**Table 2 Notes on selected full-length HIV-1/SIVcpz nucleotide sequences that appear in the alignments.**

| Sequence | Accession | Origin | Author | Reference |
|---|---|---|---|---|
| A_KE.Q23 | AF004885 | KENYA | Neilson, JR | *J Virol* **73**(5):4393–4403 (1999) |

This subtype A sequence was derived from a woman from Mombasa, Kenya, who had been recently infected with HIV-1. The blood sample was drawn in June 13, 1994. An env gene fragment from a PCR amplification from an earlier blood sample (July 1993) was published in Poss, M., *et al. ARHR* **13**(6):493–499 (1997). The full length sequence was kindly released prior to publication by M. Poss and colleagues, U. Washington. Many env sequences from this same patient are available with accession numbers AF004893 and AF047979–AF048346.

A_SE.SE8891 AF069673 SWEDEN Carr, JK *AIDS* **13**(14):1819–1826 (1999)

This sequence was generated in sweden from a PBMC co-culture of a sample taken from a 26-year-old Ugandan woman. Probable infection country is Uganda, risk factor heterosexual. The virus is CCR5+, CXCR4-. Small sections of the 5′ and 3′ LTRs are not included in this sequence.

A_SE.SE8538 AF069669 SWEDEN Carr, JK *AIDS* **13**(14):1819–1826 (1999)

This sequence was generated in Sweden from a cultured blood sample from a 24-year old Tanzanian woman. Probable country of infection is Tanzania. The virus is NSI, CCR5+ and CXCR4-. Small sections of the 5′ and 3′ LTRs are not included in this sequence.

A_SE.SE6594 AF069672 SWEDEN Carr, JK *AIDS* **13**(14):1819–1826 (1999)

This sequence was from a 29 year old male sampled in Sweden in 1993, who is believed to have been infected in Uganda via heterosexual contact, before moving to Sweden. The patient was CDC-B3 at the time of sampling. This isolate has a syncytium-inducing phenotype and uses the CXCR4 coreceptor for entry. Small sections of the 5′ and 3′ LTRs are not included in this sequence.

A_SE.SE7535 AF069671 SWEDEN Carr, JK *AIDS* **13**(14):1819–1826 (1999)

This sequence is from a 45 year old male sampled in Sweden in 1994, who is believed to have been infected in Uganda via heterosexual contact, before moving to Sweden. The patient was CDC-A1 at the time of sampling. This isolate has a syncytium-inducing phenotype and uses the CXCR4 coreceptor for entry. Small sections of the 5′ and 3′ LTRs are not included in this sequence.

A_SE.SE7253 AF069670 SWEDEN Carr, JK *AIDS* **13**(14):1819–1826 (1999)

This subtype A sequence is from a 27 year old male living in Sweden, who is thought to have been infected in Somalia via heterosexual contact. The patient was CDC stage C3 when sampled in 1994. The virus is NSI and uses the CCR5 coreceptor. The patient's CD4 count was zero. Virus was cocultured with donor PBMC before PCR amplification and direct sequencing. Small sections of the 5′ and 3′ LTRs are not included in this sequence.

**Table 2 (cont.)**

| Sequence | Accession | Origin | Author | Reference |
|---|---|---|---|---|
| A_SE.SE8131 | AF107771 | SWEDEN | Carr, JK | *AIDS* **13**(14):1819–1826 (1999) |

This subtype A sequence is from a 32 year old female living in Sweden, who is thought to have been infected in Uganda via heterosexual contact. The patient was CDC stage C3 when sampled in 1995. The virus is SI and can use both the CCR5 and CXCR4 coreceptor. Virus was cocultured with donor PBMC before PCR amplification and direct sequencing. Small sections of the 5′ and 3′ LTRs are not included in this sequence.

A_UG.U455 M62320 UGANDA Oram, JD *ARHR* **6**(9):1073–1078 (1990)

This sequence is from the 1985 Ugandan isolate U455. It was cloned in phage, and has defective env, vpr, and vpu. The env ORF in this sequence is interrupted by an in-frame stop codon beyond the COOH end of the V5 region. This sequence clusters with subtype A HIV-1.

A_UG.92UG037 U51190 UGANDA Gao, F *J Virol* **70**(3):1651–1657 (1996)

This sequence is from a complete genome PCR amplified from proviral DNA. The patient was a 31 year old asymptomatic female from Entebe, Uganda. See also U09124, U09127. 92UGO37 is one of a set of three complete genomes from a study linking the HIV-1 epidemic in the heterosexual population in Thailand to an A/E recombinant. It is obtained through the WHO Global Programme on AIDS (WHO Network, *ARHR* **10**(11) :1327–1343 (1994)) and comes from an asymptomatic 31-year old female from Entebbe, Uganda; she had not taken any anti-retroviral therapy prior to sampling. The risk factor for infection was heterosexual contact. The isolate 92UGO37 was established and propogated by short term cocultivation with normal donor lymphocytes and then the near full length genome was PCR amplified and sequenced. 92UGO37 is subtype A. An LTR sequence is available under accession number U51287 and an additional env/nef sequence with accession number U09127. There is an inframe stop codon in pol at position 3144 in this clone. The isolate from which this sequence was derived is NSI and uses CCR5 or CCR8 (Bjorndal, A., *et al., J Virol* **71**(10):7478–87 (1997) and Rucker, J., *et. al., J Virol* **71**(12):8999–9007 (1997). See also Gao, F., *et al., J Virol* **70**(10):7013-7029 (1996). This sequence was kindly made available prior to publication, and is now published (Gao F., *et al., J Virol* **72**(7):5680–5698 (1997)). Biotypes were determined by MT-2 syncytium assay; however, both syncytium-inducing (SI) and non-syncytium-inducing (NSI) variants may be present in the viral "swarm" for each isolate. Recent studies indicate that NSI isolates contain predominantly CCR5-using variants while most SI isolates contain both CXCR4 (SI) and CCR5 (NSI) variants. Some SI isolates may contain dual-tropic variants that use both CXCR4 and CCR5 co-receptors. The isolate 93UG037 is available from the NIH AIDS Reagent program, and is NSI R5.

B_AU.MBC200 AF042100 AUSTRALIA Oelrichs, RB *ARHR* **14**(9):811–4 (1998)

Isolate MBC200. Date of sample and source of isolation:18/3/1986, Melbourne, Australia. Biological source:Peripheral blood co-cultured with donor PBMC by the Victorian Infectious Diseases Reference Laboratory. The patient was a Caucasian homosexual male, diagnosed with AIDS in December 1985 at which time the T4:T8 ratio was 0.2. Virus isolation:Biological cloning by three rounds of limiting dilution in donor PBMC. Sequencing:Derived directly from the biologically cloned isolate. Hirt supernatant DNA was obtained from low-passage number donor PBMC culture and sequence derived from overlapping PCR products. All open reading frames are intact and the nucleic acid sequence clusters within subtype B in p17, pol, env and nef. Viral phenotype:Produces high levels of syncytia in PBMC and MT-2 cells. Grows well in Jurkat cells and primary macrophages (Kiernan,R. *ARHR* **6**(6):743–752)

**Table 2 (cont.)**

| Sequence | Accession | Origin | Author | Reference |
|---|---|---|---|---|
| B_AU.MBC925 | AF042101 | AUSTRALIA | Oelrichs, RB | *ARHR* **14**(9):811–4 (1998) |

Isolate MBC925. Date of isolation:15/5/1987, Melbourne, Australia. Biological source:Cerebrospinal fluid cultured with donor PBMC by the Victorian Infectious Diseases Reference Laboratory. The patient was a Caucasian homosexual male, diagnosed with AIDS in October 1986. Neurological symptoms were present. In December 1986 the T4:T8 ratio was 0.14. The patient died in May 1987. Virus isolation:Biological cloning by three rounds of limiting dilution in donor PBMC. Sequencing:Derived directly from the biologically cloned isolate. Hirt supernatant DNA was obtained from low-passage number donor PBMC culture and sequence derived from overlapping PCR products. All open reading frames are intact and the nucleic acid sequence clusters within subtype B in p17, pol, env and nef. Viral phenotype:Non-syncytium inducing in PBMC. The virus does not replicate in MT-2 cells or Jurkat cells. Good growth in primary macrophages.

| | | | | |
|---|---|---|---|---|
| B_CN.RL42 | U71182 | CHINA | Graf, M | *ARHR* **14**(3):285–288 (1998) |

RL42 was isolated from an asymptomatic IVDU, infected by needle sharing, in Dehong prefecture of Yunnan province South of China. This is near the Laos and Thailand golden drug triangle. The isolate was generated by Prof. Dr. Shao Yiming from the Chinese Academy of Preventive Medicine, Beijing, China. This sequence is of the Thai B' subtype (a subset of subtype B), which is the most prevalent subtype of HIV-1 found in the Yunnan province of Southwest China.

| | | | | |
|---|---|---|---|---|
| B_DE.HAN | U43141 | GERMANY | Sauermann, U | *ARHR* **6**(6):813–823 (1990) |

Isolate HAN2 was isolated from a 39 year old homosexual German patient with AIDS related complex, in 1986. This patient died from complications of AIDS in 1987. HAN2 was highly cytopathic in MT-2 T cell line, it was able to productively infect MT-4, H9 or Jurkat cell lines. Genomic DNA from infected MT-2 cells was used to prepare a lambda phage genomic library. Two full-length clones, HAN2/2 and HAN2/3 were purified. HAN2/3 was used for DNA sequencing, and has a defective env gene.

| | | | | |
|---|---|---|---|---|
| B_DE.D31 | U43096 | GERMANY | Kreutz, R | *ARHR* **8**(9):1619–1629 (1992) |

The patient from which this virus is derived, has never been well described. It is only shown as HIV1-D31 in figure 3 of the paper. The complete genome has been sequenced.

| | | | | |
|---|---|---|---|---|
| B_ES.89SP061 | AJ006287 | SPAIN | Olivares, I | *ARHR* **14**(18):1649–1651 (1998) |

This sequence was generated from a melocular clone. This clone in turn was derived from a biological clone out of a co-cultured sample from a 4-year-old boy who was presumably infected via his mother, who was a Spanish intravenous drug user. The sample was taken in 1989; the infection date is unknown. The virus is subtype B, the clone has the SI phenotype but does not display the 'SI-indicating' mutations in the V3 loop.

**Table 2 (cont.)**

| Sequence | Accession | Origin | Author | Reference |
|---|---|---|---|---|
| B_FR.HXB2 | K03455 | FRANCE | Wong-Staal, F | *Nature* **313**(6000):277–284 (1985) |

This sequence was from provirus cloned in lambda phage and is derived from the IIIB isolate related to LAI. This clone has been extensively studied. sites; tat mRNA and other transcript boundaries. Sequence for [25] kindly provided in computer-readable form by L.Ratner, 19-AUG-1986. The HXB2R sequence is being used as a reference genome for all the HIV entries because it has been derived from a demonstrably infectious clone. Hence not all of the 'sites' references above were concerned with this isolate. Since the earliest appearance of this sequence in the HIV database and in GenBank (prior to the 1987 publication of [3]), the 5′LTR appears to have been derived from BH10 rather than HXB2. With this corrected version, the 3′LTR only is shown with annotation of differences from [1] so as to facilitate coordination with [3] and with the latest information on splice sites. Revisions were made by [5], [6], and [7] with approval of the principal author of [3]. These affect restriction site analyses, in particular upstream of the gag cds start and in the envelope cds; the latter affect the RRE sequence at 7266. Many of the revisions bring the HXB2 sequence closer to the BH10 sequence, yet these were the differences receiving the greatest attention in [3]. Possibly mutation has occurred in the sub-cloning and outgrowth, or possibly the more recent HXB2 subclones are BH10 contaminants. Be that as it may, HXB2 clones currently in use have a sequence most closely approximated by the sequence below. For a full comparison of the IIIb/LAV sibling sequences, see Part III. 2 additional changes to the sequence of HIVHXB2R were made per MarvinReitz (personal correspondence), 6/91:the "t" at site 8383 was changed to "c", and the "g" at site 8427 was changed to "a". The vpU cds not annotated below do not possess a start codon in the normal position (bases 5608 to 5610; "ACG"). Schwartz *et al., J. Virol.* **64**: 2519–2529, state that HXB2 does not produce vpU protein (it remains an infectious clone). The minimal continuous RRE (CAR) of 204 nt is defined by [7] to start at coordinate 7327. Dr. Seth Pincus *et al.* [8] report a single base deletion after codon 686 of the envelope gene in an "E variant" set of clones, which results in premature termination of translation and the production of a truncated gp160, causing a marked decrease in the expression of envelope on the surface of the infected cell. This sequence is from the French isolate LAI (formerly BRU) which is also referred to as IIIB. (Wain-Hobson85). Also see:(Alizon86), (Lukashov95b) and (Wain-Hobson91). GenBank accession numbers K02013, L23090–L23103, X01762, L48380–L48399, M64178–M64223, M64406–M64415 and M64768–M64775, AF033819. Other sequences which are of this type include:PV22, K02083; MFA, M33943 (Stevenson90); un-named, Z11530; BH8, K02011; BH10, M15654; TH4, L31963; MCK1, D86068; PM213, D86069; F12CG, Z11530; and HXB, K03455, M38432, M64775 and M14100. The variation of the IIIB isolate in culture was studied by (Lockey96), GenBank accession numbers U54647, U54649, U54651, U54653, U54655, U54657, U54659, U54665, U54667, U54681, U54683, U54685 and U54689. The variation of IIIB/LAI in 9 years of infection in a chimpanzee has been studied by Fultz et al. unpublished, GenBank accession numbers U56866–U56883 and U56888–U56899. The IIIB/LAI isolate of HIV-1 has also been extensively studied in cases such as the infected lab worker. See for example (Reitz94), (Pincus94) U12030–U12055. The tropism of isolates from the lab worker for primary PBMCs and failure to grow in T-cell lines was localized to the V3-loop by Lishan Su et al.(Su97). Recombinant virus pNL4-3, with envelope from LAI(BRU) and gag-pol from NY5 has also been studied:(Adachi86) GenBank accession number M19921, (Duensing95) GenBank accession number L42371 and (Salminen95) GenBank accession number U26942. Other GenBank entries with IIIB-LAI sequences can be found in the patented sequences section and in the cloning vector section (for example U19867 and A00647)

**Table 2 (cont.)**

| Sequence | Accession | Origin | Author | Reference |
|---|---|---|---|---|
| B_GA.OYI | M26727 | GABON | Huet, T | *AIDS* **3**(11):707–715 (1989) |

This sequence is derived from the Gabonese isolate OYI, designated elsewhere as isolate 397, was obtained from a healthy HIV-1 infected individual presenting an atypical Western Blot. This sequence is from a lambda phage clone, the cloned provirus being functionally defective. The vpu gene does not have a start codon. Phylogenetic analysis reveals that the sequence is closely related to the North American isolate SF2 and the European virus HAN (across the genome). This is the first report of a virus from Africa that clusters with North American rather than Zairean viruses:OYI and SF2 differ by approximately 7% in envelope. The single C -> S substitution at residue 22 of the OYI tat protein renders it inactive, but may not account for the avirulence of the virus. Sibling sequences for OYI(397) are available (see the 1989 HIV Database Compendium page I-A-181).

| | | | | |
|---|---|---|---|---|
| B_GB.MANC | U23487 | UNITED KINGDOM | Zhu, T | *Nature* **374** (6522): 503–504 (1995) |

Kindly provided in electronic form by Dr. David Ho, Aaron Diamond AIDS Research Center, New York City, (212)-725-0018. This sequence ostensibly represents HIV-1 captured by PCR amplified from the 1959 sample "Manchester sailor" kidney tissue (see Corbitt. G., *et al., Lancet* **336**:51 (1990)). The sequence of the complete genome is available, and it is indistinguishable from contemporary subtype B HIV-1 sequences in phylogenetic analysis. This information together with the observation that additional tissue samples were HIV PCR negative, suggests that the HIV clone that came from MANC kidney sample was very likely to be a contemporary clinical contaminant. The sequence was assembled from multiple PCR amplified fragments. All reading frames in this sequence are intact.

| | | | | |
|---|---|---|---|---|
| B_GB.CAM1 | D10112 | UNITED KINGDOM | McIntosh, AA | Unpublished (1991) |

This sequence is from the British isolate CAM1. It has a defective vpu gene. McIntosh A, and Karpas A, Thesis (1991), Cambridge University, England. GenBank accession numbers D10112, D00917

| | | | | |
|---|---|---|---|---|
| B_NL.3202A21 | U34604 | NETHERLANDS | Guillon, C | *ARHR* **11**(12):1537–1541 (1995) |

This sequence is from a complete genome of a macrophage tropic, NSI clone, from an isolate taken from the PBMC of a patient who was in transition from NSI to SI phenotype bulk phenotype. An SI isolate from the same patient has also been completely sequenced, ACH320.2A.1.2 The patient, isolates and phenotype of the molecular clones are described in Groenink, M., *et al., J Virol* **65**:1968–1975 (1991)

| | | | | |
|---|---|---|---|---|
| B_TW.TWCYS | AF086817 | TAIWAN Province of China | Huang, LM | Unpublished |

This as-yet unpublished sequence contains a frameshifting single-base insertion in the pol gene, at 3203-3209 where a run of 7 "A"s should have only 6.

| | | | | |
|---|---|---|---|---|
| B_US.SF2 | K02007 | UNITED STATES | Sanchez-Pescador, R | *Science* **227** (4686): 484–492 (1985) |

This sequence is from an infectious phage clone from the US isolate ARV-2. ARV-2/SF2 was isolated from the PBMC of a patient with oral candidiasis after co-culture with mitogen stimulated PBMCs, (Levy, J., *et al., Science* **225**:840–842, (1984)). It is a standard reference strain, and has been used for vaccine studies. The isolate SF2 is available from the NIH AIDS Reagent program, and is SI R5X4.

**Table 2 (cont.)**

| Sequence | Accession | Origin | Author | Reference |
|---|---|---|---|---|
| B_US.DH123 | AF069140 | UNITED STATES | Shibata, R | *J Virol* **69**(7):4453–4462 (1995) |

The DH12 isolate has been extensively charachterized. It is dual-tropic, using CCR3, CCR5 and CXCR4 coreceptors. Chimeric molecular clones with the macrophage-tropic AD8 isolate have been made, showing that either V1-V2 or V3 regions of env from DH12 can confer the ability to use CXCR4 onto AD8. The DH12 isolate was passaged in human and chimpazee PBMCs prior to cloning. A complete genome sequence is available. See also GenBank accession number AF069139.

| | | | | |
|---|---|---|---|---|
| B_US.NY5CG | M38431 | UNITED STATES | Willey, RL | *PNAS USA* **83**(14):5038–5042 (1986) |

This sequence is from the 1984 New York T-cell tropic isolate NY5. It was cloned in lambda phage and is not replication competent. It has a defective vpu gene due to the loss of the start codon. Supernatant DNA extracted from A3.01 cells infected with the NY5 HIV isolate stock was digested with EcoRI and cloned into lambda WESB. The insert is an EcoRI permuted single LTR clone and was then transferred into pBR322. See also GenBank accession number K03346, for an env sequence of this isolate. A computer readable copy of this sequence was kindly provided by Chuck Buckler, 01-NOV-1988.

| | | | | |
|---|---|---|---|---|
| B_US.AD8 | AF004394 | UNITED STATES | Theodore, TS | *ARHR* **12**(3):191–194 (1996) |

AD8 is a molecular clone that was replication competent after reconstruction. It has a defective vpu due to the loss of the start codon. It was derived from the macrophage tropic isolate AD87, which in turn was isolated from ADA, described in Westervelt, P., *et al., PNAS* **88**:3097–3101 (1991). The pAD8 HIV-1 was originally cloned as an EcoRI permuted linear DNA from a 1 LTR circle. It was converted to a 2 LTR infectious clone as described [1]. The sequence presented begins at the 5′ end of the LTR and continues through the HIV-1 sequences followed by the vector sequences. The 225 bases at the 3′ end represent a duplication of bases 8851..9075 produced during the construction of the two LTR provirus. The original AD8 clone had a defective vpu gene containing a valine codon instead of an initiator methionine. A functional vpu gene was constructed by changing the valine codon at 6058 to one for the initiator methionine by oligonucleotide-directed mutagenesis. Macrophage tropic. Subtype B. Other receptors defined for this isolate are:CCR3, GPR15, STRL33, CCR8. Second receptor usage for this isolate was defined by Choe *et al, Cell* **85**:1135 (1996), Farzan *et al., J Exp Med* **186**:405 (1997) and Rucker *et al, J Virol* **71**(12):8999–9007 (1997). WARNING:Observed phenotype may not correspond with sequence. Further coreceptor usage data was obtained by Choe *et al J Virol* **72**(3):2509–2515 (1998) by subcloning regions of the env gene from a dual-tropic HIV-1 (isolate DH12) into the AD8 clone. The isolate ADA-M is available from the NIH AIDS Reagent program, and is NSI R5.

| | | | | |
|---|---|---|---|---|
| B_US.WCIPR9018 | U69591 | UNITED STATES | Fang, G | Unpublished (1996) |

This subtype B sequence is from a sample taken in 1990. No information is available about the patient or properties of the virus.

**Table 2 (cont.)**

| Sequence | Accession | Origin | Author | Reference |
|---|---|---|---|---|
| B_US.YU2 | M93258 | UNITED STATES | Li, Y | *J Virol* **65**(8):3973–3985 (1991) |

Other receptors defined for this isolate are:CCR3, GPR15. Second receptor usage for this isolate was defined by Choe *et al, Cell* **85**:1135 (1996) and Farzan *et al, J Exp Med* **186**:405 (1997). WARNING:Observed phenotype may not correspond with this exact sequence, but to the same isolate. A complete genome from another isolate from this same patient (clone YU-10 accession number M93259) has also been sequenced. YU-2 is a lambda phage clone that is replication competent. It was from the uncultured brain tissue of a patient with AIDS dementia complex. YU2 and YU10 differ by 0.26% in the nucleotide sequence. YU2 was fully replication competent after reconstruction in both primary T lymphocytes and monocyte-macrophages. YU-2 has a defective vpu gene due to the loss of the start codon. See also Li, Y., *et al., J Virol* **65**:3973–3985 (1991). Kindly provided in computer-readable format by Beatrice Hahn.

B_US.JRCSF M38429 UNITED STATES O'Brien, WA *Nature* **348**(6296):69–73 (1990)

JRCSF and JRFL (see U63632) were isolated from cerebral spinal fluid and brain tissue respectively of the patient JR, who died with Kaposi's sarcoma and severe AIDS encephalopathy (*Science* **236**, 819–822, 1987). JRCSF is from an infectious lambda phage clone of the 1986 isolate JRCSF. Both clones are infectious, but JRFL productively infects macrophages while JRCSF does not. (Peripheral blood was not available from the patient). Second receptor usage for this isolate was defined by Simmons *et al., J Virol* **70**:8355 (1996) and Zhang *et al, Nature* **383**:768 (1996). The JRCSF and JRFL env nucleotide sequences differ by at least 3%. Both manifest insertions in nef previously reported for HIVBRVA. Patient JR was a 29 year old homosexual male with a history of multiple sexual partners, including one who died of AIDS with AIDS related dementia. Patient JR died in June 1986 and had AIDS related dementia. Complete autopsy findings are reported in [3]. The entry with accession number U45960 is from a mutated clone of JR-CSF which shows a wider range of tropism for cultured cells and increased syncytium-inducing capability. The sequence was kindly provided in computer-readable form by Irvin Chen, UCLA School of Medicine, Los Angeles. On July 16, 1997, Christopher Buck wrote:... [there are] some variations between the published sequence of proviral clone JR-CSF and the actual sequence of plasmid pYK-JRCSF (AIDS Reagent Program catalog #2708). The substitutions are as follows:C7808->G G7809->C G8232->A resulting in these amino acid substitutions in envelope reading frame:R525->A G666->D The isolate JR-CSF is available from the NIH AIDS Reagent program, and is NSI R5. References:Simmons *et al, J Virol* **70**:8355 (1996) MEDLINE:97126031 Zhang *et al, Nature* **383**:768 (1996) MEDLINE:97048157 *Science* **236**:819-822 (1987) MEDLINE:87206194.

B_US.MNCG M17449 UNITED STATES Gurgo, C *Virology* **164**(2):531–536 (1988)

MN is from one of the earliest available isolates, and is a commonly used reference and vaccine strain. The MN isolate was taken from a 6 year old male pediatric AIDS patient from the area of Newark, New Jersey, USA in 1984. His mother was an IV drug user who died of pneumonia in 1982. His father was also HIV seropositive. Other sequences from this patient from the 1984 blood sample and from a 1987 sample taken shortly before death (U72495) are available. See also L48364–L48379. The MN sequence was cloned from the isolate in lambda phage. The coding sequences for pol, nef and vpu are prematurely truncated; pol shows an in-frame stop codon at 3783, nef and vpu are prematurely truncated at position 9357 and position 6142 respectively. A set of V3 sequences from this isolate are available (Accession #s L48364–L48379, Lukashov, V. and Goudsmit, J., *AIDS* **9**:1307–1311 (1995). In Lori et al [2], the pol gene is not noted to be defective in the ST.1 clone of HIV-1 isolate MN. Another complete genome of the MN isolate is available with accession number AF075719 and it too has defective genes, although not pol nor vpu. This sequence was kindly provided in computer readable format by M. Reitz, N.C.I., Bethesda, MD. 20892 U.S.A. The isolate MN is available from the NIH AIDS Reagent program, and is X4.

**Table 2 (cont.)**

| Sequence | Accession | Origin | Author | Reference |
|---|---|---|---|---|
| B_US.BC | L02317 | UNITED STATES | Kong, LI | *Science* **240**(4858):1525–1529 (1988) |

This is one of the most cytopathic isolates of HIV-1 in both human and chimpanzee lymphocytes and does not replicate in monocyte-macrophages. The isolate BC was obtained in 1987 from an individual from the U.S. with terminal AIDS. The clone was obtained after short term co-culture of HUT-78 cells with PHA-stimulated PBMCs. Total cellular DBA from a BC-infected HUT 78 cell culture was harvested, cut with EcoRI, and ligated into Lambda-pGEM-12 to create a genomic library. Three proviral clones (lambda SG3, lambda SG15 and lambda SG39) were identified as containing full-length HIV-1 BC genomes. Only lambda SG3 was replication competent and produced virions whose replicative and cytopathic properties were equivalent to those of the parental BC isolate. The viral DNA in lambda SG3 was subcloned into the pTZ 19U plasmid (USB) and the resulting clone pSG3.1 was used for sequencing. The pSG3.1 plasmid was transfected into COS-1 cells and infectious viral stocks were obtained for furtherm analysis. BCSGS3 is the sequence from the provirus SG3;the full length provirus was cloned intact in lambda phage. BC and BCSG3 replicate more efficiently in chimpanzee than in human lymphocytes. All typical HIV-1 genes are intact except vpu, which is normally 243 bp in length;vpu was disrupted after 70 nucleotides by a 23 bp deletion that results in a frameshift and truncation of an open reading frame.

B_US.P896 U39362 UNITED STATES Collman, R *J Virol* **66**(12):7517–7521 (1992)

89.6 is cloned from a highly cytopathic primary macrophage-tropic and syncytium-inducing isolate. It comes from an AIDS patient living in Philadelphia for 15 years, although the patient was originally from Jamaica. At the time of viral isolation, this 47 year old man had received no antiretroviral therapy and had advanced immunodeficiency. 89.6 was cloned in phage and was replication competent after reconstruction. Other receptors defined for this isolate are:CCR3, CCR2b, CCR8, V28. Second receptor usage for this isolate was defined by Farzan *et al, J Exp Med* **186**:405 (1997) and Rucker *et al., J Virol* **71**(12):8999–9007 (1997). WARNING:Observed phenotype may not correspond to the sequence. Also see Kim, F., *et al., J Virol* **69**:1755–1761 (1995). The 89.6 env gene was substituted into the SHIV-4 chimeric simian-human virus by (Reimann *et al J Virol* **70**:3198 (1996)) see accession numbers U89134 and AF038398. The isolate 89.6 is available from the NIH AIDS Reagent program and is listed as SI R5X4.

B_US.WEAU160 U21135 UNITED STATES Ghosh, SK Unpublished (1995)

Sequence kindly provided by Sajal K. Ghosh, UAB, Birmingham. A cytopathic HIV-1 virus was cloned from an acutely infected patient in 1990. The clone WEAU 1.60 is replication competent and upon transfection produces highly cytopathic T-cell tropic virus. The clone and the viral isolate from which it was derived are syncytium-inducing (SI). Genbank accession number U21135. The WEAU 1.60 clone was obtained from a coculture of the patient's PBMCs, first with normal donor PHA-stimulated lymphocytes for 14 days, then with the H9 T-cell line for another 14 days. The patients' blood specimen was obtained 15 days after the onset of clinical symptoms of acute (primary) infection, and 35 days after a single sexual encounter with a partner whose virus was proven phylogenetically to be responsible for the transmission event. The patient is identified as "Patient #1" in *N. Engl. J. Med.* **324**:954–960 (1991) and as "WEAU 0575" in *Science* **259**:1749–1754 (1993). The patient is also discussed in Borrow *et al., Nat Med*, **3**:205–11 (1997). The WEAU 1.60 clone has been completely sequenced from a plasmid. It is subtype B. There is a deletion of a single T at position 9069, resulting in a frameshift mutation and premature termination of nef. The frameshifting deletion in nef was NOT present in the patients' unclutured PBMCs where instead there is a "T". The nef gene was not interrupted in 10 of 10 clones analyzed by PCR sequencing from the uncultured PBMCs. It has been sequenced in its entirety by two different labs (G. Shaw and L. Hood) with 100% concordance.

**Table 2 (cont.)**

| Sequence | Accession | Origin | Author | Reference |
|---|---|---|---|---|
| B_US.RF | M17451 | UNITED STATES | Starcich, BR | *Cell* **45**(5):637–648 (1986) |

RF (also designated HAT because the virus was isolated from a patient of Haitian descent) is among the first isolates, and is among the commonly used reference and vaccine study strain. The sequence is from the full-length lambda phage clone HAT-3, from isolate RF, cultured in HUT-78 cells. RF is from a 28 year old symptomatic Haitian male, who moved to the U.S. at age 25, in 1980 and was sampled in 1983, shortly before his death in Dec. 1983. He had no history of IV drug use, homosexuality or blood transfusions. In October 1983, he had 20 lb weight loss, giardia with diarrhea, thrush, and diffuse lymphadenopathy. His CD4/CD8 ratio was 0.08. Primary culture from a November 1983 blood sample was co-cultured on HUT-78 cells. GenBank accession numbers are M17451 and M1250. RF has defective gag and vpu genes. Several env genes are available from this isolate. GenBank accession numbers are U30778–U30781. The sequenced clone did not have the base 'a' at position 640 required for gag translation. Two differences in the restriction map arise in comparison to an earlier published map for lambda-HAT (Hahn,B.H. *et al., PNAS USA* **82**:4813 (1985)):i) a Bgl-II site is found at position 193 of the sequence and ii) the HindIII site reported by Hahn et al. at position 2000 is not present in this sequence. See also Reitz M., *et al., ARHR* **8**:1950 (1992)

| | | | | |
|---|---|---|---|---|
| B_US.WR27 | U26546 | UNITED STATES | Salminen, MO | *N Engl J Med* **314**(2):131–2 (1986) |

The clinical isolate WR27 was from a USA patient with a first seropositive sample in 1987 and with clinical progression to Walter Reed stage 5 (Disease stages described in Redfield *et al. New Engl. J. Med.* **314**:131–32 (1986)). Blood was drawn for viral isolation in Sept. 1988. The virus was cultured on seronegative donor PBMCs prior to PCR amplification of nearly complete genome from PBMC proviral DNA. The viral sequence had a V3 loop predictive of an SI phenotype. All reading frames were intact except for rev, which had an inframe stop codon in both exons.

| | | | | |
|---|---|---|---|---|
| C_BR.92BR025 | U52953 | BRAZIL | Gao, F | *J Virol* **70**(3):1651–1667 (1996) |

This sequence is from a PCR clone from a primary isolate that is part of a set obtained through WHO Global Programme on AIDS (WHO Network, *ARHR* **10**:1327–1344 (1994)). It is from a 23 year old male hemophilia patient from Porto Alegre, Brazil. He had seroconverted more than 1.2 months prior to the date this blood sample was collected in 1992. He was asymptomatic, and had not taken any anti-retroviral therapy prior to sampling. 92BR025 was established and propagated by short-term co-cultivation with normal donor lymphocytes, and then the near full length genome was PCR amplified and sequenced. The HIV isolate exhibited an NSI phenotype, when assayed by the WHO. See also entries with accession numbers U09126, U15121 and U52953. The full length genome is clone 8, 92BR025.8. This clone has two inframe stop codons in pol at positions 2141, and 3115, and a frame shift mutation at position 4131. This sequence was kindly made available prior to publication. Additional env, nef and ltr region sequences are available from this isolate:U09126, U09132, U51282, and U15121 Biotypes were determined by MT-2 syncytium assay; however, both syncytium-inducing (SI) and non-syncytium-inducing (NSI) variants may be present in the viral "swarm" for each isolate. Recent studies indicate that NSI isolates contain predominantly CCR5-using variants while most SI isolates contain both CXCR4 (SI) and CCR5 (NSI) variants. Some SI isolates may contain dual-tropic variants that use both CXCR4 and CCR5 co-receptors. The isolate 92BR025 is available from the NIH AIDS Reagent program, and is NSI R5. A small region of Gag, beginning very near the junction of p1/p6 (position 2131 in the HXB2R K03455 genome; 1471 in this genome) and ending before the end of p6 (position 2237 in HXB2R; position 1579 in this genome) was discovered to be subtype B at the HIV Database. This region includes 3 in/dels that are highly indicative of either subtype B or C, that contribute to the B-like nature of this region in 92BR025, but which would be overlooked if gapstripping were used prior to bootscanning.

**Table 2 (cont.)**

| Sequence | Accession | Origin | Author | Reference |
|---|---|---|---|---|
| C_BW.96BW0402 | AF110962 | BOTSWANA | Novitsky, VA | *J Virol* **73**(5):4427–32 (1999) |
| C_BW.96BW1104 | AF110969 | BOTSWANA | Novitsky, VA | *J Virol* **73**(5):4427–32 (1999) |
| C_BW.96BW15C02 | AF110974 | BOTSWANA | Novitsky, VA | *J Virol* **73**(5):4427–32 (1999) |
| C_BW.96BW0502 | AF110967 | BOTSWANA | Novitsky, VA | *J Virol* **73**(5):4427–32 (1999) |
| C_BW.96BW01B03 | AF110959 | BOTSWANA | Novitsky, VA | *J Virol* **73**(5):4427–32 (1999) |
| C_BW.96BW16B01 | AF110976 | BOTSWANA | Novitsky, VA | *J Virol* **73**(5):4427–32 (1999) |
| C_BW.96BW1210 | AF110972 | BOTSWANA | Novitsky, VA | *J Virol* **73**(5):4427–32 (1999) |
| C_BW.96BW17B03 | AF110980 | BOTSWANA | Novitsky, VA | *J Virol* **73**(5):4427–32 (1999) |

These subtype C sequences are from Botswana. They have 4 NF-Kappa B binding sites where most subtype C have 3, and most other subtypes have just 2. C_BW.96BW0402 was kindly provided as a reference strain prior to publication by Dr. Vlad Novitsky, and is part of a study of multiple full length C subtype sequences from Botswana. The pol protease region of the 96BW17 genomes clusters with two isolates from Zimbabwe (Z1226 AF083262, and Z84 AF083267) quite outside the subtype C clade. These 3 isolates may thus be representatives of a circulating recombinant form, with 3 isolates from 2 countries identified to date (Aug 1999). However, presently there are no three complete genomes available, a requirement for an 'official' CRF.

C_ET.ETH2220 U46016 ETHIOPIA Alaeus, A *ARHR* **12**(14):1329–1339 (1996)

ETH2220 is the first reported (almost full length) subtype C sequence from Ethiopia. The patient from which this clone was obtained was taken in 1986. In its genomic organization, this clone closely resembles subtype A, B, and D isolates except that the core promoter contains three potential binding sites for the transcription factor NF-kB instead of containing two. This is a feature that was preserved in other Ethiopian C subtype samples, as well as C viruses from Zambia. This sequence was cloned as a PCR amplified near full length genome, and has a defective tat gene.

C_IN.93IN999 AF067154 INDIA Lole, KS *J Virol* **73**(1):152–160 (1999)

This subtype C sequence is one of several complete genomes from India. This subtype C sequence is one of several complete genomes from India. It is derived from primary PBMC cocultures taken April 24, 1993 from a 52 year old man from Pune in Maharashtra State, India who seroconverted in 1992; risk factors included sex with men and commercial sex workers; NSI phenotype. This sample is available in plasmid form through the NIH AIDS Research and Reference Reagent Program and is named 93IN999, and is R5. A small section of the 5′ LTR sequence present in the 301999 virus is not included in this sequence.

C_IN.95IN21068 AF067155 INDIA Lole, KS *J Virol* **73**(1):152–160 (1999)

A small section of the 5′ LTR sequence present in the 21068 virus is not included in this sequence. It is derived from primary PBMC coculture taken Feb. 18, 1995 from a 21 year old man from Pune in Maharashtra State, India who seroconverted in 1994. His only identified risk factor for HIV infection was genital ulcer disease. This sample is available in plasmid form through the NIH AIDS Research and Reference Reagent Program and is named p95IN21068.

**Table 2 (cont.)**

| Sequence | Accession | Origin | Author | Reference |
|---|---|---|---|---|
| C_IN.93IN905 | AF067158 | INDIA | Lole, KS | *J Virol* **73**(1):152–160 (1999) |

This subtype C sequence is one of several complete genomes from India. It is derived from an primary culture PBMC sample taken March 27, 1993 from a woman from Pune in Maharashtra State, India who seroconverted in 1992 following a blood transfusion; NSI phenotype.A small section of the 5′ LTR sequence present in the 301905 virus is not included in this sequence. This sample is available in plasmid form through the NIH AIDS Research and Reference Reagent Program and is named 93IN905, and is R5.

C_IN.93IN904 AF067157 INDIA Lole, KS *J Virol* **73**(1):152–160 (1999)

This subtype C sequence is one of several complete genomes from India. It is derived from a cocultured PBMC sample taken March 27, 1993 from a 28 year old woman from Pune in Maharashtra State, India who seroconverted in 1992 following a blood transfusion; NSI phenotype. A small section of the 5′ LTR sequence present in the 301904 virus is not included in this sequence. This sample is available in plasmid form through the NIH AIDS Research and Reference Reagent Program and is named p93IN301904. The isolate 93IN904 is available from the NIH AIDS Reagent program, and is R5. Reference:Lole, K.S. et al., J Virol 73(1):152-60 (1999)

C_IN.94IN11246 AF067159 INDIA Lole, KS *J Virol* **73**(1):152–160 (1999)

This subtype C sequence is one of several complete genomes from India. It is derived from primary PBMC cocultures taken Oct. 25, 1994 from a 26 year old man from Pune in Maharashtra State, India who seroconverted in 1994. His major risk factor for HIV infection was sexual contact with commercial sex workers. A small section of the 5′ LTR sequence present in the 11246 virus is not included in this sequence. This sample is available in plasmid form through the NIH AIDS Research and Reference Reagent Program and is named p94IN11246.

D_UG.94UG1141 U88824 UGANDA Gao, F *J Virol* **72**(7):5680–98 (1998)

Sample 94UG114 was obtained from an asymptomatic 31-year-old man from Butuku, Uganda as part of the WHO/UNAIDS study. He had not taken any anti-retroviral therapy prior to sampling. His risk factor for infection was heterosexual contact. The near full length genome was PCR amplified from a short term culture of a PBMC sample and sequences. The isolate from which this sequence was derived is NSI by an MT-2 assay. This sequence was kindly made available prior to publication by Feng Gao. Biotypes were determined by MT-2 syncytium assay; however, both syncytium-inducing (SI) and non-syncytium-inducing (NSI) variants may be present in the viral "swarm" for each isolate. Recent studies indicate that NSI isolates contain predominantly CCR5-using variants while most SI isolates contain both CXCR4 (SI) and CCR5 (NSI) variants. Some SI isolates may contain dual-tropic variants that use both CXCR4 and CCR5 co-receptors. The isolate 94UG114 is available from the NIH AIDS Reagent program, and is NSI R5.

D_CD.NDK M27323 (FORMER ZAIRE) Dem Rep of the Congo Spire, B *Gene* **81**(2):275–284 (1989)

Kindly provided prior to publication by J.-C. Chermann, Pasteur Institute, Marseille. This is an infectious molecular clone of an isolate estimated to be 10,000 times more cytopathic in vitro than a prototypical HIV-1 (IIIB/LAI). Only minor sequence differences appear to be responsible for the 'acute biological effect'. This sequences clusters with HIV-1 subtype D in phylogenetic analysis. It was cloned in phage and is replication competent. All reading frames in this sequence are intact. The NDK virus was isolated from a Zairean man residing in France. He had AIDS and infected his wife; she infected her lover. The cytopathicity of this virus was found not to be localized to env, and might be regulated by regions of gag, vpr and env.

**Table 2 (cont.)**

| Sequence | Accession | Origin | Author | Reference |
|---|---|---|---|---|
| D_CD.ELI | K03454 | (FORMER ZAIRE) Dem Rep of the Congo | Alizon, M | *Cell* **46**(1):63–74 (1986) |

This sequence is of a phage clone derived from the Zairean isolate ELI. ELI was recovered in 1983 from a 24 year old woman with AIDS. All reading frames in this sequence are intact. This sequence is from the Zairean isolate ELI. The complete genomic sequence and an infectious clone are available. In the 1995 Compendium (pages III-45 and III-47), ELI was listed as an unlikely D/A mosaic, with only gp41 being weakly A-like. Entry with accession number M27949 is from this same isolate.

| | | | | |
|---|---|---|---|---|
| D_CD.Z2Z6 | M22639 | (FORMER ZAIRE) Dem Rep of the Congo | Srinivasan, A | *J Virol* **72**(7):5680–98 (1998) |

An infectious molecular clone of this virus was reconstructed. It was cloned in phage and sequenced from the isolate Z2, also called CDC-Z34. All reading frames in this sequence are intact. Total cell DNA extracted from A3.01 cells, infected with the Z2 HIV isolate stock, was digested with XhoI and cloned into lambda J1. The insert is an XhoI permuted single LTR clone, which was then transferred into pIBI. In the sequence below, position one is the first base of the single LTR of the clone, while the last base (9081) is the one just before the LTR of the intact circle. HIVZ2Z6 is an infectious clone and is from the same isolate (CDC Z34;P.Feorino) as HIVZ6. The entry with accession numbers K03458 and M16322, is also from the same isolate. It is defective in Vpr, whereas this entry has an intact Vpr CDS.

| | | | | |
|---|---|---|---|---|
| D_CD.84ZR085 | U88822 | (FORMER ZAIRE) Dem Rep of the Congo | Gao, F | *J Virol* **72**(7):5680–98 (1998) |

Sample 84ZR085 was obtained from an AIDS patient from Zaire. The near full length genome was cloned in phage and sequenced. D-84ZR085 is subtype D. There was a frame shift mutation in gag/pol, position 1692. This isolate was obtained from Thomas Jefferson University, and isolate phenotyping information was not available. This sequence was kindly made available prior to publication. Phylogenetic analysis of the pol protease region done at the HIV-DB indicates that this sequence region is an outlier of the subtype D clade.

| | | | | |
|---|---|---|---|---|
| F1_BE.VI850 | AF077336 | BELGIUM | Carr, JK | Unpublished |

Small sections of the 5′ and 3′ LTRs are not included in this sequence. This sequence was isolated from a Belgian man in 1993 whose wife was infected in Zaire (now called the Dem Rep of the Congo). This sequence was kindly provided prior to publication by J. Carr *et al.* This sequence was originally classified as subtype F, but because F is readily subdivided into two distinct categories, it has been reclassified as subtype F1. See Triques *et al, Virology.* **259**(1):99–109 (1999).

**Table 2 (cont.)**

| Sequence | Accession | Origin | Author | Reference |
|---|---|---|---|---|
| F1_BR.93BR020.1 | AF005494 | BRAZIL | Gao, F | *J Virol* **72**(7):5680–98 (1998) |

This sequence was originally classified as subtype F, but because F is readily subdivided into two distinct categories, it has been reclassified as subtype F1. See Triques *et al, Virology.* **259**(1):99–109 (1999). This sample is part of a set of sequences generated through the WHO Global Programme on AIDS (WHO Network, *ARHR* **10**:1327–1344 (1994) and came from an asymptomatic HIV seropositive bisexual contact. The isolate 92BR020 was established and propogated by short term co-cultivation with normal donor lymphocytes, and then the near full length genome was PCR amplified and sequenced. The isolate 92BR020 was described as syncytium inducing (SI) using an MT-2 assay. An envelope gene sequence from this isolate is described in Gao, F., *et al., J Virol* **70**:1651–1657 (1996). This sequence was kindly made available prior to publication, and was eventually published in Gao, F. *et al., J Virol* **72**(7):5680–98 (1998). There were no defective genes. A summary of isolates with known co-receptor usage can be found in the HIV database reviews.

| | | | | |
|---|---|---|---|---|
| F1_FI.FIN9363 | AF075703 | FINLAND | Laukkanen, T | Unpublished |

The virus was isolated in 1993 from a Finnish male who claims that he was most likely infected in Finland in 1985 by a Kenyan woman. However, he had had multiple partners. This sequence was originally classified as subtype F, but because F is readily subdivided into two distinct categories, it has been reclassified as subtype F1. See Triques *et al., ARHR* in press 1999.

| | | | | |
|---|---|---|---|---|
| F1_FR.MP411 | AJ249238 | FRANCE | Triques, K | *ARHR* **16**(2):139–151 (2000) |

Prior to 1999, the F subtype was made up of 3 distinctive clusters and very diverse viruses. Triques and colleagues have defined a new subtype K, and broken F into two subclusters, F1 and F2. This sequence is a complete genome reference for F1. This isolate was from a French patient who believes he was infected when deployed in Chad or Yugoslavia.

| | | | | |
|---|---|---|---|---|
| F2_CM.MP255 | AJ249236 | CAMEROON | Triques, K | *ARHR* **16**(2):139–151 (2000) |

Prior to 1999, the F subtype was made up of 3 distinctive clusters and very diverse viruses. Triques and colleagues have defined a new subtype K, and broken F into two subclusters, F1 and F2. This sequences is a complete genome reference for F2.

| | | | | |
|---|---|---|---|---|
| F2_CM.MP257 | AJ249237 | CAMEROON | Triques, K | *ARHR* **16**(2):139–151 (2000) |

Prior to 1999, the F subtype was made up of 3 distinctive clusters and very diverse viruses. Triques and colleagues have defined a new subtype K, and broken F into two subclusters, F1 and F2. This sequences is a complete genome reference for F2.

**Table 2 (cont.)**

| Sequence | Accession | Origin | Author | Reference |
|---|---|---|---|---|
| G_BE.DRCBL | AF084936 | BELGIUM | Debyser, Z | *ARHR* **14**(5):453–459 (1998) |

Clinical details are discussed in *ARHR* **14**(5):453–9 (1998) and the analysis of the complete genome is in Oelrichs *et al., ARHR* **15**(6):585–9 (1999). A pregnant 26 year old women was sampled who had lived in Zaire (now called the Dem Rep of the Congo) until 1993, then moved to Belgium. She was diagnosed with AIDS and had a low CD4 when sampled in 1996. Her G subtype virus was not detected by Amplicor Monitor or Nasba RNA kits, although she was found to have a high viral load by branched DNA. The sequence was kindly provided prior to publication by R. Oelrichs *et al*. This complete genome sequence shows the same pattern of phylogenetic associations as 92NG083 (U88826), HH8793 (AF061640) and SE6165 (AF061642). These four (and other subtype G sequences) form their own clade (subtype G) when the complete gag, pol or env genes are included in the analysis. However, all G subtype genomes have some ambiguous A/G regions in the central part of the genome in a phylogenetically indistinct region in the the accessory gene region from the beginning of vif to the beginning of vpu. The coordinates of this region are 5055 to 6297 on HXB2, and there are several subtypes which become difficult to resolve unambiguously in this region:A, G, and the circulating recombinant forms which resemble the prototypes AE(CM240) and AG(IbNG) (per. comm., Jean Carr). Of particular note is that a region of gp41 from these viruses clusters with the AE and IbNG circulating recombinant forms. See J. Carr *et al Virology* **247**:22–31 (1998) and F. Gao *et al J Virol* **72**(7):5680–5698 (1998) for analyses of the other genomes with this pattern. At this time (Jan. 1999) it is not clear whether the AE(CM240) circulating recombinant form is AEG triple recombinant, or if the above four genomes are AEG triple recombinant, or if an evolutionary anomaly rather than recombination is the basis for this pattern.

G_FI.HH8793-12.1 AF061641 FINLAND Salminen, MO *ARHR* **8**(9):1733–1742 (1992)

Sections of the 5′ and 3′ LTRs are not included in this sequence. This sample was taken in June 1993 (Jean Carr, Pers. Communication)

G_NG.92NG083 U88826 NIGERIA Gao, F *J Virol* **72**(7):5680–98 (1998)

This sequence is from a PCR clone from a primary culture from the NSI isolate 92NG083; the sample was taken in 1992 from an AIDS patient from Jos, Nigeria. The isolate was originally called JV1083, but was renamed 92NG083 to be consistent with WHO nomenclature. The full length clone has an altered initiation codon at position 157, an inframe stop codon at position 360 in gag, and a vpu frameshift mutation at position 5462.

G_SE.SE6165 AF061642 SWEDEN Carr, JK *Virology* **247**(1):22–31 (1998)

Siblings sequences from the same blood sample:L40743, L40761, L40752 Set:Two female sex parters of this individual Set_IDs:L40744, L40745, L40753, L40754, L40762, L40763. Sections of the 5′ and 3′ LTRs are not included in this sequence. This patient (6165) was infected in Congo and moved to Sweden. He had tested HIV positive (ELISA and Western Blot) aproximately 18 months prior to infecting patient 6167 via heterosexual intercourse and 19 months prior to infecting patient 6168, also via heterosexual intercourse. He had low (35 per ul) CD4 count and dermatological problems but no AIDS defining illness at the time of sampling in 1993.

H_BE.VI991 AF190127 BELGIUM Laukkanen, T Unpublished

H_BE.VI997 AF190128 BELGIUM Laukkanen, T Unpublished

**Table 2 (cont.)**

| Sequence | Accession | Origin | Author | Reference |
|---|---|---|---|---|
| H_CF.90CF056 | AF005496 | CENTRAL AFRICAN REPUBLIC | Murphy, E | *ARHR* **9**(10):997–1006 (1993) |

This sequence clusters with available HIV-1 subtype H sequences in phylogenetic analysis, and is the first available full length H subtype sequence . The isolate comes from Bangui, in the Central African Republic, and was sampled in 1990, from an asymptomatic individual, who had no anti-retroviral therapy. The isolate had an NSI phenotype by an MT-2 assay, and the sample was obtained from the Pasteur Institute, Bangui. The isolate 90CF056 was established and propagated by short term cocultivation with normal donor lymphocytes, and then the near full length genome was PCR amplified, cloned and sequenced. The isolate was at one point designated 90CR056, but was changed to 90CF056 as CR stands for Costa Rica, and CF for Central African Rebuplic. The first genetic characterization of this virus isolate (an env V3 sequence designated 4056, GB accession number L11497, Murphy *et al., ARHR* **9**:997–1006 (1993)) left the subtype designation as unclassified, but a second study of this env region sequence classified it as subtype H (W. Janssens, *ARHR* **10**:877–879 (1994)). This sequence was kindly made available prior to publication, and was the first subtype H full length genome available (Gao, F. *et al., J Virol* **72**(7):5680–98 (1998)) The patient was heterosexual, asymptomatic, and the biological phenotype of the isolate was NSI. There were no defective genes in the sequence.

J_SE.SE7022 AF082395 SWEDEN Laukkanen, T *ARHR* **15**(3):293–297 (1999)

This sequence is from a woman who was infected in Zaire (now called the Democratic Republic of the Congo) between 1981 and 1986. Blood for sequencing was drawn in 1993. She was asymptomatic with a CD4 count of 184. The sequence was kindly provided prior to publication by M. Salminen. This sequence is from the same individual as SE7022 described by T. Leitner *et al ARHR* **11**(8):995–997 (1985), see accession numbers L41177 and L41179 for env and gag genes from this individual. Other examples of subtype J have been found in Gambia, see accession numbers U33099, U33100 and U33102.

J_SE.SE7887 AF082394 SWEDEN Laukkanen, T *ARHR* **15**(3):293–297 (1999)

This sequence is from a male who was infected in Sweden between 1993 and 1994. Blood for sequencing was drawn in 1994. He was asymptomatic with a CD4 count of 567. The sequence was kindly provided prior to publication by M. Salminen. This sequence is from the same individual as SE7887 described by T. Leitner *et al ARHR* **11**(8):995–997 (1985), see accession numbers L41176 and L41178 for env and gag genes from this individual.

K_CM.MP535 AJ249239 CAMEROON Triques, K *ARHR* **16**(2):139–151 (2000)

Prior to 1999, the F subtype was made up of 3 distinctive clusters and very diverse viruses. Triques and colleagues have defined a new subtype K, and broken F into two subclusters, F1 and F2. This sequences is a complete genome reference for subtype K.

K_CD.EQTB11C AJ249235 (FORMER ZAIRE) Dem Rep of the Congo Triques, K *ARHR* **16**(2):139–151 (2000)

Prior to 1999, the F subtype was made up of 3 distinctive clusters and very diverse viruses. Triques and colleagues have defined a new subtype K, and broken F into two subclusters, F1 and F2. This sequences is a complete genome reference for subtype K.

**Table 2 (cont.)**

| Sequence | Accession | Origin | Author | Reference |
|---|---|---|---|---|
| CRF01_AE_CF.90CF402 | U51188 | CENTRAL AFRICAN REPUBLIC | Gao, F | *J Virol* **70**(10):7013–7029 (1996) |

One of a set of three complete genomes from a study linking the HIV-1 epidemic in the heterosexual population in Thailand to an A/E recombinant. 90CF402, previously named CAR-E 4002 or 90CR402, was obtained from a man from Bangui, Central African Republic, who had lymphadeopathy, diarrhea, severe weight loss and recurrent respiratory infections. He was infected through heterosexual contact, but the year of infection is unknown. 90CR402, GenBank accession number U51188, was first adapted to growth in chimpanzee cells, expanded in chimpanzee cells, and then re-expanded in human PBMCs before lambda cloning and sequencing. The sequence has a defective vpu due to the loss of a start codon, and a defective vif gene. The pattern of subtype A-E recombination of breakpoints is shared between A-E subtype sequences from Thailand and from the Central African Republic, suggesting a shared ancestral recombined virus that arose prior to the subsequent epidemics in the two areas.

| | | | | |
|---|---|---|---|---|
| CRF01_AE_TH.93TH253 | U51189 | THAILAND | Gao, F | *J Virol* **70**(10):7013–7029 (1996) |

This virus was isolated in 1993 from a 21 year old man from Chiang Mai, Thailand, who had end-stage AIDS. The isolate was previously designated CMU010, or 302053. The isolate was expanded in donor PBMCs, then in H9 cells, then a lambda phage clone was generated and sequenced. The sequence has a defective vpu due to the lack of a start codon and a defective env gene. Like other 'E' subtype viruses from both Asia and Africa, large stretches of the genome are associated with the A subtype, and all share a common mosaic pattern of A/E breakpoints, suggesting that the currently identified A/E recombinants all share a common ancestor. The breakpoints are mapped in Robertson, D., part III pages 25–30, of the 1997 compendium.

| | | | | |
|---|---|---|---|---|
| CRF01_AE_TH.CM240 | U54771 | THAILAND | Carr, JK | *J Virol* **70**(9):5935–5943 (1996) |

Blood from an asymptomatic heterosexual 21-year-old Thai man was transported from Thailand to the USA where PBMCs were separated and co-cultivated with PHA-stimulated donor PBMCs. DNA from p24 antigen-positive culture was used to amplify the proviral DNA. The complete genomic sequence of the provirus was determined by the compilation of three clones containing different parts of the viral genome. CM240 is an example of a Thai subtype E virus, which is a mosaic of a clade A virus and clade E virus, with the gag gene (and other regions) of sutbype E viral genome falling within clade A in phylogenetic analysis. This is the pattern of A-E sequences found throughout Asia and Africa, and no full length E subtype reference strain has been identified (as for 93TH253). Carr *et al.*, provide detailed analysis of the breakpoints, and point out that the A/E mosaic genomes have a natural pseudotype structure where the external envelope protein spikes on the virion essentially are contributed by the E subtype, and the rest of the viral proteins have a subtype A origin. See also the env sequence from the same isolate (L14572), Mascola J., *et al., JID* **169**:48–54 (1993).

| | | | | |
|---|---|---|---|---|
| CRF02_AG_FR.DJ263 | AF063223 | FRANCE | Carr, JK | *Virology* **247**(1):22–31 (1998) |

Carr *et al.* states that this virus was from a French foreign legion soldier assigned to peacekeeping duties in Djbouti, referencing Louwagie *et al J Virol* **69**(1):263–271 (1995). However the Louwagie paper does not mention the French soldiers and only states that the blood sample was from Djbouti. The sample was isolated in 1991. There are several sequences which share AG recombination breakpoints with IbNG, and are essentially the same recombinant recirculating form; IbNG is the prototype, sharing a similar structure with DJ264 and DJ263.

**Table 2 (cont.)**

| Sequence | Accession | Origin | Author | Reference |
|---|---|---|---|---|
| CRF02_AG_FR.DJ264 | AF063224 | FRANCE | Carr, JK | *Virology* **247**(1):22–31 (1998) |

A small section of LTR sequence present in the DJ263 virus is not included in this file. Carr *et al.* states that this virus was from a French foreign legion soldier assigned to peacekeeping duties in Djbouti, referencing Louwagie *et al J Virol* **69**(1):263–271 (1995). However the Louwagie paper does not mention the French soldiers and only states that the blood sample was from Djbouti. The sample was isolated in 1991. There are several sequences which share AG recombination breakpoints with IbNG, and are essentially the same recombinant recirculating form; IbNG is the prototype, sharing a similar structure with DJ264 and DJ263.

| | | | | |
|---|---|---|---|---|
| CRF02_AG_NG.IBNG | L39106 | NIGERIA | Carr, JK | *Virology* **247**(1):22–31 (1998) |

HIV-1 IbNg was isolated from the PBMCs of an apparently healthy 23 year old man from Nigeria. The patient's PBMCs were cocultured with PHA-stimulated donor PBMCs from an HIV sero-negative donor. After confirming HIV infection in the culture, a mixture of cells and culture supernatant were used to infect a second culture of donor PBMCs, with fresh PHA-stimulated PBMCs added on days 4 and 6. The cultured cells were harvested on day 8 and cytoplasmic RNA was harvested. RT-PCR was used to amplify the complete HIV-1 genome in 5 overlapping segments. The partial env gene sequence (U48628) was originally designated subtype A (Howard, T., et al., ARHR 10:1755-1757 (1994));as was the full length genome. The full length sequence was eventually shown to be an A-G recombinant with multiple cross-over points (Gao F, *et al., J. Virol* **70**:7013 (1996)). The breakpoints are mapped in (Robertson, D., *et al.*, part III pages 25–30, of the 1997 compendium). The IbNg sequence has a 16 bp insertion within the Lys-tRNA primer binding site, just 3′ of the 5′ LTR. It also has a single nucleotide deletion in tat cds at position 5449. See also the entry with accession number U48628, which is from another isolate taken from this same individual. There are several sequences which share recombination breakpoints with IbNG, and are essentially the same recombinant circulating form; IbNG is the prototype, sharing a similar structure with DJ264 and DJ263 (Carr *et al., Virology* **247**:22–31 (1998)).

| | | | | |
|---|---|---|---|---|
| CRF03_AB_RU.KAL153-2 | AF193276 | RUSSIAN FEDERATION | Liitsola, K | *AIDS* **12**(14):1907–19 (1998) |

This is the first complete genome sequence of the gag-A/env-B circulating recombinant form which is common among IV drug users in the Kaliningrad region of Russia. The nonrecombinant subtype A and subtype B parents of this recombinant are common in southern Ukraine and in Russia. A gag gene sequence from this same patient is available with accession number AF082414.

**Table 2 (cont.)**

| Sequence | Accession | Origin | Author | Reference |
|---|---|---|---|---|
| CRF04_cpx_CY.94CY032-3 | AF049337 | CYPRUS | Gao, F | *J Virol* **72**(12):10234–41 (1998) |

This sample, like others in this study (see also subtypes A, B, C, and F) were collected in February 1994 from the AIDS clinic in Nicosia, Cyprus. Patient HO31 was a 24 year old asymptomatic female known to have been HIV seropositive for at least 5 years. Patient HO32 was a 35 year old asymptomatic male, also seropositive for at least 5 years. Both were IVDUs who had lived in Greece and used IV drugs there, before moving to Cyprus. DNA was extracted from patient PBMCs and PCR amplified. Products were cloned and sequenced. Two env gene clones from HO32 and one from HO31 were sequenced (accession numbers U28672, U28673 and U28685). Complete genome is only available for HO32 (CY032). For patient information see Kostrikis,L.G. *et al J Virol* **69**:6122–6130 (1995). This sequence has the same genetic recombination breakpoints as PVCH and PVMY, and 94CY032 is the prototype of the circulating recombinant form. The Gao *et al.* 1998 paper characterizes subtype I in greater detail, presenting the fist published account of this full length genome. The analysis of C2-V3 env gene sequences confirmed that 94CY032.3 was closely related to sequences previously classified as subtype I. However, the remainder of its genome various regions in which 94CY032.3 was significantly clustered with either subtype A or subtype G. Only regions in vpr, nef, and the middle portions of pol and env, formed independent lineages roughly equidistant from all other known subtypes. Since these latter regions most likely have a common origin, Gao et al. classified them all as subtype I, and report that 94CY032 represents a triple recombinant (A/G/I) with at least 11 points of recombination crossover. Since subtype I is now obsolete and has been found to consist of segments of subtype K as well as unknown regions, the new designation for this CRF should be AGKU.

CRF04_cpx_GR.97PVMY AF119819 GREECE Nasioulas, G *ARHR* **15**(8):745–758 (1999)

The sequence was isolated from a 13 year old whose mother and father were IVDUs. The isolate is also called GR84. This sequence has the same genetic recombination breakpoints as CY032 and PVCH, and is one of the circulating recombinant forms of which CY032 is the prototype. See the previous entry for more information on the mosaic pattern.

CRF04_cpx_GR.97PVCH AF119820 GREECE Nasioulas, G *ARHR* **15**(8):745–758 (1999)

The sequence isolated from patient GR11 (accession AF049292) is from the same patient as 97PVCH AF049292. The patient was a 32 year old male IVDU with symptoms (CDC stage B3) in 1991, when sampled. He is no longer living. For patient information see information on patient GR11 in Nasioulas, G. *et al. ARHR* **14**(8):685–90 (1998). This sequence has the same genetic recombination breakpoints as CY032 and PVMY, and is one of the circulating recombinant forms of which CY032 is the prototype.

AC_IN.21301 AF067156 INDIA Lole, KS *J Virol* **73**(1):152–160 (1999)

A small section of the 5′ LTR sequence present in the 21301 virus is not included in this sequence. This subtype AC sequence is one of several complete genomes from India. It is derived from primary PBMC cocultures taken April 3, 1993 from a 40 year old man from Pune in Maharashtra State, India who seroconverted in 1995; his risk factor was sex with commercial sex workers. This sample is available in plasmid form through the NIH AIDS Research and Reference Reagent Program and is named p95IN21301. Phylogenetic analysis at the HIV-DB shows that the subtype C regions of this genome cluster with other Indian subtype C sequences. This indicates that the A/C recombination event likely took place in India, rather than suggesting two introductions of HIV-1 (C and A/C) from outside.

**Table 2 (cont.)**

| Sequence | Accession | Origin | Author | Reference |
|---|---|---|---|---|
| AC_RW.92RW009 | U88823 | RWANDA | Gao, F | *J Virol* **72**(7):5680–98 (1998) |

See also U16220, U08793, for envelope gene sequences from this same patient. The isolate 92RW009 is part of a set generated through the WHO Global Programme on AIDS. The virus was derived from an asymptomatic 24 year old female from Kigali, Rwanda, whose route of infection is thought to be due to heterosexual contact. The blood sample was taken in 1992. The clone has a frameshift mutation in gag in position 213. The original env and gag sequences clustered with HIV-1 subtype A sequences (Gao, F., *et al., ARHR* **10**:1359–1368 (1994)), however, subsequent in-depth analysis of the full length genome sequence from the isolate suggests it is an AC mosaic sequence with multiple crossover points. The full length sequence and the information concerning its mosaic nature was kindly made available prior to publication and the manuscript is currently submitted (Gao, F., *et al.*, 1997). The breakpoints are marked in Robertoson, D., *et al.*, part III, pages 25–30, of the 1997 compendium. The cytopathic effects of both the primary isolate of this patient, and from early passage cultured virus was NSI on MT-2 cells when analyzed by two different WHO labs in 1993-1994 (WHO Network, *ARHR* **10**:1327–1343 (1994), and De Wolf, F., *et al., ARHR* **10**:1387–1400, 1994. The NIAID 1997 Reference Reagent Catalog classifies it as NSI. However, more recent papers (Zhang, L., *et al., ARHR* **13**(16):1357–1366, 1997, classify it as SI. This sequence is from PCR-amplified proviral DNA harvested from PBMCs from the NIH NIAID reagent repository. The NSI phenotyping was determined using earlier passages, and SI was associated with later passages, (J. Bradac, personal communication). Both the full length clone 92RW009.6 and the partial gp 120 clone 92RW009.14 (accession U08793) were amplified from the same DNA extracted from a short term primary PBMC culture prepared by Ogden Bioservices, stored in the NIAID Research Reagent Repository, and sent to Drs. Gao and Hahn on 2/3/93. It was lot 1 (22 Jan 93) B. Hahn, personal communication. Other env sequences from this isolate are :U16221, U08793, U13441, U16220, and U16222, and a gag sequence is U86545. Second receptor usage of this isolate was defined by [Zhang,L., Huang,Y., He,T., Cao,Y. and Ho,D.D. HIV-1 subtype and second-receptor use. Nature 383(6603), 768 (1996) MEDLINE 97048157] and [Zhang,L., Carruthers, C.D., He,T., Huang,Y., Cao,Y., Wang,G., Hahn,B. and Ho,D.D. HIV type 1 subtypes, coreceptor usage, and CCR5 polymorphism. *ARHR* **13**(16):1357–1366 (1997) MEDLINE 98022612]. The CYTOPATHIC EFFECT (SI/NSI) phenotype of this isolate was not determined from the same passage as this sequence, but from 1-2 passages later of this same virus. The full length sequence and the information concerning its mosaic nature was kindly made available prior to publication, and the manuscript is currently submitted (F. Gao *et al.*, (1997)). Biotypes were determined by MT-2 syncytium assay; however, both syncytium-inducing (SI) and non-syncytium-inducing (NSI) variants may be present in the viral "swarm" for each isolate. Recent studies indicate that NSI isolates contain predominantly CCR5-using variants while most SI isolates contain both CXCR4 (SI) and CCR5 (NSI) variants. Some SI isolates may contain dual-tropic variants that use both CXCR4 and CCR5 co-receptors. The isolate 92RW009 is available from the NIH AIDS Reagent program, and is SI R5X4.

AC_SE.SE9488 AF071474 SWEDEN Carr, JK Unpublished

Although the features source line in the original GenBank entry specified that this was a recombinant between subtypes A and D, it is actually a recombinant between subtypes A and C. The GenBank record has also been corrected. Small sections of the 5′ and 3′ LTRs are not included in this sequence.

**Table 2 (cont.)**

| Sequence | Accession | Origin | Author | Reference |
|---|---|---|---|---|
| AC_ZM.ZAM184 | U86780 | ZAMBIA | Alaeus, A | *Virology* **213**(1):80–86 (1995) |

The envelope gene from this sample, from a Zambian woman taken in 1990 (J. Louwagie *et al., J. Virol.* **69**:263–271 1995), proved to be an A/C recombinant in subsequent phylogenetic analyses (D. Robertson *et al., Nature* **374**:124–126 (1995)). The full length provirus was recovered using PCR, and sequenced, according to the method described in M. O. Salminen *et al., Virology* **213**:80–86 (1995). Additional clones from serial samples from the index case ZAM184, and from her spouse who was also HIV-1 positive (GenBank Accession # U86768–U86781, inclusive) represent gag, env, or additional full sequences from this couple. The A/C mosaic pattern of the full length genomic sequence is presented in M. O. Salminen *et al., J. Virol.* **71**:2647–2655 (1997). See also entry with accession number L22955. , and also in Robertson, D., *et al.*, part III pages 25–30, of the 1997 compendium. All coding sequences are intact.

| | | | | |
|---|---|---|---|---|
| ACD_SE.SE8603 | AF075702 | SWEDEN | Carr, JK | *AIDS* **13**(14):1819–1826 (1999) |

Small sections of the 5′ and 3′ LTRs are not included in this sequence.

| | | | | |
|---|---|---|---|---|
| AD_SE.SE7108 | AF071473 | SWEDEN | Carr,JK | *AIDS* **13**(14):1819–1826 (1999) |

Small sections of the 5′ and 3′ LTRs are not included in this sequence.

| | | | | |
|---|---|---|---|---|
| ADH_NO.NOGIL3 | AJ237565 | NO | Jonassen, TO | *ARHR* **16**(1):49-58 (2000) |

Unpublished 38685 Three Norwegians, a couple and their daughter, died from AIDS in 1976 after up to 10 years of clinical manifestations of HIV infection (Lindboe *et al., Acta Pathol. Microbiol, Immunol. Scand.* **94**:117–123 (1986); Froland *et al., Lancet* 1344–1345 (1988)). In phylogenetic analysis, the obtained sequences of the HIV pol and vif genes clustered with the HIV-1 group O clade. The genotyping was confirmed by detection of antibodies against HIV-1 group O in blood samples from the father and the mother.

| | | | | |
|---|---|---|---|---|
| ADU_CD.MAL | X04415 | (FORMER ZAIRE) Dem Rep of the Congo | Gao, F | *J Virol* **72**(12):10234–10241 1998 |

This sequence is from a lambda phage clone derived from the Zairean isolate MAL. MAL was recovered in 1985 from a 7 year old boy with ARC, probably infected by a blood transfusion in 1981, as his parents were seronegative. All reading frames are intact except for vpu, due to the loss of a start codon. A reconstructed infectious clone is available. MAL was one of the first African sequences characterized and soon after the initial characterization it was determined to be a mosaic (Li, W. H, *et al., Genetics* **116**:s44 (1987)). Recent analysis suggests that it has three different distinct subtype associations, A, D, and I, and some regions that are difficult to characterize. (see Robertson, D., *et al.*, part III pages 25–30, of this compendium, 1997). The I association is based on phylogenetic analysis with a clone from Cyprus that was the first I subtype characterized (94CY032), this clone 94CY032, itself, appears to be A-G-I mosaic (see Robertson, D., *et al.*, part III pages 25–30, of this compendium, 1997). A region similar to the vpu cds of HIVELI exists from positions 5636 (starts with ‘ata’ instead of ‘atg’) to 5881.

**Table 2 (cont.)**

| Sequence | Accession | Origin | Author | Reference |
|---|---|---|---|---|
| AG_NG.92NG003 | U88825 | NIGERIA | Abimiku, AG | *ARHR* **10**(11):1581–1583 (1994) |

This sequence is from a PCR clone from an NSI primary culture from isolate G3, renamed 92NG003 to be consistent with WHO nomenclature. The sample was taken in 1992 from a 27 year old, asymptomatic HIV seropositive female prostitute from Jos, Nigeria. (Abimiku, A., *ARHR* **10**:1581–1583 (1994), env sequence accession number U13208). The isolate came from the Institute of Human Virology, Baltimore, MD, and is NSI. While originally described as subtype G in env, this genome has some A/G regions, with A-like regions in the central part of the genome that are similar to CRF_AG(IbNG), suggesting it may be a recombinant of a G subtype virus and an IbNG like virus (J. Carr, pers. comm.) A and G breakpoints are mapped in Robertson, D., et al., part III pages 25–30, of the 1997 compendium, and in Gao 1998. There are frameshift mutations associated with 10-16 bp deletions in vpr and vpu, at positions 5024 and 5485, as well as deletions totaling 33 bp near the 3′ end of the V3 loop. Nef has an altered initiation codon at position 8113. The sequence with accession number U13208 is from this same isolate. In the pol gene protease region, several other sequences cluster with this one in phylogenetic analysis (AF025737, AF075676 and AF075672), with at least 4 isolates obtained from Nigeria, suggesting 92NG003 may turn out to be a representative of a circulating recombinant form.

| | | | | |
|---|---|---|---|---|
| AGU_CD.Z321 | U76035 | (FORMER ZAIRE) Dem Rep of the Congo | Choi, DJ | *ARHR* **13**(4):357–361 (1997) |

Z321B is a later passage of isolate Z321 (see GenBank Accession Number M15896); Z321 was isolated from a 1976 Zairian serum sample; Z321 was grown to industrial scale in a chronically infected T-cell line to manufacture an inactive, therapeutic HIV-1 immunogen; Z321B was established from this industrial scale stock; as with Z321, Z321B contains a mutation in the termination codon of the tat gene, (bases 2294-2296) so that the tat gene of HIVZ321 extends further downstream (bases 2342-2344), and has a defective vpr and vpu. Recent analysis suggests that it has three different distinct subtype associations, A, G and I, and some regions that are difficult to characterize as part of any known subtypes (see Robertson, D., *et al.*, part III pages 25–30, of this compendium, 1997). The I association is based on phylogenetic associations with a clone from a Cyprus isolate that was the first I subtype characterized (95CY032). The 95CY032 clone also appears to be an A-G-I mosaic (Robertson, D., *et al.*, part III pages 25–30, of this compendium, 1997) but with different recombination breakpoints.

| | | | | |
|---|---|---|---|---|
| AGJ(BFP90)_AU.BFP90 | AF064699 | AUSTRALIA | Oelrichs, RB | *ARHR* **14**(16):1495–1500 (1998) |

HIV-1 from Burkina Faso, identified in Australia. The patient is a 32 year old African male who acquired the infection heterosexually in 1991. The patient was diagnosed in August 1996 at which time he had a CD4 count of 125. This sequence was derived by PCR directly from patient PBMCs when therapy-naive. The genome of this virus is a mosaic between subtypes A, G and J. The LTR is subtype J, a segment of about 950 bp at the beginning of gag is subtype A, the remainder of gag and part of the protease sequence are subtype G. Most of the pol gene cannot be assigned to a subtype. The mid-genome accessory region is mostly subtype J. gp120 is subtype G. The 3′ region of gp41, the third exons of tat and rev, and the nef gene are subtype J. The subtype G protease region clusters tightly in phylogenetic analysis, with three sequences from the Ivory Coast (accession numbers AF000482, AF000491 and AF000492), indicating that this is possibly a circulating recombinant form with 4 isolates from 2 different countries obtained to date (Aug 1999).

**Table 2 (cont.)**

| Sequence | Accession | Origin | Author | Reference |
|---|---|---|---|---|
| AGJ(BFP90)_ML.95ML84 | AJ245481 | MALI | Montavon, C | Unpublished |

This sequence is AG?J recombinant with an identical mosaic pattern to AG?J(BFP90) with accession number AF064699. The subtype G protease region clusters tightly in phylogenetic analysis, with three sequences from the Ivory Coast (accession numbers AF000482, AF000491 and AF000492), indicating that this is possibly a circulating recombinant form with 4 isolates from 2 different countries obtained to date (Aug 1999).

| | | | | |
|---|---|---|---|---|
| AJU_BW.98-2117 | AF192135 | BOTSWANA | Novitsky, VA | Unpublished (1999) |
| BF_BR.93BR029.4 | AF005495 | BRAZIL | Gao,F | *J Virol* **72**(7):5680–98 (1998) |

The isolate 93BR029 is part of a set of isolates obtained through the WHO Global Programme on AIDS (WHO Network, *ARHR* **10**:1327–1344 (1994)), and came from an asymptomatic HIV seropositive, 17 year old male, with unknown risk factor, from Sao Paulo, Brazil, sampled in 1993. The isolate had an NSI phenotype by an MT-2 assay. The isolate 93BR029 was established and propagated by short term cocultivation with normal donor lymphocytes, and then the near full length genome was PCR amplified and sequenced. An envelope gene sequence from this isolate is described in F. Gao *et al., J. Virol.* **70**:1651–1657 (1996). The envelope gene was first described as an F subtype, however subsequent phylogenetic analysis of the full length genome indicated the full length genome is a mosaic with regions of B and F subtype. The breakpoints are mapped in Robertson, D., *et al.*, part III pages 25–30, of the 1997 compendium. There are two frame shift mutations in gag, in positions 269 and 472. Biotypes were determined by MT-2 syncytium assay; however, both syncytium-inducing (SI) and non-syncytium-inducing (NSI) variants may be present in the viral "swarm" for each isolate. Recent studies indicate that NSI isolates contain predominantly CCR5-using variants while most SI isolates contain both CXCR4 (SI) and CCR5 (NSI) variants. Some SI isolates may contain dual-tropic variants that use both CXCR4 and CCR5 co-receptors. The isolate 93BR029 is available from the NIH AIDS Reagent program, and is NSI R5.

| | | | | |
|---|---|---|---|---|
| MO_CM.97CAMP645MO | AJ239083 | CAMEROON | Peeters, M | *J Virol* **73**(9):7368–7375 (1999) |

Sequence from a 29 year old Cameroonean woman, risk factors heterosexual sex and blood transfusion. The sample was taken in March 1997; at the time the woman was asymptomatic. The patient appeared to be dually infected with a group M (subtype AG_IBNG) and a recombinant M/O virus:envelope sequences from both groups could be found, but only group M gag and pol sequences.

| | | | | |
|---|---|---|---|---|
| O_CM.ANT70 | L20587 | CAMEROON | Vanden Haesevelde, M | *J Virol* **68**(3):1586–1596 (1994) |

ANT70 was isolated from the first O group infection discovered, and the very divergent LTR sequence was published in 1990 (de Leys, R., *et al., J Virol* **64**:1207–1216 (1990)). The isolate came from CDC stage II infected 19 year old female with unusual serological reactivity, who progressed to CDC stage III before publication of [1]. Both a man and his wife were infected with an O group HIV-1. The wife seroconverted in March, 1987. Both the husband and wife were originally from Cameroon, and living in Belgium. The husband was CDC stage III at the time virus was isolated from the wife for sequencing. The wife had a CD4:CD8 ratio of 0.25. Supernatant from the original coculture of wife PBMCs plus PHA-stimulated donor PBMCs, was used to infect MOLT4 clone 8 cells and MT-4 cells. Syncytia were formed in both these cell lines. After several weeks culture, chronically infected cell lines were obtained that shed virus, and supernatant from these stable lines were used for viral RNA isolation. HIV-1 O group viruses have the same genetic organization as HIV-1 M group viruses. For a review see Korber, B., *et al.*, Human Retroviruses and AIDS Database, Part III, 41–56 1996.

**Table 2 (cont.)**

| Sequence | Accession | Origin | Author | Reference |
|---|---|---|---|---|
| O_CM.MVP5180 | L20571 | CAMEROON | Gurtler, LG | *J Virol* **68**(3):1581–1585 (1994) |

This isolate was derived from a Cameroonian woman, sampled in 1991, who died of AIDS in 1992. The viral isolate MVP-5180 was grown in several human T-cell lines and the monocytic U937 line. The isolate MVP5180 is available from the NIH AIDS Reagent program, and is SI R5X4. A summary of isolates with known co-receptor usage can be found in the HIV database reviews.

N_CM.YBF30 AJ006022 CAMEROON Simon, F *Nature Med* **4**(9):1032–1037 (1998)

YBF30 was also tested on cell lines expressing CCR2b and CCR3 and did not utilze these coreceptors. YBF30 grew on chimpanzee PBMCs, but did not replicate in T-cell lines (MT-2 and HUT78). YBF30 infection of cells was entirely blocked by RANTES alone or in combination with MIP-1alpha and MIP-1beta. YBF30 was isolated from a 40 year old woman who had never travelled outside Cameroon. She presented with Histoplama capulatum infection of the colon in May, 1995 and died of AIDS (cachexia, neurological involvement and suspected disseminated histoplasmosis) in December, 1995. YBF30 was isolated from a May, 1995 blood sample and YBF31 from a December 1995 blood sample. YBF30 and YBF31 are greater than 98% identical to each other and less than 85% identical to HIV-1 M group, HIV-1 O group and SIV-CPZ sequences. The authors propose labelling this and similar viruses as N (between M and O, and also non-O non-M) group HIV-1. The N group designation is based upon sero-epidemilogical surveys in Cameroon, and another sequence (YBF105 not yet submitted to the databases) which indicate that more than one patient is infected with this clade of HIV-1. 700 stored sera collected between 1988 and 1997 were serologically tested, with a peptide-based EIA. 611 (87%) were reactive with M-group. 65 (9%) were reactive with O group. 8 were indeterminate. 16 (2%) were reactive with SIV-CPZ and not M or O group, 3 of these were strongly reactive with YBF30 peptides. A partial Pol gene was sequenced for one (YBF105).

SIVCPZUS AF103818 United States Gao, F *Nature* **397**(6718):436–441 (1999)

This full length molecular clone of a simian immunodeficiency virus, which infected an African wild-caught chimpanzee (Marilyn) who was the only chimpanzee identified as virus infected during a serosurvey of 98 chimpanzees in 1985. Marylin had never been used in AIDS research and had not recieved human blood products after 1969. She died in captivity in 1985 after giving birth to stillborn twins. The complete genome was sequenced from 4 overlapped PCR fragments, amplifed in 1998 from spleen tissues frozen at autopsy in 1985. Recovery of infectious virus from the frozen tissue was attempted but unsuccessful.

SIVCPZGAB X52154 GABON Huet, T *Nature* **345**(6273):356–359 (1990)

The CPZ genome is more closely related to HIV-1s than to any other HIV or SIV viral sequences, but it is more divergent from prototypical HIV-1 than any other isolate, with possible exception of the partially characterized ANT70. CPZ is especially different with respect to the vpU gene product. Also see CPZGAB2, U11495 for a sequence fragment from an additional chimpanzee caught in Gabon

**Table 2 (cont.)**

| Sequence | Accession | Origin | Author | Reference |
|---|---|---|---|---|
| SIVCPZANT | U42720 | (FORMER ZAIRE) Dem Rep of the Congo | Vanden Haesevelde, MM | *Virology* **221**(2):346–350 (1996) |

CPZANT is a simian immunodeficiency virus, phylogenetically linked to HIV-1. It was isolated from a captured wild chimpanzee from Zaire. This is the third SIV strain linked to HIV-1, after SIVCPZ-GAB (X52154) and SIVCPZ-GAB2 (U11495) were isolated from chimps in Gabon. Another chimpanzee virus was sequenced in 1998, and published in 1999 with accession number AF103818 (CPZ-US). The chimpanzee viral sequences are genetically more closely related to the HIV-1 sequences derived from infected humans than are HIV-2 strains or other SIVs. SIVCPZ-ANT is considered to be an outgroup of HIV-1 and is used to suggest the possibility of various introductions of HIV-1 into the human population.