Chromosome Research 7: 177–183, 1999.

Localization of HTLV-1 and HIV-1 proviral sequences in chromosomes of persistently infected cells

Liubov A. Glukhova[1], Sergei V. Zoubak[2,3], Alla V. Rynditch[2], Galina G. Miller[4], Irina V. Titova[4], Nadezjda Vorobyeva[5], Zoya V. Lazurkevitch[2], Alexander S. Graphodatskii[5], Alla A. Kushch[1] & Giorgio Bernardi[3]*
[1]*Ivanovsky Institute of Virology, Russian Academy of Medical Sciences, 16 Gamaleya str., 123098 Moscow, Russia;* [2]*Institute of Molecular Biology and Genetics, Ukrainian Academy of Sciences, 150 Zabolotnogo str., 252627 Kiev, Ukraine;* [3]*Laboratoire de Génétique Moléculaire, Institut Jacques Monod, 2 Place Jussieu, 75005 Paris, France; Tel. (33-1) 4329-5824; Fax (33-1) 44277977; E-mail: bernardi@citi2.fr;* [4]*Gamaleya Institute of Epidemiology and Microbiology, Russian Academy of Medical Sciences, 18 Gamaleya str., 123098 Moscow, Russia;* [5]*Institute of Cytology and Genetics, Sibirian Branch of Russian Academy of Sciences, 10 Lavrentyeva prosp., 630090 Novosibirsk*
**Correspondence*

Received 27 July 1998; received in revised form and accepted for publication by B. Dutrillaux 22 December 1998

Key words: chromosomal integration, HIV-1, HTLV-1, isochores

Abstract

Integration sites for HTLV-1 and HIV-1proviruses were detected by FISH on the chromosomes of HTHIV27 cells persistently infected by HIV-1 (strain IIIB). HTLV-1 signals were found on 9 loci of chromosomes 4, 6, 9, 15 and 16. Integration sites of GC-rich HTLV-1 provirus are located in GC-rich isochores, confirming an 'isopycnic' integration, namely an integration in which the GC level of the host sequences around the integration site match the GC level of the provirus. This conclusion is not only derived from the compositional map of human chromosomes, but also from HTLV-1 hybridization on compositional fractions of human DNA. Integration of GC-poor HIV-1 provirus was found on 4 loci of chromosomes 2, 7, 17 and 19. One copy of a complete HIV-1 provirus, which is active, was integrated in H1 isochores, whereas other defective copies were located in GC-poor L isochores. These results are discussed in terms of regional integration of retroviral sequences.

Introduction

HIV-1 infection is characterized by a variable clinical course and the prognosis of HIV infection can be predicted from the amount of virus in plasma, a persistent high rate of virus production, measured as HIV-1 RNA, indicating a poorer prognosis (Mellors *et al.* 1996). The factors that influence virus production are not clear. One of them certainly is the efficient transcription of viral DNA (Coffin 1990, Sakai *et al.* 1993), which in turn is dependent on the sites of integration (Zerial 1986, Rynditch *et al.* 1991, Zoubak *et al.* 1994, Rynditch *et al.* 1998).

Integration sites can be studied by using a compositional approach to the localization of proviral sequences in the host genome (reviewed in Bernardi 1995). Such an investigation is possible for two reasons: (1) The genomes of warm-blooded vertebrates are mosaics of isochores, long DNA segments, which are homogeneous in base composition, but belong to different families characterized by different CG levels. In the human genome, two GC-poor isochore

families, L1 and L2, form 63% of the genome, and three GC-rich isochore families, H1, H2 and H3, make up 24%, 7.5% and 4–5% of the genome, respectively, the rest being represented by satellite and ribosomal DNA. Gene concentration is low in GC-poor isochores and increasingly higher in GC-rich isochores (Bernardi 1995; Zoubak *et al.* 1996).

(2) The large DNA fragments, 50–100 kb in size, which form routine DNA preparations and reflect the base composition of the isochores from which they are derived, can be fractionated by preparative equilibrium centrifugation in Cs_2SO_4 density gradients in the presence of sequence-specific DNA ligands. Hybridization with appropriate probes of the integrated viral sequences in compositional fractions of host cell DNA has shown that, in all cases investigated so far, integration of expressed viral sequences is 'compartmentalized' and 'isopycnic', i.e. takes place in isochore families that show a compositional match with the viral sequences. This compartmentalized integration was found for the GC-rich sequences from bovine leukaemia virus (BLV), hepatitis B virus (HBV), Rous sarcoma virus (RSV), and human T-cell leukaemia virus (HTLV-1), which all integrate in GC-rich regions of the host genome, and for the GC-poor sequences from mouse mammary tumour virus (MMTV), which integrate in GC-poor regions (Kettmann *et al.* 1979, Zerial *et al.* 1986, Salinas *et al.* 1987, Rynditch *et al.* 1991, Zoubak *et al.* 1994). Moreover, it has been shown for GC-rich viral sequences that transcribed copies appear to be located in compartments characterized by GC levels closer to those of proviruses compared to untranscribed sequences (Zerial *et al.* 1986, Rynditch *et al.* 1991, Zoubak *et al.* 1994).

On the other hand, *in-situ* hybridization of compositional DNA fractions corresponding to different isochore families provided a classification of chromosomal bands that is related to GC level and to gene distribution in chromosomes (Saccone *et al.* 1993, 1996). Therefore, *in-situ* localization of integrated viral sequences parallels their distribution in compositional DNA fractions.

In the present work, we assessed the localization of HIV-1 sequences on human chromosomes and compositional DNA fractions using the T-lymphoblastoid cell line HTHIV27 persistently infected with HIV1/IIIB. HTHIV27 cells were obtained from the MT4 line, which contains poorly expressed HTLV-1 proviral copies, and were selected as cells characterized by a high growth rate, active production and long-term stable expression of HIV-1 in 100% cells (Miller *et al.* 1992, Kushch *et al.* 1995).

Materials and methods

Cells

The MT4 line, obtained by cocultivating lymphocytes from the umbilical cord with the cells of a T-cell leukaemia patient (Miyoshi *et al.* 1981) was infected with HIV-1/IIB. The HTHIV-27 cell line was established by selection of these infected cells according to their resistance to the cytolytic activity of HIV-1 (Miller *et al* 1992, Kushch *et al.* 1995). The cells were grown in RPMI-1640 medium (Serva) with 10% FCS (Flow), glutamine and gentamycin, recultured every 4–5 days for 8 years.

Chromosome preparation and in-situ hybridization

pBH10-R3, containing a HIV-1 provirus lacking 190 bp in 5′LTR and leader sequences (Fisher *et al.* 1985) and pMT-2, containing HTLV-I provirus with a single LTR (Clarke *et al.* 1983) were used as probes for *in-situ* hybridization on the metaphase spreads. Probes obtained from Biotech Research Laboratories, Inc., Rockville, MD, were labelled by nick-translation according to established procedures using biotin-16-dUTP (Singer *et al.* 1989). Hybridization *in situ* was performed as described (Lawrence *et al.* 1990), with some modifications. For each slide, 80–130 ng of labelled probes were applied in 20 μl of hybridization buffer. Probes were detected using avidin conjugated to fluorescein isothiocyanate, FITC (Vector). The signal was amplified by two series of incubation with avidin conjugated to FITC and biotinylated goat anti-avidin antibodies (Vector). Chromosomes were counterstained with propidium iodide. The same methaphase spreads were then rinsed and stained by G-banding and rephotographed.

The precise location of the fluorescent signals on the chromosomes was limited by three factors: (1) differences in intensity of some spots in each metaphase spread; (2) different number of signals per metaphase; (3) difficulties in obtaining high resolution of G-banding after *in-situ* hybridization. For these reasons, the integrations of the two viral sequences are presented using pairs of chromosomes

after FISH and G-banding from different metaphases. In spite of this, we unequivocally visualized integration loci for HIV-1 and HTLV-1 proviruses analysing 103 metaphase spreads.

Localization of HIV-1/HTLV-1 sequences in compositional DNA fractions

High-molecular-weight DNA in 0.1 mol/L Na_2SO_4/1 mol/L borate pH 9.4 from HTHIV27 cells was fractionated by Cs_2SO_4/BAMD density gradient centrifugation (Cortadas *et al.* 1977, Salinas *et al.* 1987). The average sizes of DNA fragments were higher than 50 kb as estimated by gel electrophoresis. After dialysis to eliminate Cs_2SO_4 and BAMD, total DNA and aliquots of DNA from fractions were digested with XbaI, which does not cut inside HIV proviral sequences, and investigated by Southern blot hybridization with pMT-2 and pBH10-R3 under stringent conditions.

Results and discussion

Localization of HTLV-1 sequences

The localization of HTLV-1 sequences in compositional DNA fractions (Figure 1A) is characterized by the fact that different hybridization bands having the same size may be present in a different number of fractions and exhibit different intensities. The reasons for the broad and variable band distribution have been discussed elsewhere (Zoubak *et al.* 1994). In order to identify the maxima of viral sequence distribution, band intensities have been compared, the strongest band corresponding, on a first approximation, to such a maximum. It should be stressed, however, that since in the present work, equal amounts of DNA were loaded for each fraction, loads were proportionally larger for the fractions containing less DNA, namely, for the GC-rich fractions (above 39–40% GC which is the modal GC level of human DNA). Assessments of hybridization maxima should, therefore, take into account the overestimation of hybridization intensities in GC-rich fractions. Then, for example, among the two apparently equally strong bands of 19 kb, the left one (fraction 4) corresponds to the largest amount of proviral sequences.

Eleven bands of hybridization with HTLV-1 probe were detected in compositional DNA fractions in HTHIV27 cells. The number of hybridization bands may not, however, correspond to the number of viral copies because of the possible loss of XbaI site in defective proviruses (Saggioro *et al.* 1990).

HTLV-1 sequences were distributed in the 41–46%

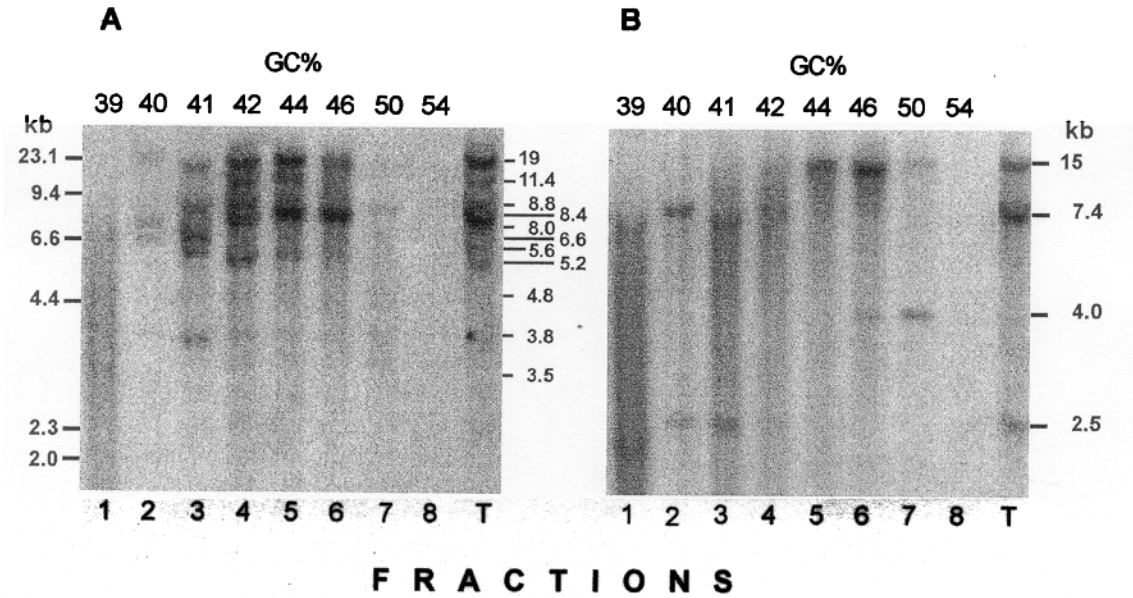

Figure 1. Localization of HTLV-1 (**A**) and HIV-1 (**B**) proviral copies in DNA compositional fractions from HTHIV-27 cell line. Total DNA (8 μg) and equal amounts (3 μg) of DNA from each fraction were digested with XbaI, electrophoresed on 0.8% agarose gels and hybridized under stringent conditions with the insert of pMT-2 or pBH10-R3.

GC range of the host genome. The peaks of distribution of the most intense bands (19 kb, 8 kb) and of the less intense bands (11.4 kb, 8.8 kb, 5.2 kb) were localized in 42–44% GC isochores (see above), where both transcriptionally active and transcriptionally inactive HTLV-1 sequences were previously found (Zoubak *et al.* 1994). Other bands, which could correspond to a small population of the cells that diverged from the main one because of viral or cellular DNA rearrangements, were located at 41% GC which is the localization of transcriptionally inactive HTLV-1 sequences in the human genome. This means that HTLV-1 sequences were integrated either in regions where transcriptionally inactive HTLV-1 sequences were previously found or in the borderline regions where both transcriptionally active and inactive proviruses are located. HTLV-1 sequences were not observed in the regions where only transcriptionally active proviruses integrate.

The analysis of metaphase chromsomes showed that HTLV-1 integration sites were distributed in 9 chromosomal loci (Figure 2A, B, C). Most freqeuntly, HTLV-1 sequences integrated into 6 chromosomal loci: 4p15(16), 16q22, 6p21, 9q21, 16p13 and 15q24(25) (Table 1). In all clear-cut cases, HTLV-1 integration sites were observed in R(everse) bands, except for one, localized in 9q21, a G(iemsa) band. R bands are non-uniform in GC levels and comprise three distinct subsets of bands at a 400-band resolution level: $H3^+$, $H3^*$ and $H3^-$, which harbour large amounts, moderate amounts and no H3 isochores, respectively (Saccone *et al.* 1996). $H3^+$ bands largely coincide with the T(elomere) bands of Dutrillaux (1973). The integration sites in the main cell population were located in $H3^*$ bands. These bands may comprise isochores from all families. 9q21 corresponded to one of the G bands, which consist mainly of GC-poor L isochores and, to a lesser extent, H1 isochores (Saccone *et al.* 1996). Expectedly, when moving from the standard 400-bands to the 850-bands resolution, this band was shown to be heterogeneous (Francke *et al.* 1994) and to comprise both $H3^-$ R bands and G bands (Saccone *et al.*, unpublished data). Two other integration sites, 4p15(16) and 15q24(25), could be localized with equal probability in G or R bands. The 850-bands resolution showed that the 4p15 band consisted of $H3^+$, $H3^-$ and G bands and that the 4p16 band consisted mainly of G bands and of one $H3^-$ R band. Another band, 15q24(25), corresponded to $H3^+$ and $H3^-$ chromosomal bands (Saccone *et al.*, unpublished data). The comparison of the data on the chromosomal localization of the most frequently observed HTLV-1 integration sites with the results on the determination of the peaks of the most intense hybridization with HTLV-1 probe at 42–44% GC suggest that the highest probability of localization of HTLV-1 sequences is in H1 isochores. Two less frequently observed sites of HTLV-1 integration, 6p23 and 6q12(13), are H3 bands (in the latter case, a possible alternative is a G band).

The present data revealed that HTLV-1 are not localized in the GC-richest H3 isochores and are not observed in the regions where only transcriptionally active proviruses are integrated. This is consistent with the characteristics of HTLV-1 in the MT4 cell line (which was used to obtain the HTHIV27 cell line) and shows a low level of expression of HTLV-1 RNA and proteins due to heavy methylation of HTLV-1 proviruses (Saggioro *et al.* 1990). Defective non-infective subviral HTLV-1 particles have been observed in MT4 cells (Susloparov *et al.* 1996). The presence of HTLV-1 viral particles has been confirmed in HTHIV27 cell line by electron microscopy where they were found in an approximately 1 : 100 ratio to HIV virions (Kushch *et al.* 1995). These data on the low level of HTLV-1 proviral expression are consistent with the results on the localization of HTLV-1 around 41–43% GC sequences where non-expressed HTLV-1 sequences are located and which comprise the borderline region between non-expressed and expressed proviruses.

Table 1. Characteristics of proviral HTLV-I DNA integration sites in the chromosomes of persistently HIV-1 infected HTHIV-27 cells.

No.	Chromosome locus	Frequency per population (cell number, %)	Cytogenetic/ compositional band[a]
1	4p15(16)	85	G/R($H3^*$)
2	4q12	25	R($H3^-$)
3	6p21	55	R($H3^*$)
4	6p23	20	R($H3^-$)
5	6q12(13)	10	G/R($H3^-$)
6	9q21	55	G
7	15q24(25)	32.5	R($H3^*$)/G
8	16p13	50	R($H3^*$)
9	16q22	72.5	R($H3^*$)

[a]At 400-band resolution.

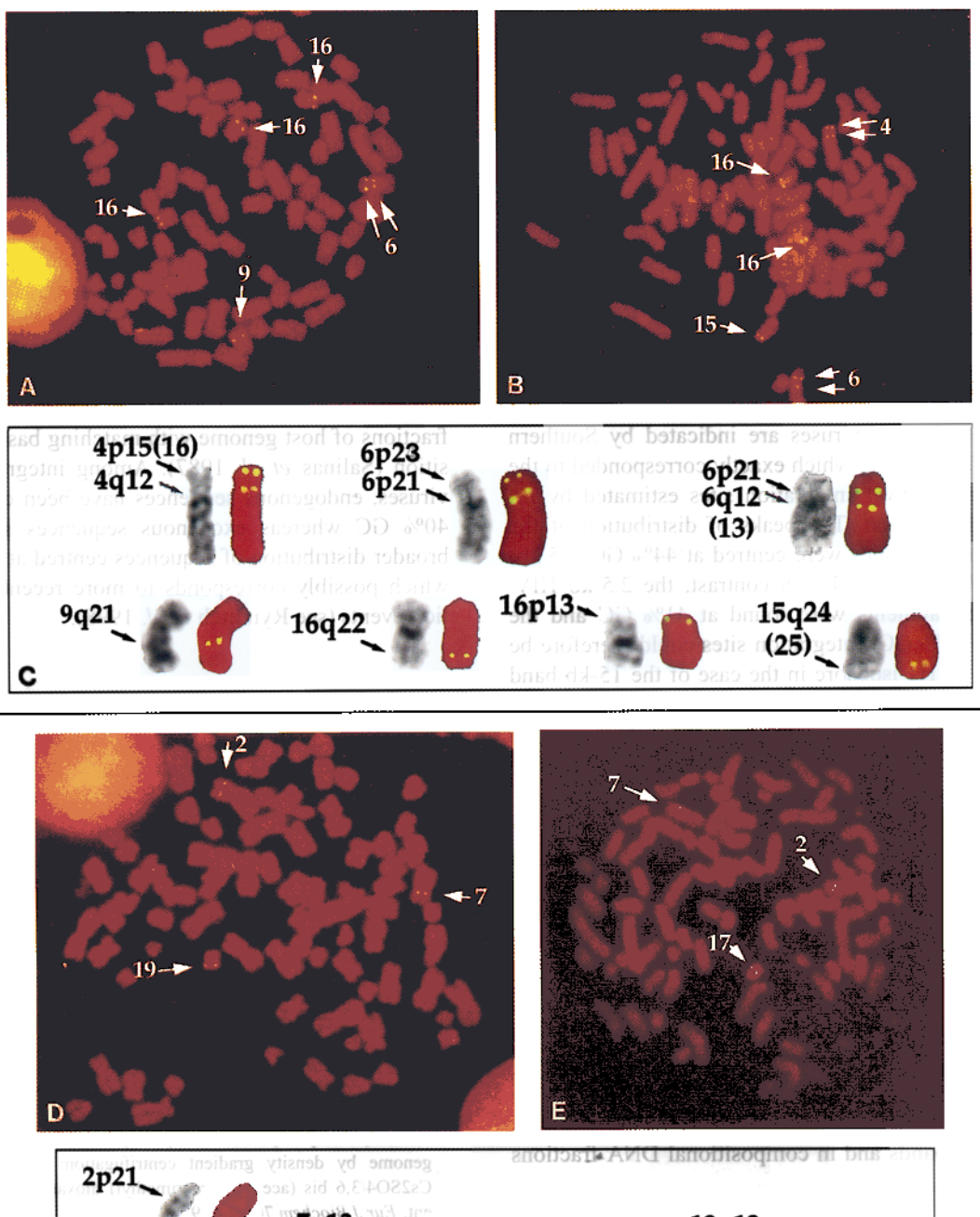

Figure 2. Fluorescent *in-situ* hybridization (FISH) with the pMT-2 and pBH10-R3 probes to metaphases of HTHIV27 cell line. **A, B.** Two metaphase spreads after FISH with pMT-2 probe. **C.** Pairs of chromosomes after FISH and G-banding from different metaphases (see Materials and methods) showing integration of HTLV-1 provirus in 9 loci. **D, E.** Two metaphases after FISH with pBH10-R3 probe. **F.** Pairs of chromosomes showing integration of HIV-1 provirus in 4 loci.

Compositonal distribtuion of HIV-1

The analysis of 103 metaphase spreads revealed HIV-1 proviruses in only four sites of integration (Figure 2D, E, F). Of the HTHIV27 cells, 90–70% contained HIV proviruses in bands 2p21 and 7p13. HIV-1 integration was less frequent in chromosomes 17 and 19 (Table 2). G-staining of chromsomes revealed that all integration sites are located in H3* R bands.

HIV-1 sequences were observed in compositional fractions corresponding to 40–46% GC (Figure 1B). The integration sites of individual proviruses are characterized by different XbaI fragments containing HIV-1. Since XbaI does not cut into the HIV-1 genome, four proviruses are indicated by Southern blot hybridization, which exactly corresponded to the number of HIV-1 integration sites estimated by *in-situ* hybridization. The peaks of distribution of the most intense bands were centred at 44% GC (15 kb) and 40% GC (7.4 kb). In contrast, the 2.5-kb HIV-specific fragments were found at 41% GC and the 4 kb at 46% GC. Integration sites could therefore be located in H1 isochore in the case of the 15-kb band or in L isochore in the case of the 7.4-kb band. The integration sites which were seen less frequently could be localized in L or H2 isochore families, respectively. The size of the bands representing HIV-1 proviral sequences indicated that only one, the 15-kb band, could correspond to the complete provirus. The HTHIV27 cell line is characterized by a long-term stable production of non-defective HIV-1 particles and therefore at least one complete HIV-1 provirus should be transcribed. This suggests that the 15-kb integrant is the only possible candidate for it. Cytogenetically, this integration site as the most strongly hybridized could be positioned at 2p21 chromosomal locus.

In conclusion, the localization of HIV-1 on chromosomal bands and in compositional DNA fractions revealed four integrated proviruses in the HTHIV27 cell line. The localization of the complete HIV-1 provirus which should be transcriptionally active as suggested by the production of complete HIV-1 viral particles is most likely in the H1 isochore family. Two other HIV-1 defective proviruses are distributed most probably in L isochores. While the latter correspond to an isopycnic location, the former is in a GC-rich environment. This situation was already found in some MMTV integrates and may be associated with the infection of HTHIV27 cells. Indeed, a similar situation has been observed for another GC-poor retrovirus MMTV (43.3% GC). MMTV proviral sequences have been localized in the compositional fractions of host genome with matching base composition (Salinas *et al.* 1987). Among integrated proviruses, endogenous sequences have been centred at 40% GC whereas exogenous sequences showed a broader distribution of sequences centred at 43% GC which possibly corresponds to more recent integration events (see Rynditch *et al.* 1998).

Table 2. Localization of HIV-1 proviral integration sites in human chromosomes and compositional DNA fractions.

Chromosome locus	Frequency per population	Bands	Fragment size (kb)	GC (%)
2p21	89	R(H3*)	15	44
7p13	70	R(H3*)	7.4	40
17q21	23	R(H3*)	2.5	41
19p13	15	R(H3*)	4	46

Acknowledgements

We are grateful to Anatolii Michailik for help in experiments and Salvatore Saccone for discussions.

References

Bernardi G (1995) The human genome: Organization and evolutionary history. *Ann Rev Genet* **29**: 445–476.

Clarke MF, Gelman EP, Reitz M (1983) Homology of human T-cell leukaemia virus envelope. *Nature* **305**: 60–62.

Coffin JM (1990) Retroviridae and their replication. In: Fields BN, Knipe DM, *Virology*, 2nd edn. New York: Raven Press, pp. 1437–1500.

Cortadas J, Macaya G, Bernardi G (1977) An analysis of the bovine genome by density gradient centrifugation: fractionation in Cs2SO4/3,6 bis (acetatomercurimethyl) dioxane density gradient. *Eur J Biochem* **76**: 13–19.

Dutrillaux B (1973) Nouveau système de marquage chromosomique: les bandes T. *Chromosoma* **41**: 395–402.

Fisher AG, Collati E, Ratrner L, Gallo RC, Wongtaal F (1985) A molecular clone of HTLVII with biological activity. *Nature* **316**: 262–265.

Francke U (1994) Digitized and differentially shaded human chromsomes ideorrams for genomic applications. *Cytogenet Cell Genet* **65**: 206–208.

Kettmann R, Meunier-Rotival M, Cortadas J *et al.* (1979) Integration of bovine leukemia virus DNA in the bovine genome. *Proc Natl Acad Sci USA* **76**: 4822–4826.

Kushch AA, Glukhova LA, Miller GG (1995) Cytogenetic analysis of continuous cell lines sensitive to and infected by HIV-1. In: Lvov DK, ed. *Virology Reviews*. New York: Harwood Academic Publishers, pp. 83–112.

Lawrence JB, Marselle LM, Byron KS, Johnson CV, Sullivan JL, Singer RH (1990) Subcellular localization of low-abundance human immunodeficiency virus nucleic acid sequences visualized by fluorescence in situ hybridization. *Proc Natl Acad Sci USA* **87**: 5420–5424.

Mellors JW, Rinaldo CR, Jr Gupta P, White RM, Todd A, Kingsley LA (1996) Prognosis in HIV-1 infection predicted by the quantity of virus in plasma. *Science* **272**: 1167–1170.

Miller G, Bykovsky APh, Pokidisheva LN, Titova IV (1992) Establishment and characterization of a new virus-producing permanent T-lymphoblastoid cell line. *VIII International Conference on AIDS*, Amsterdam, 19–23 June, Abstract Book, 3: 29.

Miyoshi I, Kubonishi I, Yoshimoto S *et al.* (1981) Type C-virus particles in a cord T-cell line derived by cocultivating normal human cord leukocytes and human leukemic T cells. *Nature* **294**: 770–771.

Rynditch A, Kadi F, Geryk J, Zoubak S, Svoboda J, Bernardi G (1991) The isopycnic, compartmentalized integration of Rous sarcoma virus sequences. *Gene* **106**: 165–172.

Rynditch A, Zoubak S, Tsyba L, Gouley N, Bernardi G (1998) The regional integration of retroviral sequences into the mosaic genomes of mammals. *Gene* **222**: 1–16.

Saccone C, De Sario A, Wiegant J, Rap AK, Dell Valle G, Bernardi G (1993) Correlations between isochores and chromosomal bands in the human genome. *Proc Natl Acad Sci USA* **90**: 11929–11933.

Saccone S, Cacciò S, Kusuda J, Andreozzi L, Bernardi G (1996) Identification of the gene-richest bands in human chromosomes. *Gene* **174**: 85–94.

Saggioro D, Panozzo M, Chieco-Bianchi Z (1990) Human T-lymphotropic virus type I transcriptional regulation by methylation. *Cancer Res* **50**: 4968–4973.

Sakai M, Kawamura M, Sakuragi JJ *et al.* (1993) Integration is essential for efficient gene expression of human immunodeficiency virus type 1. *J Virol* **67**: 1169–1174.

Salinas J, Zerial M, Filipski J, Crépin M, Bernardi G (1987) Nonrandom distribution of MMTV proviral sequences in the mouse genome. *Nucl Acids Res* **15**: 3009–3022.

Singer RH, Byron KS, Lawrence JB, Sullivan JL (1989) Detection of HIV-1-infected cells from patients using nonisotopic in situ hybridization. *Blood* **74**: 2295–2301.

Susloparov MA, Krendelschikov AV, Bakhtina MM *et al.* (1996) Cloning of proviral DNA squences of human T-cellular leukemia virus type 1 (HTLV-1) from the genome of MT-4 cell culture. *Mol Gen Mikrobiol Virusol* **2**: 32–39.

Zerial M, Salinas J, Filipski J, Bernardi G (1986) Genomic localization of hepatitis B virus in a human hepatoma cell line. *Nucl Acids Res* **14**: 8373–8386.

Zoubak S, Richardson J, Rynditch A *et al.* (1994) Regional specificity of HTLV-I proviral integration in the human genome. *Gene* **143**: 155–163.

Zoubak S, Clay O, Bernardi G (1996) The gene distribution of the human genome. *Gene* **174**: 95–102.