virions may be due to amino acid replacements arising from PCR misincorporation or the amplification of a minor variant. For example, the *env* gene of 92HT596.4 contains a mutation of the cysteine residue that is located between V1 and V2. This envelope construct is very weakly positive in the CAT assay, consistent with the finding that even substantial deletions and alterations in the V1/V2 loop structures are functionally tolerated (102). However, it is unlikely that clone 4 represents the most predominant viral form in the 92HT596 culture, since this isolate replicated with an SI phenotype (compare Table 1). Another example is clone 93TH975.15, which has an unusual amino acid substitution in C3 in which one of two acidic residues (boldface type) known to be important for CD4 binding (75) is altered (GG**D**P/L**E** in the HIV-1 consensus is changed to GG**D**L**K** in 93TH975.15 [Fig. 3]). Future studies will have to determine whether this *env* clone is representative of the predominant virus in the 93TH975 culture.

One other interesting finding was that clone 91US005.11, which was derived from a patient with rapid CD4 cell decline (78, 79), was strongly positive in the CAT complementation assay although its gp41 was prematurely truncated (lacking 17 residues from the carboxy terminus). On the basis of previous reports implicating the intracellular portion of gp41 in glycoprotein transport, steady-state expression levels, packaging, and early entry processes (11, 17, 28, 51, 55, 104), we expected virions complemented with this envelope to exhibit relatively poor infectivity. Instead, 91US005.11 resulted in CAT conversion rates comparable to or higher than those of the positive HXB2 control (Fig. 7). To explain this unusual phenotype, we examined the deduced amino acid sequence of 91US005.11

Downloaded from http://jvi.asm.org/ on October 16, 2019 by guest

more closely. Interestingly, this clone also contained an amino acid substitution at position 721, where a tyrosine was replaced by a cysteine (Y721C). LaBranche et al. recently demonstrated that mutation of the corresponding tyrosine residue in SIV_{MAC} (Y723C) resulted in markedly increased expression of envelope glycoprotein on the cell surface (46). Identical observations were made by Rowell et al., who expressed an HIV-1 envelope glycoprotein mutated in this same tyrosine residue (Y707A) in a recombinant vaccinia virus (84). Most recently, Sauter et al. demonstrated that this tyrosine residue, which is highly conserved among all primate lentiviruses, is part of an endocytosis signal which is responsible for the rapid internalization of envelope molecules from the cell surface via clathrin-coated pits (89). In light of these reports, it seems likely that the cysteine mutation in 91US005.11 would also result in an increase of the amount of glycoprotein on the cell surface. This, in turn, may compensate any potentially negative effect that a truncated gp41 molecule may exert on viral infectivity and replication. Studies are under way to investigate this possibility as well as the question whether the unusual clinical presentation of patient 91US005, who exhibited an extremely high virus load at the time of seroconversion (79) and rapid CD4 T-cell decline over the period of several months (78), is related to these observations.

Utility of the generated envelope clones. The availability of envelope constructs with documented biological activity should now facilitate efforts aimed at identifying biological and immunological correlates of different sequence subtypes of HIV-1. Functional envelope constructs will be useful for the generation of clade-specific antigens for cytotoxic T-lymphocyte analyses, the construction of SIV-HIV chimeric proviruses for in vivo studies in the macaque animal model (98), the generation of recombinant poliovirus replicons (68, 80), and the generation of other subunit vaccines. The reagents described should also facilitate mechanistic studies relating to the structure and function relationships of envelope glycoproteins from primary HIV-1 isolates. Using the same CAT complementation assay, Sullivan et al. recently reported that the fusogenicity of envelope glycoproteins from a few subtype B primary isolates can be enhanced by the binding of soluble CD4 or certain human monoclonal antibodies directed against the V3 loop or the CD4-binding site-related epitopes (95). In an independent study, Moore et al. found that the infectivity of some primary isolates derived from several different subtypes (A to F and I) can be enhanced by HIV-1-positive human sera, in an isolate-dependent (but not serum-dependent) manner (42, 64). Whether these observations are linked is not yet known. However, the envelope constructs described here will now allow the dissection of fusion enhancement in *env* complementation and infectivity assays. This, in turn, may reveal whether this phenomenon represents a significant obstacle to the development of AIDS vaccines that are based on the induction of humoral immunity (63).

All envelope constructs have been submitted to the National Institutes of Health Research and Reference Program, Bethesda, Md., and all sequences have been recorded in GenBank and are available on line through the Los Alamos database. These reagents are thus available to interested investigators and manufacturers involved in the development and testing of HIV vaccines.

ACKNOWLEDGMENTS

We thank Y. Cao and D. D. Ho, Aaron Diamond AIDS Research Center, New York, N.Y., and the NIH AIDS Research and Reference Reagent Program, Rockville, Md., for providing expanded PBMC cultures of HIV-1 isolates; A. C. Siepel, E. E. Allen, and B. T. M. Korber, Los Alamos National Laboratory, Los Alamos, N.Mex., for independent phylogenetic analyses; J. A. Hoxie, University of Pennsylvania, Philadelphia, Pa., for unpublished information concerning the endocytosis mechanisms of HIV/SIV envelope glycoproteins; J. P. Moore, Aaron Diamond AIDS Research Center, for helpful discussions; A. Bobkov and M. Garaev, Ivanovski Institute of Virology, Moskow, Russia, for providing blood samples; Y.-L. Chen, B. Jian, and C. Carruthers for excellent technical assistance; the members of the WHO and NIAID Networks on HIV Isolation and Characterization for their continuing collaborative interactions; and J. B. Wilson and K. T. Agnew for artwork and preparation of the manuscript.

This work was supported by grants from the National Institute of Health (RO1AI25291 and NO1AI35170), by shared facilities of the UAB Center for AIDS Research (DNA Sequencing Core; P30 AI27767), and by the Birmingham Veterans Administration Medical Center. D.L.R. is supported by a studentship from the University of Nottingham.

REFERENCES

1. **Abimiku, A. G., T. L. Stern, A. Zwandor, P. D. Markham, C. Calef, S. Kyari, W. C. Saxinger, R. C. Gallo, M. Robert-Guroff, and M. S. Reitz.** 1994. Subgroup G HIV type 1 isolates from Nigeria. AIDS Res. Hum. Retroviruses **10:**1581–1583.
2. **Alizon, M., S. Wain-Hobson, L. Montagnier, and P. Sonigo.** 1986. Genetic variability of the AIDS virus: nucleotide sequence analysis of two isolates from African patients. Cell **46:**63–74.
3. **Baskar, P. V., S. C. Ray, R. Rao, T. C. Quinn, J. E. K. Hildreth, and R. C. Bollinger.** 1994. Presence in India of HIV type 1 similar to North American strains. AIDS Res. Hum. Retroviruses **8:**1039–1041.
4. **Bobkov, A., R. Cheingsong-Popov, M. Garaev, A. Rzhaninova, P. Kaleebu, S. Beddows, M. H. Bachmann, J. I. Mullins, J. Louwagie, W. Janssens, G. van der Groen, F. McCutchan, and J. Weber.** 1994. Identification of an *env* G subtype and heterogeneity of HIV-1 strains in the Russian Federation and Belarus. AIDS **8:**1649–1655.
5. **Bobkov, A., G. M. M., A. Rzhaninova, P. Kaleebu, R. Pitman, J. N. Weber, and R. Cheingsong-Popov.** 1994. Molecular epidemiology of HIV-1 in the former Soviet Union: analysis of *env* V3 sequences and their correlation with epidemiologic data. AIDS **8:**619–624.
6. **Bolognesi, D. P., L. Corey, S. H. Vermund, and D. F. Hoth.** 1994. HIV vaccine development: a progress report. Ann. Intern. Med. **8:**603–611.
7. **Bruce, C., C. Clegg, A. Featherstone, J. Smith, B. Biryahawaho, R. Downing, and J. Oram.** 1994. Presence of multiple genetic subtypes of human immunodeficiency virus type 1 proviruses in Uganda. AIDS Res. Hum. Retroviruses **10:**1543–1550.
8. **Charneau, P., A. M. Borman, C. Quillent, D. Guétard, S. Charmaret, J. Cohen, G. Rémy, L. Montagnier, and F. Clavel.** 1994. Isolation and envelope sequence of a highly divergent HIV-1 isolate: definition of a new HIV-1 group. Virology **205:**247–253.
9. **Delwart, E. L., E. G. Shpaer, J. Louwagie, F. E. McCutchan, M. Grez, H. Rübsamen-Waigmann, and J. I. Mullins.** 1993. Genetic relationships determined by a DNA heteroduplex mobility assay: analysis of HIV-1 *env* genes. Science **262:**1257–1261.
10. **Dietrich, U., M. Grez, H. von Briesen, B. Panhans, M. Geinbendörfer, H. Kühnel, J. Maniar, G. Mahambre, W. B. Becker, M. B. Becker, and H. Rübsamen-Waigmann.** 1993. HIV-1 strains from India are highly divergent from prototypic African and US/European strains, but are linked to a South African isolate. AIDS Res. Hum. Retroviruses **8:**23–27.
11. **Dubay, J. W., S. J. Roberts, B. H. Hahn, and E. Hunter.** 1992. Truncation of the human immunodeficiency virus type 1 transmembrane glycoprotein cytoplasmic domain blocks virus infectivity. J. Virol. **66:**6616–6625.
12. **Dumitrescu, O., M. L. Kalish, S. C. Kliks, C. I. Bandea, and J. A. Levy.** 1994. Characterization of human immunodeficiency virus type 1 isolates from children in Romania: identification of a new envelope subtype. J. Infect. Dis. **169:**281–288.
13. **Faulkner, D. M., and J. Jurka.** 1988. Multiple aligned sequence editor (MASE). Trends Biochem. Sci. **13:**321–322.
14. **Felsenstein, J.** 1985. Confidence limits on phylogenies: an approach using the boostrap. Evolution **39:**783–791.
15. **Felsenstein, J.** 1992. PHYLIP (Phylogeny Inference Package), 3.5c ed. Department of Genetics, University of Washington, Seattle.
16. **Fuerst, R. T., E. G. Niles, W. Studier, and B. Moss.** 1986. Eukaryotic transient-expression system based on recombinant vaccinia virus that synthesizes bacteriophage T7 RNA polymerase. Proc. Natl. Acad. Sci. USA **83:**8122–8126.
17. **Gabuzda, D. H., A. Lever, E. Terwilliger, and J. Sodroski.** 1992. Effects of deletions in the cytoplasmic domain on biological functions of human immunodeficiency virus type 1 envelope glycoproteins. J. Virol. **66:**3306–3315.
18. **Gao, F., and B. H. Hahn.** Unpublished observations.
19. **Gao, F., L. Yue, S. Craig, C. L. Thornton, D. L. Robertson, F. E. McCutchan, J. A. Bradac, P. M. Sharp, and B. H. Hahn.** 1994. Genetic

Downloaded from http://jvi.asm.org/ on October 16, 2019 by guest

variation of HIV type 1 in four World Health Organization-sponsored vaccine evaluation sites: generation of functional envelope (glycoprotein 160) clones representative of sequence subtypes A, B, C, and E. AIDS Res. Hum. Retroviruses **10:**1359–1368.
20. **Gao, F., L. Yue, S. C. Hill, D. L. Robertson, A. H. Graves, M. S. Saag, G. M. Shaw, P. M. Sharp, and B. H. Hahn.** 1994. HIV-1 sequence subtype D in the United States. AIDS Res. Hum. Retroviruses **10:**625–627.
21. **Graham, B. S., and P. F. Wright.** 1995. Candidate AIDS vaccines. N. Engl. J. Med. **333:**1331–1339.
22. **Grez, M., U. Dietrich, P. Balfe, H. von Briesen, J. K. Maniar, G. Mahambre, E. L. Delwart, J. I. Mullins, and H. Rübsamen-Waigmann.** 1994. Genetic analysis of human immunodeficiency virus type 1 and 2 (HIV-1 and HIV-2) mixed infections in India reveals a recent spread of HIV-1 and HIV-2 from a single ancestor for each of the viruses. J. Virol. **68:**2161–2168.
23. **Gürtler, L. G., P. H. Hauser, J. Eberle, A. von Brunn, S. Knapp, L. Zenkeng, J. M. Tsague, and L. Kaptue.** 1994. A new subtype of human immunodeficiency virus type 1 (MVP-5180) from Cameroon. J. Virol. **68:** 1581–1585.
24. **Haynes, B. F.** 1993. Scientific and social issues of human immunodeficiency virus vaccine development. Science **260:**1279–1286.
25. **Helseth, E., M. Kowalski, D. Gabuzda, U. Olshevsky, W. Haseltine, and J. Sodroski.** 1990. Rapid complementation assays measuring replicative potential of human immunodeficiency virus type 1 envelope glycoprotein mutants. J. Virol. **64:**2416–2420.
26. **Higgins, D. G., and P. M. Sharp.** 1989. Fast and sensitive multiple sequence alignments on a microcomputer. Comput. Appl. Biosci. **5:**151–153.
27. **Ho, D. D., T. Moudgil, and M. Alam.** 1989. Quantitation of human immunodeficiency virus type 1 in the blood of infected persons. N. Engl. J. Med. **321:**1621–1625.
28. **Hunter, E.** 1994. Macromolecular interactions in the assembly of HIV and other retroviruses. Semin. Virol. **5:**71–83.
29. **Jain, M. K., T. J. John, and G. T. Keusch.** 1994. A review of human immunodeficiency virus infection in India. J. Acquired Immune Defic. Syndr. **7:**1185–1194.
30. **Jameel, S., M. Zafrullah, M. Ahmad, G. S. Kapoor, and S. Sehgal.** 1995. A genetic analysis of HIV-1 from Punjab, India reveals the presence of multiple variants. AIDS **9:**685–690.
31. **Janssens, W., L. Heyndrickx, Y. V. de Peer, A. Bouckaert, K. Fransen, J. Motte, G.-M. Gershy-Damet, M. Peeters, P. Piot, and G. van der Groen.** 1993. Molecular phylogeny of part of the env gene of HIV-1 strains isolated in Côte d'Ivoire. AIDS **8:**21–26.
32. **Janssens, W., L. Heyndrickx, K. Fransen, J. Motte, M. Peeters, J. N. Nkengasong, P. M. Ndumbe, E. Delaporte, J.-L. Perret, C. Atende, P. Piot, and G. Van Der Groen.** 1994. Genetic and phylogenetic analysis of env subtypes G and H in central Africa. AIDS Res. Hum. Retroviruses **10:**877–879.
33. **Janssens, W., L. Heyndrickx, K. Fransen, M. Temmerman, A. Leonaers, T. Ivens, J. Motte, P. Piot, and G. van der Groen.** 1994. Genetic variability of HIV type 1 in Kenya. AIDS Res. Hum. Retroviruses **10:**1577–1579.
34. **Kaleebu, P., A. Bobkov, R. Cheingsong-Popov, P. Bieniasz, M. Garaev, and J. Weber.** 1995. Identification of HIV-1 subtype G from Uganda. AIDS Res. Hum. Retroviruses **11:**657–659.
35. **Kalish, M. L., C.-C. Luo, B. G. Weniger, K. Limpakarnjanarat, N. Young, C.-Y. Ou, and G. Schochetman.** 1994. Early HIV type 1 strains in Thailand were not responsible for the current epidemic. AIDS Res. Hum. Retroviruses **10:**1573–1575.
36. **Kimura, M.** 1980. A simple method for estimating evolutionary rates of base substitutions through comparitive studies of nucleotide sequences. J. Mol. Evol. **16:**111–120.
37. **Korber, B. T. M., E. E. Allen, A. D. Farmer, and G. L. Myers.** 1995. Heterogeneity of HIV-1 and HIV-2. AIDS **9:**S5–S18.
38. **Korber, B. T. M., G. Learn, G. I. Mullins, B. H. Hahn, and S. Wolinsky.** 1995. Protecting HIV databases. Nature (London) **378:**242–243.
39. **Korber, B. T. M., K. MacInnes, R. F. Smith, and G. Myers.** 1994. Mutational trends in V3 loop protein sequences observed in different genetic lineages of human immunodeficiency virus type 1. J. Virol. **68:**6730–6744.
40. **Korber, B. T. M., S. Osmanov, J. Esparza, G. Myers, and WHO Network for HIV Isolation and Characterization.** 1994. The World Health Organization Global Programme on AIDS proposal for standardization of HIV sequence nomenclature. AIDS Res. Hum. Retroviruses **10:**1355–1358.
41. **Kostrikis, L. G., E. Bagdades, Y. Cao, L. Zhang, D. Dimitriou, and D. D. Ho.** 1995. Genetic analysis of human immunodeficiency virus type 1 strains from patients in Cyprus: identification of a new subtype designated subtype I. J. Virol. **69:**6122–6130.
42. **Kostrikis, L. G., Y. Cao, H. Ngai, J. P. Moore, and D. D. Ho.** 1996. Quantitative analysis of serum neutralization of human immunodeficiency virus type 1 from subtypes A, B, C, D, E, F, and I: lack of direct correlation between neutralization serotypes and genetic subtypes and evidence for prevalent serum-dependent infectivity enhancement. J. Virol. **70:**445–458.
43. **Kuiken, C. L., G. Zwart, E. Baan, R. A. Coutinho, J. A. R. van der Hoek, and J. Goudsmit.** 1993. Increasing antigenic and genetic diversity of the V3 variable domain of the human immunodeficiency virus envelope protein in the course of the AIDS epidemic. Proc. Natl. Acad. Sci. USA **90:**9061–9065.
44. **Kunanusont, C., H. M. Foy, J. K. Kreiss, S. Rerks-Ngarm, P. Phanuphak, S. Raktham, C.-P. Pau, and N. L. Young.** 1995. HIV-1 subtypes and male-to-female transmission in Thailand. Lancet **345:**1078–1083.
45. **Kusumi, K., B. Conway, S. Cunningham, A. Berson, C. Evans, A. K. N. Iversen, D. Colvin, M. V. Gallo, S. Coutre, E. G. Shpaer, D. V. Faulkner, A. deRonde, S. Volkman, C. Williams, M. S. Hirsch, and J. I. Mullins.** 1992. Human immunodeficiency virus type 1 envelope gene structure and diversity in vivo and after cocultivation in vitro. J. Virol. **66:**875–885.
46. **LaBranche, C. C., M. M. Sauter, B. S. Haggarty, P. J. Vance, J. Romano, T. K. Hart, P. J. Bugelski, M. Marsh, and J. A. Hoxie.** 1995. A single amino acid change in the cytoplasmic domain of the simian immunodeficiency virus transmembrane molecule increases envelope glycoprotein expression on infected cells. J. Virol. **69:**5217–5227.
47. **Leitner, T., A. Alaeus, S. Marquina, E. Lilja, K. Lidman, and J. Albert.** 1995. Yet another subtype of HIV type 1? AIDS Res. Hum. Retroviruses **11:**995–997.
48. **Leonard, C. K., M. W. Spellman, L. Riddle, R. J. Harrison, J. N. Thomas, and T. J. Gregory.** 1990. Assignment of intrachain disulfide bonds and characterization of potential glycosylation sites of the type 1 recombinant human immunodeficiency virus envelope glycoprotein (gp120) expressed in Chinese hamster ovary cells. J. Biol. Chem. **265:**10373–10382.
49. **Li, W.-H., M. Tanimura, and P. M. Sharp.** 1988. Rates and dates of divergence between AIDS virus nucleotide sequences. Mol. Biol. Evol. **5:**313–330.
50. **Li, Y., J. C. Kappes, J. A. Conway, R. W. Price, G. M. Shaw, and B. H. Hahn.** 1991. Molecular characterization of human immunodeficiency virus type 1 cloned directly from uncultured human brain tissue: identification of replication-competent and -defective viral genomes. J. Virol. **65:**3973–3985.
51. **Lodge, R., H. Göttlinger, D. Gabuzda, E. A. Cohen, and G. Lemay.** 1994. The intracytoplasmic domain of gp41 mediates polarized budding of human immunodeficiency virus type 1 in MDCK cells. J. Virol. **68:**4857–4861.
52. **Loussert-Ajaka, I., M.-L. Chaix, B. T. M. Korber, F. Letourneur, E. Gomas, E. Allen, T.-D. Ly, F. Brun-Vezinet, F. Simon, and S. Saragosti.** 1995. Variability of human immunodeficiency virus type 1 group O strains isolated from Cameroonian patients living in France. J. Virol. **69:**5640–5649.
53. **Louwagie, J., W. Janssens, J. Mascola, L. Heyndrickx, P. Hegerich, G. van der Groen, F. E. McCutchan, and D. S. Burke.** 1995. Genetic diversity of the envelope glycoprotein from human immunodeficiency virus type 1 isolates of African origin. J. Virol. **69:**263–271.
54. **Louwagie, J., F. E. McCutchan, M. Peeters, T. P. Brennan, E. Sanders-Buell, G. A. Eddy, G. van der Groen, K. Fransen, G.-M. Gershy-Damet, R. Deleys, and D. S. Burke.** 1993. Phylogenetic analysis of *gag* genes from 70 international HIV-1 isolates provides evidence for multiple genotypes. AIDS **7:**769–780.
55. **Lydy, S. L., and R. W. Compans.** 1993. Role of the cytoplasmic domains of viral glycoproteins in antibody-induced cell surface mobility. J. Virol. **67:** 6289–6294.
56. **Mammano, F., E. Kondo, J. Sodroski, A. Bukovsky, and H. G. Göttlinger.** 1995. Rescue of human immunodeficiency virus type 1 matrix protein mutants by envelope glycoproteins with short cytoplasmic domains. J. Virol. **69:**3824–3830.
57. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual, p. 269–295. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
58. **McCutchan, F. E., P. A. Hegerich, T. P. Brennan, P. Phanuphak, P. Singharaj, A. Jugsudee, P. W. Berman, A. M. Gray, A. K. Fowler, and D. S. Burke.** 1992. Genetic variants of HIV-1 in Thailand. AIDS Res. Hum. Retroviruses **8:**1887–1895.
59. **McCutchan, F. E., B. L. P. Ungar, P. Hegerich, C. R. Roberts, A. K. Fowler, S. K. Hira, L. Perine, and D. S. Burke.** 1992. Genetic analysis of HIV-1 isolates from Zambia and an expanded phylogenenic tree for HIV-1. J. Acquired Immune Defic. Syndr. **5:**441–449.
60. **Miotti, P., J. Canner, J. Chiphangwi, G. Liomba, A. Saah, and G. Dallabetta.** 1994. Preparations for AIDS vaccine evaluations: rate of new HIV infection in a cohort of women of childbearing age in Malawi. AIDS Res. Hum. Retroviruses **10:**S239–S241.
61. **Modrow, S., B. H. Hahn, G. M. Shaw, R. C. Gallo, F. Wong-Staal, and H. Wolf.** 1987. Computer-assisted analysis of envelope protein sequences of seven human immunodeficiency virus isolates: prediction of antigenic epitopes in conserved and variable domains. J. Virol. **61:**570–578.
62. **Moore, J., and R. Anderson.** 1994. The WHO and why of HIV vaccine trials. Nature (London) **372:**313–314.
63. **Moore, J. P.** 1995. Back to primary school. Nature (London) **376:**115.
64. **Moore, J. P., Y. Cao, J. Leu, L. Qin, B. T. M. Korber, and D. D. Ho.** Inter- and intraclade neutralization of human immunodeficiency virus type 1: genetic clades do not correspond to neutralization serotypes but partially correspond to gp120 antigenic serotypes. J. Virol. **70:**427–444.
65. **Moore, J. P., Y. Cao, L. Qing, Q. J. Sattentau, J. Pyati, R. Koduri, J. Robinson, C. F. Barbas III, D. R. Burton, and D. D. Ho.** 1995. Primary isolates of human immunodeficiency virus type 1 are relatively resistant to neutralization by monoclonal antibodies to gp120, and their neutralization is not predicted by studies with monomeric gp120. J. Virol. **69:**101–109.

Downloaded from http://jvi.asm.org/ on October 16, 2019 by guest

66. **Moore, J. P., and D. D. Ho.** 1995. HIV-1 neutralization: the consequences of viral adaptation to growth on transformed T cells. AIDS **9:**S117–S136.
67. **Morgado, M. G., E. C. Sabino, E. G. Shpaer, V. Bongertz, L. Brigido, M. D. C. Cuimaraes, E. A. Castilho, B. Galvao-Castro, J. I. Mullins, R. M. Hendry, and A. Mayer.** 1994. V3 region polymorphisms in HIV-1 from Brazil: prevalence of subtype B strains divergent from North American/European prototype and detection of subtype F. AIDS Res. Hum. Retroviruses **10:**569–576.
68. **Morrow, C. D., D. C. Porter, D. C. Ansardi, Z. Moldoveanu, and P. N. Fultz.** 1994. New approaches for mucosal vaccines for AIDS: encapsidation and serial passage of poliovirus replicons that express HIV-1 proteins on infection. AIDS Res. Hum. Retroviruses **10:**S61–S66.
69. **Murphy, E., B. Korber, M.-C. Georges-Courbot, B. You, A. Pinter, D. Cook, M.-P. Kieny, A. Georges, C. Mathiot, F. Barre-Sinoussi, and M. Girard.** 1993. Diversity of V3 region sequences of human immunodeficiency viruses type 1 from the Central African Republic. AIDS Res. Hum. Retroviruses **9:**997–1007.
70. **Myers, G.** 1994. HIV: between past and future. AIDS Res. Hum. Retroviruses **10:**1317–1324.
71. **Myers, G., B. Korber, S. Wain-Hobson, K.-T. Jeang, L. E. Henderson, and G. N. Pavlakis.** 1994. Human retroviruses and AIDS: a compilation and analysis of nucleic acid and amino acid sequences. Los Alamos National Laboratory, Los Alamos, N.M.
72. **Nelson, K. D., D. Vlahov, N. Galai, J. Astemborski, and L. Solomon.** 1994. Preparations for AIDS vaccine trials: incident human immunodeficiency virus (HIV) infections in a cohort of injection drug users in Baltimore, Maryland. AIDS Res. Hum. Retroviruses **10:**S201–S205.
73. **Nelson, K. E., C. Beyrer, C. Natpratan, S. Eiumtrakul, D. D. Celentano, and C. Khamboonruang.** 1994. Preparatory studies for possible HIV vaccine trials in Northern Thailand. AIDS Res. Hum. Retroviruses **10:**S243–S246.
74. **Nkengasong, J. N., W. Janssens, L. Heyndrickx, K. Fransen, P. M. Ndumbe, J. Motte, A. Leonaers, M. Ngolle, J. Ayuk, P. Piot, and G. van der Groen.** 1994. Genetic subtypes of HIV-1 in Cameroon. AIDS **8:**1405–1412.
75. **Olshevsky, U., E. Helseth, C. Furman, J. Li, W. Haseltine, and J. Sodroski.** 1990. Identification of individual human immunodeficiency virus type 1 gp120 amino acids important for CD4 receptor binding. J. Virol. **64:**5701–5707.
76. **Ou, C.-Y., Y. Takebe, B. G. Weniger, C.-C. Luo, M. L. Kalish, W. Auwanit, S. Yamazaki, H. D. Gayle, N. L. Young, and G. Schochetman.** 1993. Independent introductions of two major HIV-1 genotypes into distinct high-risk populations in Thailand. Lancet **341:**1171–1174.
77. **Pantaleo, G., J. F. Demarest, H. Soudeyns, C. Graziosi, F. Denis, J. W. Adelsberger, P. Borrow, M. S. Saag, G. M. Shaw, R. P. Sekaly, and A. S. Fauci.** 1994. Major expansion of $CD8^+$ T cells with a predominant Vβ usage during the primary immune response to HIV. Nature (London) **370:**463–467.
78. **Piatak, M., Jr., M. S. Saag, L. C. Yang, S. J. Clark, J. C. Kappes, K.-C. Luk, B. H. Hahn, G. M. Shaw, and J. D. Lifson.** 1993. Determination of plasma viral load in HIV-1 infection by quantitative competitive polymerase chain reaction. AIDS **7:**S65–S71.
79. **Piatak, M., Jr., M. S. Saag, L. C. Yang, S. J. Clark, J. C. Kappes, K.-C. Luk, B. H. Hahn, G. M. Shaw, and J. D. Lifson.** 1993. High levels of HIV-1 in plasma during all stages of infection determined by competitive PCR. Science **259:**1749–1754.
80. **Porter, D. C., D. C. Ansardi, W. C. Choi, and C. D. Morrow.** 1993. Encapsidation of genetically engineerd poliovirus minireplicons which express human immunodeficiency virus type 1 Gag and Pol proteins upon infection. J. Virol. **67:**3712–3719.
81. **Potts, K. E., M. L. Kalish, C. I. Bandea, G. M. Orloff, M. S. Louis, C. Brown, N. Malanda, M. Kavuka, G. Schochetman, C.-Y. Ou, and W. L. Heyward.** 1993. Genetic diversity of human immunodeficiency virus type 1 strains in Kinshasa, Zaire. AIDS Res. Hum. Retroviruses **9:**613–618.
82. **Robertson, D. L., B. H. Hahn, and P. M. Sharp.** 1995. Recombination in AIDS viruses. J. Mol. Evol. **40:**249–259.
83. **Robertson, D. L., P. M. Sharp, F. E. McCutchan, and B. H. Hahn.** 1995. Recombination in HIV-1. Nature (London) **374:**124–126.
84. **Rowell, J. F., P. E. Stanhope, and R. F. Siliciano.** 1995. Endocytosis of endogenously synthesized HIV-1 envelope protein. J. Immunol. **155:**473–488.
85. **Rübsamen-Waigmann, H., H. von Briesen, H. Holmes, A. Björndal, B. T. M. Korber, R. Esser, S. Ranjbar, P. Tomlinson, B. Galvao-Castro, E. Karita, S. Sempala, C. Wasi, S. Osmanov, E.-M. Fenyö, and WHO Network for HIV Isolation and Characterization.** 1994. Standard conditions of virus isolation reveal biological variability of HIV type 1 in different regions of the world. AIDS Res. Hum. Retroviruses **10:**1401–1408.
86. **Ruff, A., J. Coberly, N. A. Halsey, R. Boulos, J. Desormeaux, A. Burnley, D. J. Joseph, M. McBrien, T. Quinn, P. Losikoff, K. L. O'Brien, M. A. Louis, and H. Farzadegan.** 1994. Prevalence of HIV-1 DNA and p24 antigen in breast milk and correlation with maternal factors. J. Acquired Immune Defic. Syndr. **7:**68–73.
87. **Sabino, E. C., E. G. Shpaer, M. G. Morgado, B. T. M. Korber, R. S. Diaz, V. Bongertz, S. Cavalcante, B. Gâlvao-Castro, J. I. Mullins, and A. Mayer.** 1994. Identification of human immunodeficiency virus type 1 envelope genes recombinant between subtypes B and F in two epidemiologically linked individuals from Brazil. J. Virol. **68:**6340–6346.
88. **Saitou, N., and M. Nei.** 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4:**406–425.
89. **Sauter, M. M., A. Pelchen-Matthews, R. Bron, M. Marsh, C. C. LaBranche, P. J. Vance, J. Romano, B. S. Haggarty, T. K. Hart, W. M. F. Lee, and J. A. Hoxie.** An internalization signal in the simian immunodeficiency virus transmembrane protein cytoplasmic domain modulates expression of envelope glycoproteins on the cell surface. J. Cell. Biol., in press.
90. **Sharp, P. M., D. L. Robertson, F. Gao, and B. H. Hahn.** 1994. Origins and diversity of human immunodeficiency viruses. AIDS **8:**S27–S42.
91. **Sharp, P. M., D. L. Robertson, and B. H. Hahn.** 1995. Cross-species transmission and recombination of 'AIDS' viruses. Philos. Trans. R. Soc. London Ser. B **349:**41–47.
92. **Smith, R. F., and T. F. Smith.** 1990. Automatic generation of primary sequence patterns from sets of related protein sequences. Proc. Natl. Acad. Sci. USA **87:**118–122.
93. **Smith, R. F., and T. F. Smith.** 1992. Pattern-induced multi-sequence alignment (PIMA) algorithm employing secondary structure-dependent gap penalties for use in comparative protein modeling. Protein Eng. **5:**35–41.
94. **Spire, B., J. Sire, V. Zachar, F. Rey, F. Barré-Sinoussi, F. Galibert, A. Hampe, and J.-C. Chermann.** 1989. Nucleotide sequence of HIV1-NDK: a highly cytopathic strain of the human immunodeficiency virus. Gene **81:**275–284.
95. **Sullivan, N., Y. Sun, J. Li, W. Hofmann, and J. Sodroski.** 1995. Replicative function and neutralization sensitivity of envelope glycoproteins from primary and T-cell line-passaged human immunodeficiency virus type 1 isolates. J. Virol. **69:**4413–4422.
96. **Thompson, J. D., D. G. Higgins, and T. J. Gibson.** 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.
97. **Vanden Haesevelde, M., J. L. Decourt, R. J. De Leys, B. Vanderbotght, G. van der Groen, H. van Heuverswijn, and E. Saman.** 1994. Genomic cloning and complete sequence analysis of a highly divergent human immunodeficiency virus isolate. J. Virol. **68:**1586–1596.
98. **Warren, J. T., and M. Dolatshahi.** 1994. Annual updated survey of worldwide HIV, SIV, and SHIV challenge studies in vaccinated nonhuman primates. J. Med. Primatol. **23:**184–225.
99. **Wasi, C., B. Herring, S. Raktham, S. Vanichseni, T. D. Mastro, N. L. Young, H. Rübsamen-Waigmann, H. von Briesen, M. L. Kalish, C.-C. Luo, C.-P. Pau, A. Baldwin, J. I. Mullins, E. L. Delwart, J. Esparza, W. L. Heyward, and S. Osmanov.** 1995. Determination of HIV-1 subtypes in injecting drug users in Bangkok, Thailand, using peptide binding enzyme immunoassay and the heteroduplex mobility assay: evidence of increasing infection with HIV-1 subtype E. AIDS **9:**943–949.
100. **WHO Network for HIV Isolation and Characterization.** 1994. HIV type 1 variation in World Health Organization-sponsored vaccine evaluation sites: genetic screening, sequence analysis, and preliminary biological characterization of selected viral strains. AIDS Res. Hum. Retroviruses **10:**1327–1343.
101. **Wrin, T., and J. H. Nunberg.** 1994. HIV-1MN recombinant gp120 vaccine serum, which fails to neutralize primary isolates of HIV-1, does not antagonize neutralization by antibodies from infected individuals. AIDS **8:**1622–1623.
102. **Wyatt, R., J. Moore, M. Accola, E. Desjardin, J. Robinson, and J. Sodroski.** 1995. Involvement of the V1/V2 variable loop structure in the exposure of human immunodeficiency virus type 1 gp120 epitopes induced by receptor binding. J. Virol. **69:**5723–5733.
103. **Yahi, N., J. Fantini, I. Hirsch, and J.-C. Chermann.** 1992. Structural variability of *env* and *gag* gene products from a highly cytopathic strain of HIV-1. Arch. Virol. **125:**287–298.
104. **Yu, X., X. Yuan, M. F. McLane, T.-H. Lee, and M. Essex.** 1993. Mutations in the cytoplasmic domain of human immunodeficiency virus type 1 transmembrane protein impair the incorporation of Env proteins into mature virions. J. Virol. **67:**213–221.

Downloaded from http://jvi.asm.org/ on October 16, 2019 by guest