JOURNAL OF VIROLOGY, Nov. 1989, p. 4982–4985
0022-538X/89/114982-04$02.00/0

Vol. 63, No. 11

Ancestry of a Human Endogenous Retrovirus Family

RENATO MARIANI-COSTANTINI,† TOBY M. HORN, AND ROBERT CALLAHAN*

Section on Oncogenetics, Laboratory of Tumor Immunology and Biology, National Cancer Institute, Bethesda, Maryland 20892

Received 3 March 1989/Accepted 13 July 1989

The human endogenous retrovirus type II (HERVII) family of HERV genomes has been found by Southern blot analysis to be characteristic of humans, apes, and Old World monkeys. New World monkeys and prosimians lack HERVII proviral genomes. Cellular DNAs of humans, common chimpanzees, gorillas, and orangutans, but not lesser ape lar gibbons, appear to contain the HERVII-related HLM-2 proviral genome integrated at the same site (HLM-2 maps to human chromosome 1). This suggests that the ancestral HERVII retrovirus(es) entered the genomes of Old World anthropoids by infection after the divergence of New World monkeys (platyrrhines) but before the evolutionary radiation of large hominoids.

Retroviral infection of various mammalian and avian species has resulted in the introduction of proviral genomes into the germ line of the host (23). These endogenous proviral genomes are transmitted to subsequent generations of the host as stable Mendelian genetic elements. Many of these proviral genomes are noninfectious but may be expressed in a tissue-specific manner. Human endogenous retrovirus (HERV) genomes can be divided into two diverse families on the basis of their relative homologies to either mammalian type C (HERVI) or mammalian types A, B, and D and avian type C (HERVII) retrovirus genomes (4–6, 16–18, 22). The HERVII genomes, although related, are highly divergent on the basis of restriction enzymes and heteroduplex analysis (T. M. Horn et al., manuscript in preparation). This family is composed of approximately 50 copies of sequences related to viral structural genes and 1,000 copies of long terminal repeat-like sequences. HERVII genomes are found on many if not all human chromosomes (13). Two members of the HERVII family, *HLM-2* and *HLM-25*, are located on human chromosomes 1 and 5, respectively (13). Southern blot analysis of *Eco*RI-digested cellular DNA from 30 unrelated humans gave identical hybridization patterns when probed with restriction fragments corresponding to different regions of the HLM-2 proviral genome (unpublished data). This suggests that the HERVII proviral genomes originated early in the human lineage. To trace the ancestry of HERVII genomes, we examined cellular DNAs of primate species, including the large hominoids, the lesser ape lar gibbon, several Old and New World monkeys, and various prosimians (Table 1), for HERVII-related sequences.

Human cellular DNA digested with *Eco*RI and probed with HLM-2 *pol* (Fig. 1) shows a complex hybridization pattern at low stringency (Fig. 2A, lanes 1 and 10) that resolves into a small number of cross-hybridizing species after the high-stringency wash (Fig. 2B, lanes 1 and 10). The HLM-2 provirus contains a characteristic 3.7-kilobase-pair (kbp) *Eco*RI *pol* fragment that is detected in the cellular DNA of mouse NIH 3T3 cells transfected with λHLM-2 DNA (Fig. 2, lanes 2) and in human cellular DNA. Three additional fragments (2.8, 2.4, and 1.9 kbp; Fig. 2B, lanes 1 and 10) in human cellular DNA having strong sequence homology with HLM-2 *pol* have been tentatively assigned to other HERVII-related proviral genomes (Horn et al., in preparation). Genomic DNA of the common chimpanzee (central Africa) lacked the 2.4-kbp fragment, but its overall pattern of hybridization to the HLM-2 *pol* probe at both low and high stringencies was otherwise remarkably similar to

TABLE 1. Primate species and number of individuals examined for HERVII endogenous proviral genomes[a]

Group and species	Common name	No. studied
Hominoids		
Homo sapiens	Human	30
Pan troglodytes	Common chimpanzee	6
Gorilla gorilla	Gorilla	5
Pongo pygmaeus	Orangutan	4
Hylobates lar	Lar gibbon	2
Old World monkeys		
Colobus guereza	Guereza	1
Presbytis obscura	Dusky leaf monkey	1
Erythrocebus patas	Patas monkey	1
Cercopithecus aethiops	Vervet monkey	2
Macaca mulatta	Rhesus monkey	6
Macaca fascicularis	Long-tailed macaque	3
Macaca nigra	Celebes macaque	1
Macaca arctoides	Stump-tailed macaque	1
Papio cynocephalus	Yellow baboon	2
Papio sphinx	Mandrill	1
New World monkeys		
Ateles paniscus	Spider monkey	1
Aotus trivirgatus	Owl monkey	4
Saimiri sciureus	Squirrel monkey	6
Cebuella pygmaea	Pygmy marmoset	1
Saguinus (*nigricollis* × *fuscicollis*)	Hybrid tamarin	1
Prosimians		
Galago crassicaudatus	Greater bushbaby	6
Galago senegalensis	Lesser bushbaby	2
Cheirogaleus medius	Fat-tailed dwarf lemur	1
Lemur coronatus	Crowned lemur	1
Lemur fulvus	Brown lemur	1

[a] Nomenclature is as described by Wolfheim (25) and Hershkovitz (12). Cellular DNAs from tissues and cells from different individuals of the primate species listed were isolated by standard procedures (15); digested with restriction endonuclease(s) *Eco*RI, *Eco*RI-*Xho*I, or *Bam*HI; electrophoretically separated on 0.75% agarose gels; and processed for Southern blotting as described in the legend to Fig. 2.

* Corresponding author.

† Present address: Institute of Pathology, Section on Experimental Oncology, University G. D'Annunzio, Chieti, Italy.

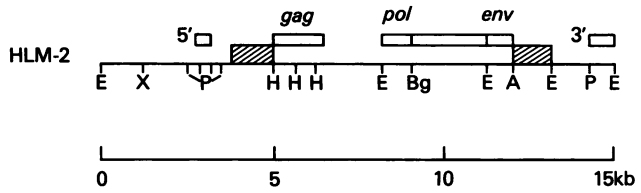

FIG. 1. λHLM-2 proviral genome. Hatched areas indicate the long terminal repeats. We show only the restriction sites used to generate the probes for this study. Abbreviations: A, *Ava*II; Bg, *Bgl*II; E, *Eco*RI; H, *Hin*dIII; P, *Pst*I; X, *Xho*I. HLM-2 was isolated from a human genomic library generated from a *Hae*III partial digest packaged in λ by using *Eco*RI linkers. The 5'-flanking sequences *gag*, *pol-env*, and *env* and the 3'-flanking sequences of HLM-2 were subcloned by standard techniques in pSP64 or pBR322 vectors.

that of human DNA (Fig. 2, lanes 3). Gorilla (central Africa) and orangutan (insular Southeast Asia) exhibited all the characteristic human internal HLM-2 *pol*-related fragments when relaxed (Fig. 2A, lanes 4 and 5) and stringent (Fig. 2B,

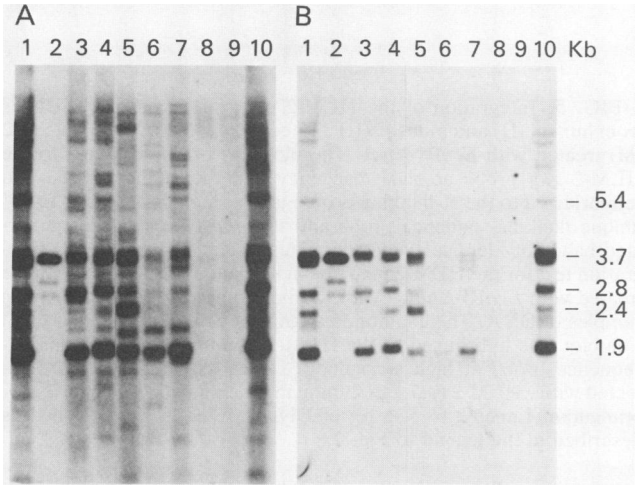

FIG. 2. Southern blot analyses of *Eco*RI digests of primate DNAs and HLM-2 *pol*. Replicate blots were hybridized at low stringency to HLM-2 *pol*, washed to low stringency, and exposed for 20 h (A); they were then washed to high stringency and reexposed for 70 h (B). Each lane contained 5 μg of DNA except for lane 2, which had 0.5 μg of DNA. Lanes: 1 and 10, Human lymphocyte; 2, mouse NIH 3T3 cells transfected with λHLM-2; 3, chimpanzee; 4, gorilla; 5, orangutan; 6, mandrill; 7, Celebes macaque; 8, owl monkey; 9, squirrel monkey. Inserts purified from agarose gels were nick translated (Bethesda Research Laboratories, Inc.) with [^{32}P]dCTP (Amersham Corp.) to use as hybridization probes. The gels were processed for Southern blotting onto Genatrans (Plasco) by using 20× SSC (1× SSC is 0.15 M NaCl and 0.015 M sodium citrate) as the transfer medium. Blots were prehybridized at 55°C in Genatrans hybridization solution (5× Denhardt solution [15], 1% sodium dodecyl sulfate, 3× SSPE [20× SSPE is 3 M NaCl–0.2 M NaH_2PO_4, pH 6.5], 2.5% dextran sulfate, 0.001 M sodium phosphate, pH 6.7) containing 50% deionized formamide. Filters were hybridized with 2.5×10^6 cpm of denatured ^{32}P-labeled probe per ml under relaxed conditions (20% formamide in Genatrans hybridization solution) at 37°C for 36 h. After being rinsed in several changes of 3× SSC–0.5% sodium dodecyl sulfate at room temperature, the filters were subjected to relaxed washes (0.5× SSC–0.1% sodium dodecyl sulfate, 55°C) and stringent washes (0.1× SSC–0.1% sodium dodecyl sulfate, 65°C). The filters were exposed to Kodak XAR film at −70°C by using Cronex intensifier screens.

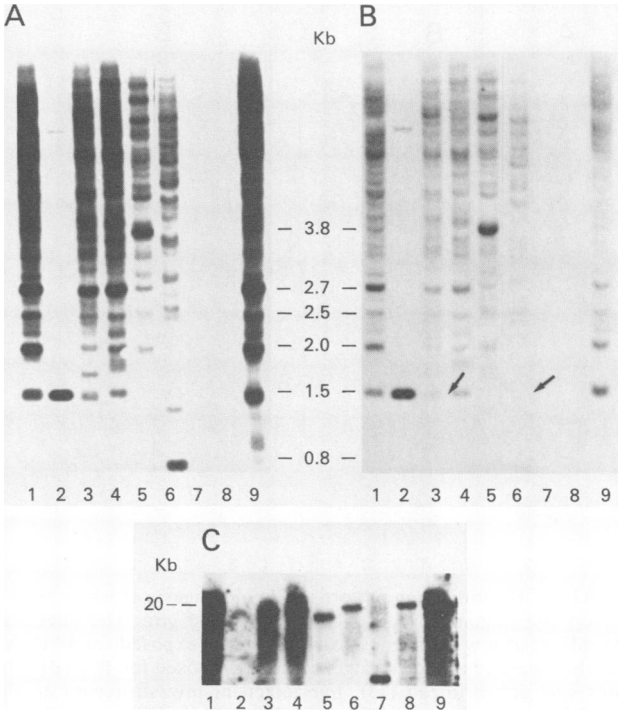

FIG. 3. Hybridization patterns of *Eco*RI digests of representative primate DNAs to HLM-2 *env*. (A) Relaxed wash, 48-h exposure; (B) stringent wash, 48-h exposure. Each lane contained 5 μg of DNA except for lane 2, which had 0.5 μg of DNA. Lanes: 1 and 9, Human lymphocyte; 2, NIH 3T3 mouse cell line transfected with λHLM-2; 3, chimpanzee; 4, gorilla; 5, orangutan; 6, Celebes macaque; 7, owl monkey; 8, squirrel monkey. (C) Control hybridization of the same blot to a human c-*sis* proto-oncogene probe. Methods were applied as described in the legend to Fig. 2.

lanes 4 and 5) wash conditions were used. The differences in intensity of individual bands might reflect differences in copy number or the degree of sequence divergence. DNA samples from the Asian lesser ape lar gibbon (data not shown) and from several species of Old World monkeys (e.g., the mandrill from central Africa and the Celebes macaque (Fig. 2, lanes 6 and 7) (Table 1) also hybridized strongly at low stringency with HLM-2 *pol* but yielded patterns after a high-stringency wash that were distinct from human DNA. No HLM-2 *pol*-related sequences were detected in the DNAs of New World monkeys, even under relaxed hybridization conditions (Fig. 2A, lanes 8 and 9).

In New World monkey or prosimian DNAs, the HLM-2 *env* probe also did not detect any cross-hybridizing fragments, even under relaxed conditions (Fig. 3A, lanes 7 and 8; other data not shown). Common chimpanzee and gorilla DNAs, however, showed many fragments corresponding in size to the human *env* fragments, even at high stringency. These included the 1.5-kbp *Eco*RI fragment characteristic of the HLM-2 provirus as well as other strongly hybridizing fragments of molecular sizes 2.0, 2.5, and 2.7 kbp (Fig. 3B; compare lanes 1, 3, and 4). While *Eco*RI digests of orangutan DNA did not contain the 1.5-kbp fragment typical of HLM-2 *env*, these DNAs retained the 2.0-, 2.5-, and 2.7-kbp fragments and, in addition, had a 3.8-kbp fragment that gave a strong hybridization signal (Fig. 3, lane 5). Celebes macaque DNA showed a weaker hybridization signal to HLM-2 *env* after high-stringency washing (Fig. 3B, lane 6, [arrow]),

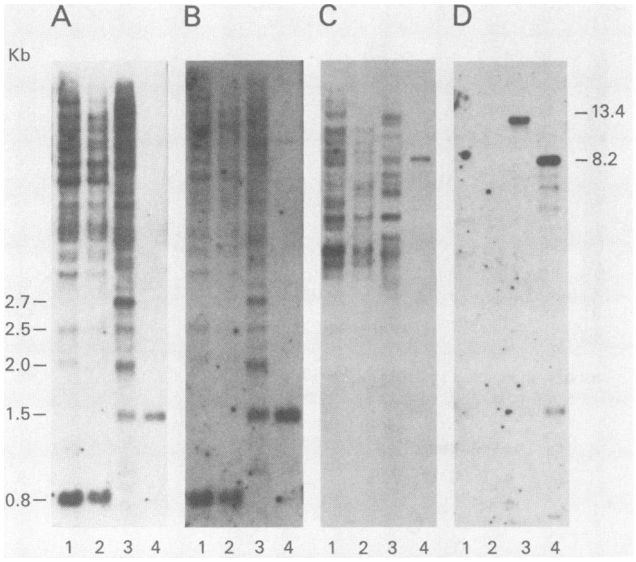

FIG. 4. Hybridization patterns of *Eco*RI digests of lesser ape lar gibbon and human DNAs to HLM-2 *env* and *gag*. The filter was hybridized at low stringency to HLM-2 *env*, exposed for 48 h (A), and then washed at high stringency and reexposed for 48 h (B). The filter was dehybridized (15), hybridized at low stringency to the HLM-2 *gag* probe, and exposed for 20 h (C); then it was washed at high stringency and reexposed for 6 days (D). Each lane contained 5 μg of DNA except for lane 4, which had 0.5 μg of DNA. Lanes: 1 and 2, Lar gibbon lymphocyte (samples from two individuals); 3, human lymphocyte; 4, NIH 3T3 mouse cell line transfected with λHLM-2. Methods were applied as described in the legend to Fig. 2.

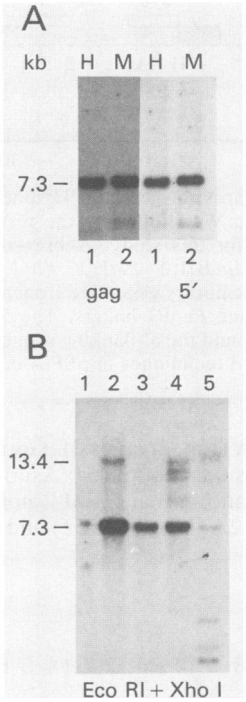

FIG. 5. Integration of the HLM-2 proviral genome. (A) DNAs from human (H) and mouse NIH 3T3 cells transfected with λHLM-2 (M) treated with *Eco*RI-*Xho*I. The blot was first hybridized to the HLM-2 *gag* probe at high stringency and then dehybridized and rehybridized to the 5′-flanking probe, also at high stringency. The 5′ unique flanking-sequence probe and the *gag* region probes generated bands of identical molecular sizes (4-day exposure). (B) Integration test for the HLM-2 provirus in hominoid DNAs. DNAs were treated with *Eco*RI-*Xho*I. (*Xho*I did not completely digest the 5-μg samples of DNA. This facilitated alignment of the autoradiograms.) The blot was hybridized to the HLM-2 5′ unique cellular flanking-sequence probe at high stringency. Lanes: 1, Mouse cells transfected with λHLM-2 (0.5 μg); 2, human; 3, chimpanzee; 4, gorilla; 5, orangutan. Lanes 2 to 5, 5 μg of DNA. Methods were applied as described in the legend to Fig. 2.

although at low stringency the major cross-hybridizing fragment was 0.8 kbp (Fig. 3A, lane 6).

Figure 4A and B depict the HLM-2 *env* hybridization patterns of the lesser ape lar gibbon (lanes 1 and 2), which is sympatric with the orangutan. Like orangutan DNA, the gibbon DNA samples have cross-hybridizing *Eco*RI fragments corresponding at 2.0 and 2.5 kbp and lack the typical HLM-2 1.5-kbp fragment (compare lanes 1 and 2 with lane 3). However, lar gibbon DNA contains an additional 0.8-kbp fragment that shows a strong signal. Gibbon DNA samples showed homology to the HLM-2 *gag* probe but only at low stringency (Fig. 4C versus D).

Unique cellular flanking-sequence probes (Fig. 1) have been used to locate individual proviral genomes to human chromosomes 1 and 5 and to detect unique integration sites (13). A single 7.3-kbp fragment generated by double digestion with *Eco*RI-*Xho*I hybridized to both the HLM-2 *gag* region and the 5′ unique cellular flanking-sequence probes (Fig. 1 and 5A). As in the transfectant and the human (Fig. 5A and B, lanes 1 and 2, respectively), common chimpanzee and gorilla DNAs showed the single 7.3-kbp *Eco*RI-*Xho*I fragment when the HLM-2 5′-flanking probe was used (Fig. 5B, lanes 3 and 4). This suggests that the genomic DNAs of the two African apes contain the HLM-2 proviral genome at the same site as does the human genomic DNA. Although *Eco*RI-*Xho*I digests of orangutan DNAs resolved a 7.3-kbp fragment, additional low-molecular-weight bands appeared (Fig. 5B, lane 5). In contrast to those of the three large apes, DNAs from the lesser ape lar gibbon did not cross-hybridize to the HLM-2 *gag* probe (Fig. 4D) or to the HLM-2 5′ unique cellular flanking-sequence probe (data not shown) at high stringency. This is consistent with the lack of characteristic HLM-2 proviral internal fragments, e.g., *env* (Fig. 4A and B) and *pol* (data not shown), and suggests that, although lar gibbon DNA has HERVII-related sequences, it does not contain the HLM-2 proviral genome.

The presence of the HLM-2 provirus in large apes and humans corresponds to their phylogenetic relationships instead of their current geographic distribution (8, 11, 12, 14, 20, 21, 25). The observed hybridization patterns suggest either that the human HLM-2 proviral genome was inherited from a common ancestor of all the extant large hominoids after the divergence of the lesser ape lineage (11) or that this region of the lar gibbon genome was lost subsequent to the divergence of lar gibbons from large hominoids. We favor the former possibility, since the restriction patterns of HLM-2 proviral *env*, *gag*, and flanking sequences are consistent with the evolutionary branching order of orangutans, African apes, and humans (11, 21). This issue should be resolved with studies using additional cellular DNA probes flanking the HLM-2 provirus or other HERVII proviral genomes. In this regard, another human endogenous proviral genome, ERV-1, which is unrelated to the HERVII family, is known to have entered the human germ line prior to the chimpanzee-human branch point (3).

The possibility that the HERVII family of endogenous retrovirus sequences entered the catarrhine lineage via an

ancient infection is suggested by the following observations: the presence of all the HERVII internal sequences in all the species of Old World simian primates tested, the similarity of HLM-2 proviral patterns in humans and large apes, and the absence of HERVII structural sequences in New World monkeys. A HERVII infection of early catarrhines would therefore have occurred after the geographic separation between the Old and the New World anthropoid lineages, i.e., approximately 40 million years ago (7, 9–11) but before the evolutionary radiation and geographic dispersal (outside Africa) of the ancestors of the extent large hominoids (estimated at 17 million years ago) (1, 2, 19, 24).

For their cooperation in helping us acquire primate tissues and cells, we thank H. M. McClure and C. Allen, Yerkes Regional Primate Center, Atlanta, Ga.; M. K. Izard, P. Wright, and M. Pereira, Duke University Center for the Study of Primate Biology and History, Durham, N.C.; M. W. Terry, Primate Information Center, Regional Primate Research Center, Seattle, Wash.; T. M. Butler, Southwest Foundation for Biomedical Research and Education, San Antonio, Tex.; A. Lock and E. Scherer, Division of Research Services, Veterinary Resources Branch, National Institutes of Health, Bethesda, Md.; S. A. Aaronson, Laboratory of Cellular and Molecular Biology, National Cancer Institute, Bethesda, Md.; and H. J. Thiel, Federal Research Centre for Virus Diseases of Animals, Tubingen, Federal Republic of Germany. R. Bassin, Laboratory of Tumor Immunology and Biology, National Cancer Institute, prepared the NIH 3T3 cell line transfectant containing the HLM-2 proviral genome. R.M.-C. gratefully acknowledges the expert assistance of E. Divizia, Department of Experimental Medicine, University "La Sapienza," Rome, Italy.

LITERATURE CITED

1. **Andrews, P.** 1981. Species diversity and diet in monkeys and apes during the Miocene, p. 25–61. *In* C. B. Stringer (ed.), Aspects of human evolution. Taylor and Francis Ltd., London.
2. **Andrews, P.** 1985. Family groups systematics and evolution among catarrhine primates, p. 14–22. *In* E. Delson (ed.), Ancestors: the hard evidence. Alan R. Liss, Inc., New York.
3. **Bonner, T. I., C. O'Connell, and M. Cohen.** 1982. Cloned endogenous retroviral sequences from human DNA. Proc. Natl. Acad. Sci. USA **79**:4709–4713.
4. **Callahan, R.** 1988. Two families of human endogenous retroviral genomes, p. 91–100. *In* Banbury report 30: eukaryotic transposable elements as mutagenic agents. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
5. **Callahan, R., I. M. Chiu, J. F. H. Wong, S. R. Tronick, B. A. Roe, S. A. Aaronson, and J. Schlom.** 1985. A new class of endogenous human retroviral genomes. Science **228**:1208–1211.
6. **Callahan, R., W. Drohan, S. Tronick, and J. Schlom.** 1982. Detection and cloning of human DNA sequences related to the mouse mammary tumor virus genome. Proc. Natl. Acad. Sci. USA **79**:5503–5507.
7. **Ciochon, R. L., and A. B. Chiarelli.** 1980. Paleobiogeographic perspectives on the origin of the platyrrhini, p. 459–493. *In* R. L. Ciochon and A. B. Chiarelli (ed.), Evolutionary biology of the New World monkeys and continental drift. Plenum Publishing Corp., New York.
8. **Ferris, S. D., A. C. Wilson, and W. M. Brown.** 1981. Evolutionary tree for apes and humans based on cleavage maps of mitochondrial DNA. Proc. Natl. Acad. Sci. USA **78**:2432–2436.
9. **Fleagle, J. G., T. M. Bown, J. D. Obradovich, and E. L. Simons.** 1986. Age of the earliest African anthropoids. Science **234**: 1247–1249.
10. **Gingerich, P. D.** 1980. Eocene adapidae, paleobiogeography and the origin of South American platyrrhini, p. 123–138. *In* R. L. Ciochon and A. B. Chiarelli (ed.), Evolutionary biology of the New World monkeys and continental drift. Plenum Publishing Corp., New York.
11. **Gingerich, P. D.** 1984. Primate evolution: evidence from the fossil record, comparative morphology and molecular biology. Yearb. Phys. Anthropol. **27**:57–72.
12. **Hershkovitz, P.** 1977. Living New World monkeys (platyrrhini) with an introduction to primates, vol. 1. The University of Chicago Press, Chicago.
13. **Horn, T. M., K. Huebner, K. C. Croce, and R. Callahan.** 1986. Chromosomal locations of members of a family of novel endogenous retroviral genomes. J. Virol. **58**:955–959.
14. **Koop, B. F., M. Goodman, P. Xu, K. Chan, and J. L. Slightom.** 1986. Primate β-globin DNA sequences and man's place among the great apes. Nature (London) **319**:234–238.
15. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
16. **May, F. E. B., and B. R. Westley.** 1986. Structure of a human retroviral sequence related to mouse mammary tumor virus. J. Virol. **60**:743–749.
17. **O'Connell, C., S. O'Brian, W. G. Nash, and M. Cohen.** 1984. ERV-3, a full length human endogenous provirus: chromosomal localization and evolutionary relationships. Virology **138**:225–235.
18. **Ono, M.** 1986. Molecular cloning and long terminal repeat sequences of human endogenous retrovirus genes related to types A and B retrovirus genes. J. Virol. **58**:937–944.
19. **Pilbeam, D.** 1985. Patterns of hominoid evolution, p. 51–59. *In* E. Delson (ed.), Ancestors: the hard evidence. Alan R. Liss, Inc., New York.
20. **Schwartz, J. H.** 1984. The evolutionary relationships of man and orangutans. Nature (London) **308**:501–505.
21. **Sibley, C. G., and J. E. Ahlquist.** 1984. The phylogeny of the hominoid primates as indicated by DNA-DNA hybridzation. J. Mol. Evol. **20**:2–15.
22. **Steel, P. E., M. A. Martin, A. B. Rabson, T. Bryan, and S. J. O'Brien.** 1986. Amplification and chromosomal dispersion of human endogenous retroviral sequences. J. Virol. **59**:545–550.
23. **Stoye, J., and J. Coffin.** 1985. Endogenous retroviruses, p. 357–404. *In* R. Weiss, N. Teich, H. Varmus, and J. Coffin (ed.), RNA tumor viruses. Molecular biology of tumor viruses, vol. 2, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
24. **Thomas, H.** 1985. The early and middle Miocene connection of the afro-arabian plate and Asia: a major event for hominoid dispersal? p. 42–50. *In* E. Delson (ed.), Ancestors: the hard evidence. Alan R. Liss, Inc., New York.
25. **Wolfheim, J. H.** 1983. Primates of the world. University of Washington Press, Seattle.