Functional Characterization of Transgene Integration Patterns by Halo Fluorescence in Situ Hybridization: Electroporation versus Retroviral Infection†

S. Goetze, Y. Huesemann, A. Baer, and J. Bode*

GBF-German Research Center for Biotechnology/Epigenetic Regulation, Mascheroder Weg 1, D-38124 Braunschweig, Germany

Received January 16, 2003; Revised Manuscript Received March 12, 2003

ABSTRACT: Gene transfer techniques are essential for bioengineering and gene therapy. Retroviral vectors constitute an important tool in this context as integrase-catalyzed genomic anchoring is one of the best defined results of nonhomologous recombination, occurring at loci with favorable expression properties. On the basis of a retroviral expression cassette, this study focuses on differences regarding the genomic targets after its transfer by either retroviral infection or electroporation. Fluorescence in situ hybridization (FISH) of clones generated by infection revealed for the majority an association of the transgene with the nuclear matrix while this was true only for a minority of single copy clones if the same construct was transferred by electroporation. Our results demonstrate that the process of integration is distinct for both gene transfer routes and confirm that retroviral survival strategies favor integration next to scaffold/matrix attachment regions. These results may be one key to explain recent consequences of gene therapy trials that have led to the deregulation of endogenous genes.

Reverse genetics, i.e., gene transfer into mammalian cells, is a tool of fundamental and practical interest (*1*). Mammalian cells readily insert foreign DNA into their chromosomes mostly by integration into preexisting breaks or, in the case of retroviral infection, by nonhomologous recombination catalyzed by retroviral integrase (IN)[1] (*2*). This step is an essential part of the retroviral replication cycle.

Following infection and reverse transcription, the viral DNA forms a complex together with IN and other viral proteins. This preintegration complex is transported into the nucleus where integration occurs. Selection of target sites is nonrandom, and the central domain of IN appears to play some role in determining the specificity (*3*). What do we know about the architecture of these sites? Although over the last 20 years many contributions have dealt with this problem, our understanding of the basic interactions of the IN−preintegration complex with both proviral and host cell DNA is still incomplete. For some cases the proximity of a DNase I hypersensitive site has been noted (*4*), indicating the presence of structural or regulatory elements. However, despite early reports to the contrary, no unique sites have been documented (*2*). We have demonstrated in 1996 that S/MARs appear to be the preferred integration targets (*5*) and have more recently emphasized the relation between these targets and fragile genomic sites (*6*). A number of subsequent reports seem to support such a conclusion:

Katz et al. (*7*) describe highly preferred in vitro retroviral integration sites within the stems of plasmid DNA cruciform structures where they occur adjacent to the loops in the cruciform. We have shown related properties for several retroviral integrants (*8*) and have argued that secondary structure formation may occur as a consequence of the demonstrated strand separation potential in these regions. This, together with the occurrence of bent sites, is a prominent characteristic of scaffold/matrix attachment regions (*5*) and may facilitate both the processing and the joining steps in the integration reaction. There is another recent report demonstrating the proximity of integrations sites and S/MARs in neoplastic cells where tumor viruses are thought to be subject to S/MAR-mediated transcriptional regulation (*9*).

A considerable part of the integration targets localizes to the vicinity of repetitive elements, and for HIV-1 Stevens and Griffith (*10*) have demonstrated that as much as 7/8 proviruses integrated at an L1 or Alu sequence. This has later been confirmed for plants (*11*). Rather than being S/MARs themselves, these retroposons are considered the markers of nearby S/MARs as they share the integration preference with retroviruses.

Due to the ordered integration mechanism the structure and location of retroviral integration sites can be defined with great precision by either inverse PCR (*5*) or the more advanced SLIP (solo long terminal repeat inverse PCR) technique (*3*). Unfortunately, this is mostly impossible for any transfection process as this is usually accompanied by deletions and even translocations (*12*) preventing the assignment of corresponding flanks and the design of the appropriate PCR primers (see discussion in ref *5*).

Scaffold/matrix attachment regions are stretches of DNA that are functionally linked to transcription. While S/MARs do not conform to any obvious sequence consensus, their recognition is governed by the above-mentioned structural features. For their characterization various in vivo and in vitro methods have been established. The former group

† This work was supported by a grant from the Deutsche Forschungsgemeinschaft (Bo 419/6) and BMBF Grant 01 KW 0003.

* To whom correspondence should be addressed. E-mail: jbo@gbf.de. Tel: +49 531 6181 251. Fax: +49 531 6181 262.

[1] Abbreviations: BUR, base-unpairing region; FISH, fluorescence in situ hybridization; IN, retroviral integrase; LTR, long terminal repeat; S/MAR, scaffold/matrix attachment region.

10.1021/bi0340907 CCC: $25.00
Published on Web 05/17/2003

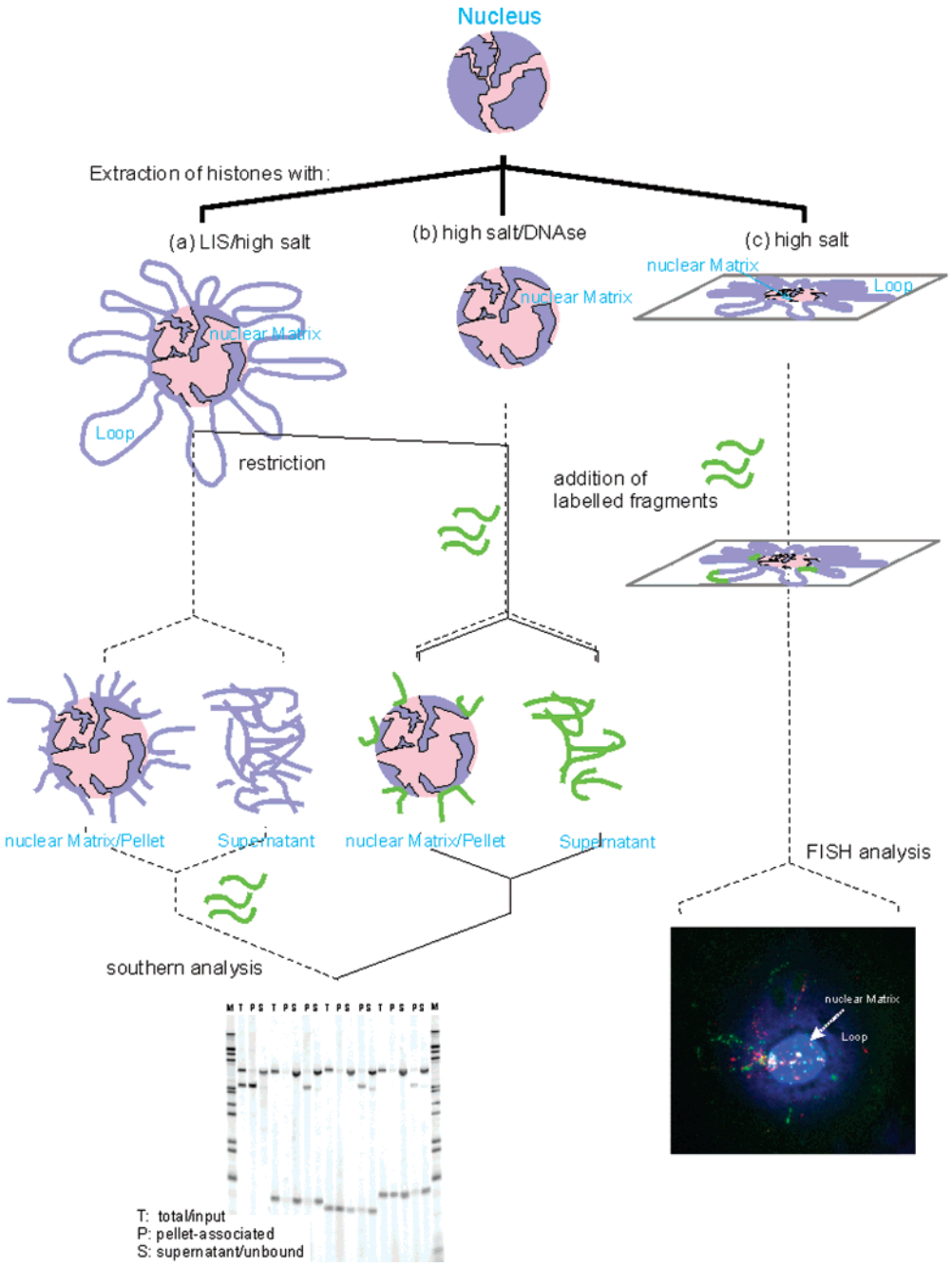

FIGURE 1: Characterization of S/MARs by conventional in vitro and by in situ (FISH) procedures. (a) Halo mapping analysis. (b) Reassociation experiments: a mixture of restriction elements is exposed to an extracted nuclear matrix; fragments with a high binding affinity are identified. The outcome of reassociation experiments is largely independent of the extraction procedure. (c) Halo-FISH in situ mapping procedure.

mainly includes minimal transcription and replication systems (*13*) whereas the latter is based on extraction procedures (Figure 1 and ref *14*). Eventually, these extraction procedures have been subject to criticism because of conditions, which were suspected to shuffle DNA-matrix contacts. These objections are largely overcome with the advent of fluorescent in situ hybridization (FISH) techniques: in a nuclear matrix protocol cell nuclei are immobilized on slides before their extraction. Since the standard protocol circumvents DNase I degradation steps, a major reorganization of the underlying structures is effectively inhibited (*15*). The extracted DNA halos are subsequently hybridized with fluorescent DNA probes. According to signal localization, S/MARs are identified and can be classified as constitutive or facultative (optionally used) elements that are involved in the potentiation of a chromatin domain (*16*). Such a high-resolution halo-FISH protocol is suited to detect single copy integrates and to resolve two probes separated as little as 1 kb (*17*, *18*).

In this study a halo-FISH technique is used to reveal integration preferences for single copy transgenes that have been introduced by either infection or electroporation. We observe a consistent, expression level-independent association potential to the nuclear matrix of retroviral infectants while electroporated proviral inserts mainly localize to the loop fraction with some more attachment for the highly producing clones. With the added observation that these classes of clones differ in their expression characteristics (*19*), we conclude that the integration pattern for both gene transfer routes is distinct.

MATERIALS AND METHODS

Cell Culture. Baby hamster kidney cells, strain BHK-A (*20*), a subclone of BHK-21 (ATCC CCL-10), were cultured in DME medium containing 10% FCS, 20 mM glutamine, 60 μg/mL penicillin, and 100 μg/mL streptomycin. The investigated clones were generated either by electroporation or by infection of the retroviral expression plasmid pM5βgeo (7.6 kb in size; see ref *19*). Individual clones were cultivated the same way as described above but with addition of 1 mg/mL G418. Unstable clones (I18, I10, 3H, and 4H), which lose gene expression when kept without selection pressure, were cultivated both with or without the selective drug.

β-Galactosidase Reporter Assay (MUG Assay). To determine β-galactosidase activity, the cells were washed with PBS and harvested with TEN [40 mM Tris-HCl (pH 7.5), 1 mM EDTA, and 150 mM NaCl]. They were pelleted by centrifugation (5 min at 200*g*) and subsequently resuspended in 200 μL of 0.25 M Tris-HCl (pH 7.8). Cell lysates were generated by three successive cycles of freezing and thawing, followed by centrifugation at 13700*g* at 4 °C for 5 min and collection of the supernatant fraction. The activity of β-galactosidase in the lysates was determined via hydrolysis of the β-Gal substrate 4-methylumbelliferyl D-galactoside (MUG) using 30 μg of MUG/mL in 200 μL of a substrate buffer (60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$, and 50 mM 2-mercaptoethanol). Fluorescence was determined at 5 min intervals in a Wallace Victor multiple counter plate reader with excitation at 365 nm and emission at 450 nm. Readings for each sample were determined in triplicate, corrected for the cell number, and referenced to a known concentration of purified β-galactosidase (Sigma).

β-Galactosidase in Situ Staining. After being washed with PBS the cells were covered with fixing solution (0.1% glutaraldehyde and 2% formaldehyde in PBS) for 2 min, washed twice with PBS, and covered with staining solution [5 mM $K_4Fe(CN)$, 6.5 mM $K_3Fe(CN)_6$, 2 mM $MgCl_2$, and 100 μg/mL X-Gal]. After at least 30 min at 37 °C blue β-galactosidase-expressing cells were counted. To determine the extent of gene silencing, the percentage of unstained cells was calculated.

Standard Reassociation Analyses. A pilot study was performed in which we traced the integration specificity of MPSV-derived vectors by the standard reassociation assay (*14*). Using identical 4 bp termini as a criterium for the assignment of corresponding 5′ and 3′ flanks, genomic DNA adjacent to the integration site was isolated for eight clones by inverse PCR (*5*) and subjected to binding analyses as described before (*14*). Binding strength was referenced to a standard from the huIFNB1 upstream S/MAR (fragment IV in ref *21*) and to a nonbinding plasmid control. PCR primers were designed in order to reconstruct, from genomic DNA of noninfected cells, 200−340 bp sequences corresponding to the preintegration state (Table 1). It is seen that all reconstructed sites have a high (58−95%) association potential, because integration occurred either at the border of a S/MAR (clones I19 and I25) or into a region with a more uniform association potential.

Preparation of Nuclear Halos. Cells were harvested by trypsinization, centrifuged at 200*g* for 2 min, and stored at −70 °C in FSB [50 mM Hepes, 10 mM NaCl, 5 mM $(CH_3CO_2)_2Mg$ (pH 7.5), and 25% glycerol]. Prior to extraction cells were thawed, washed twice with PBS, and incubated on ice with CSK buffer [10 mM Pipes, 100 mM NaCl, 0.3 M sucrose, 30 mM $MgCl_2$, and 1% Triton X-100] for 15 min. Isolated nuclei were counted, and 1.5E4 nuclei in 50 μL were pelleted onto slides using a Cytospin centrifuge (800 rpm for 5 min). Slides were treated with a 2 M NaCl buffer [2 M NaCl, 10 mM Pipes (pH 6.8), 10 mM EDTA, 0.1% digitonin, 0.05 mM spermine, and 0.125 mM spermidine] for 4 min to extract histones and soluble proteins. This extraction procedure releases DNA loops from the nuclear remnants resulting in nuclear halo structures. The slides were then subjected to a series of 10×, 5×, 2×, and 1× PBS washes followed by a series of rinses at 70%, 90%, and 100% ethanol concentrations. They were air-dried and baked for fixation at 60−70 °C for 2 h.

Table 1: S/MAR Binding Activity[a]

clone	5′-flank	3′-flank	preintegration state (bp)
I11	43	77	58 (230)
I14	41	47	75 (220)
I19	17	69	nd; adjacent to B1 DA
I20	82	71	80 (340)
I24	73	88	95 (280)
I25	64	5	90 (260)
I26	70	98	86 (200)
I28	34	72	nd

[a] S/MAR binding activity (percent of fragment in pellet fraction) for genomic DNA present adjacent 5′ or 3′ from a retroviral integration site or (after PCR reconstruction) of a fragment that represents the preintegration state (fragments with the integration site at a central position).

Fluorescence in Situ Hybridization (FISH): (A) Proof of Principle. Our FISH approach for evaluating the attachment potential is based on the proportion of signals colocalizing with the loop or matrix parts of a nuclear halo. We had to estimate to what extent an incidental colocalization of nonattached signals with the central matrix portion overestimates the attachment propensity. As a first approach in this direction we generated, by overlaying 15 images with a distance of 0.1 μM in the *Z*-direction (computated by Voxblast), a side view indicating that, in fact, there is nonattached DNA found on top of the matrix (Figure 2a). Next we used the criteria by Raetsch et al. (*22*), i.e., a centromere-specific probe (Qbiogene) that is known to be not matrix-associated (Figure 2b). Across the structure we performed five independent density scans, which, if overlaid, indicated an even distribution over the entire nuclear halo structure. This had to be substantiated by an authentic matrix-specific probe for which we chose telomeric sequences (Qbiogene) according to ref *22* (Figure 2c). As an intermediate situation we included one experiment with probes directed toward the murine chromosome 4 territory (Qbiogene) and obtained the expected bell-shaped distribution, indicating both attached and nonattached components (Figure 2d). The general correctness of this assumption was further verified by treating permeabilized nuclei with a restriction enzyme (*Xho*I) prior to DNA extraction. Such a pretreatment strongly reduced the proportion of apparently matrix-associated signals for the electroporants and to a lesser extent for the infectants (data not shown). With this information at hand we corrected the proportion of loop- and matrix-associated counts as shown in Table 2.

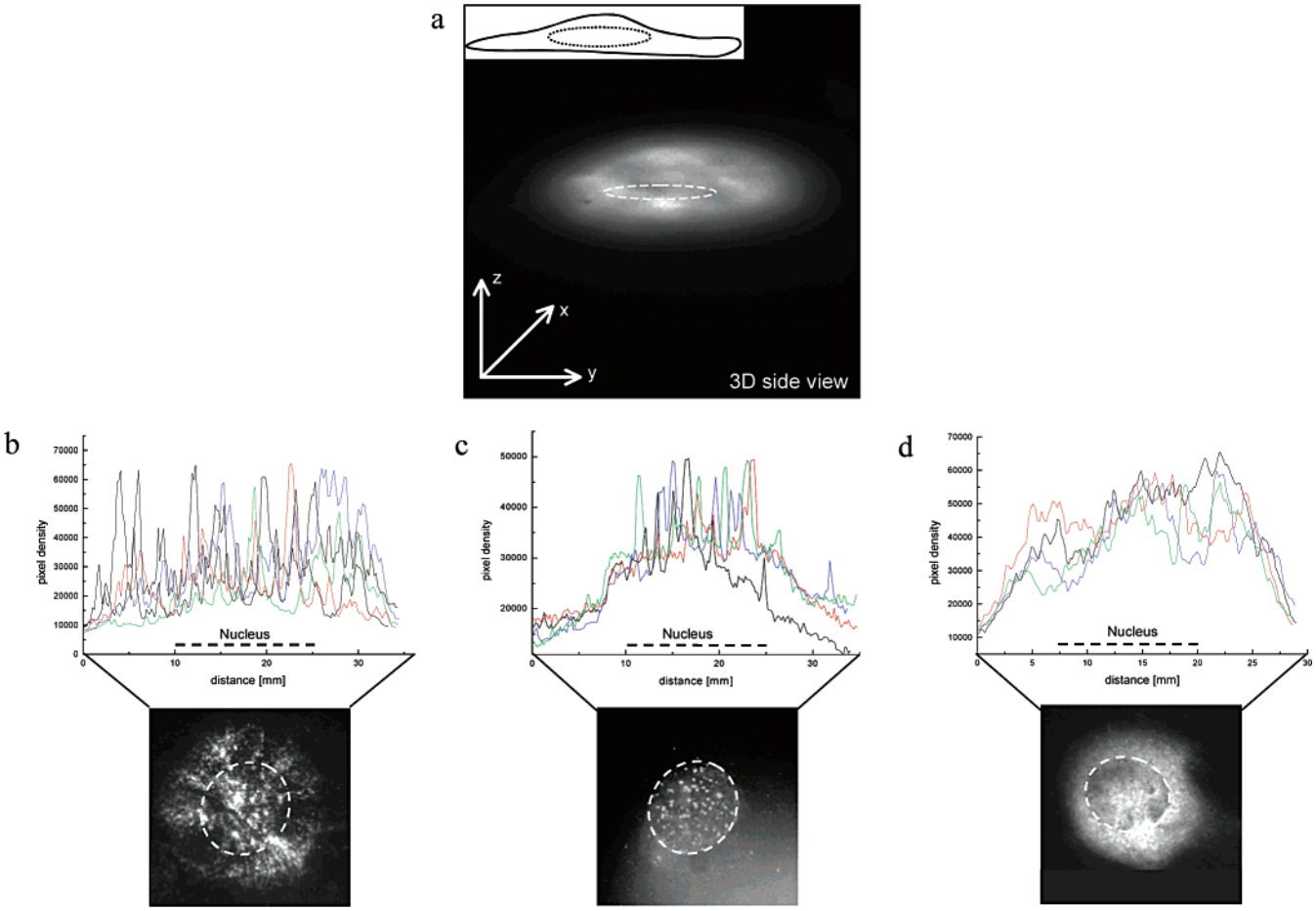

FIGURE 2: Proof of principle. (a) Side view of a nuclear halo indicating occurrence of nonattached (loop) DNA on top of the matrix portion. (b−d) Nuclear halos decorated with a nonbinding (centromeric) control (b), with a binding (telomeric) control (c), and with a probe mixture from a chromosome territory (d). Superimposition of densitograms from five diameter sections (edge−center−edge) demonstrates the even distribution of the nonattached control (b). Matrix portions have been visualized as for Figures 5 and 6, and contours have been indicated by dotted lines. On average, the matrix portion is one-third of the nuclear halo area, and this has been taken into account for data corrections in Table 2.

(B) Retroviral Inserts. The pM5βgeo transgene was detected by a nick-translated plasmid DNA probe (Roche nick translation kit). For a standard labeling reaction with 2 μg of template DNA in 20 μL total reaction volume, a dNTP mixture consisting of 1 volume of 0.4 mM digoxigenin-11-dUTP, 2 volumes of 0.4 mM dTTP, and 3 volumes of 0.4 mM dATP, dCTP, and dGTP is required. Ten microliters of this nucleotide mixture was added to the labeling reaction together with an enzyme mixture of DNA polymerase I and DNase I. After incubation at 15 °C for 2 h fragment sizes were determined on an agarose gel, and the reaction was stopped once a length distribution between 100 and 500 bp was attained. To remove the nonincorporated nucleotides, the DNA was subjected to a Sephadex G50 column (Amersham).

For hybridization 2 μL of the probe was denatured with 1 μL of mouse Cot-DNA in 12 μL of hybridization buffer [12.5% dextran sulfate, 62.5% deionized formamide, 2.5 × SSC, and 1% herring sperm DNA (pH 7.4)] at 80 °C for 10 min. DNA halos were fixed in methanol/acetic acid (10 min at room temperature) before denaturation in a formamide solution [70% formamide in 2 × SSC] at 72 °C for 5 min. Slides were dehydrated in 70%, 90%, and 100% ethanol, air-dried, and hybridized at 37 °C overnight. For the signal detection procedure the HNPP fluorescent detection set (Roche) was employed. Prior to addition of anti-Dig AP antibody the slides were rinsed in 2 × SSC (room temperature, 72 and 37 °C). After a 1 h incubation at 37 °C they were rinsed three times with washing buffer [100 mM Tris-HCl (pH 7.5), 150 mM NaCl, and 0.05% Tween 20] and twice with detection buffer [100 mM Tris-HCl (pH 8), 100 mM NaCl, and 10 mM $MgCl_2$]. Fifty microliters of HNPP/Fast Red TR mix was applied, and the slides were incubated for 30 min at room temperature. To detect the single copy transgene, the incubation step with the fluorescent agent was repeated once subsequent to another rinse with washing buffer. After a final washing step in H_2O the slides were air-dried and counter stained with 10 μL of DAPI (0.187 μg/mL in Vectashield mounting medium). Samples were examined with a Zeiss Axiovert 135TV microscope equipped with epifluorescence and filter sets from Omega Optical (Brattleboro, VT). Images were acquired with a Photometrics (Tucson, AZ) high-resolution, cooled charge-coupled device camera (PXL 1400, grade 2) for 12-bit image collection using custom software from SignalAnalytics (IPLab Spectrum). This software includes a coloring of black/white pictures taken at different wavelengths and a subsequent overlay.

RESULTS

Halo-FISH Analysis of Retroviral Infectants and Electroporants. For gene therapy, retrovirus-mediated transfer has gained particular interest as a means of mediating a high

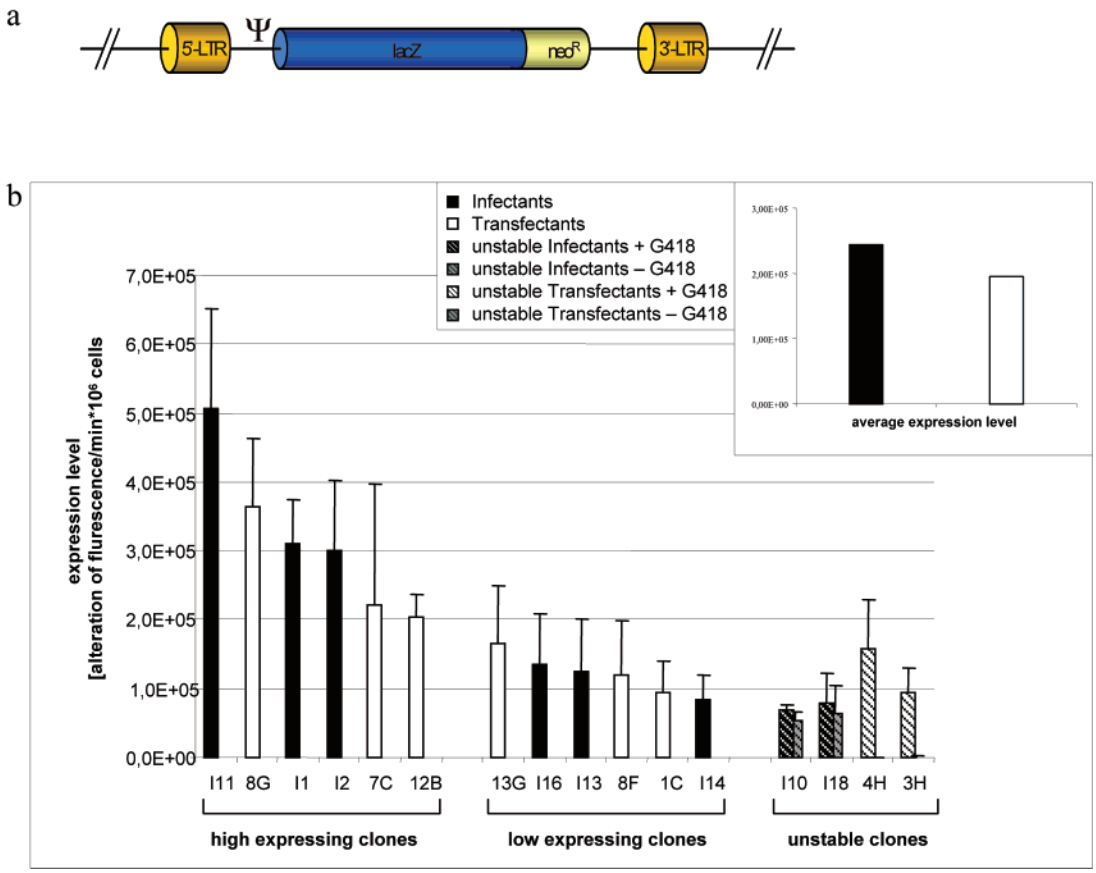

FIGURE 3: Expression properties of a proviral construct depending on the gene transfer route. (a) Structure of the transgene for both infectants and electroporants generated from the pM5βgeo plasmid. The vector contains the long terminal repeats (LTR) and the MPSV packaging signal (Ψ) and is used for the modification of BHK-A cells. (b) Expression levels for infectants (black bars) and for electroporated clones (white bars). Activities of the β-galactosidase reporter were quantified on cell extracts using the MUG assay and given as fluorescence change per 10^6 cells in 5 min. While the two classes of stable clones are similar regarding their expression range, unstable electroporants show a complete shutoff after the withdrawal of the selective drug (G418) whereas unstable infectants only contain a small population of shutoff clones (compare with Figure 5). Insert: average expression levels for infectants and electroporants, respectively.

and long-lasting expression, but there is recent concern that particular integration preferences may be the source of cancer (*23*). Present work compares clones that have been generated by two alternative gene transfer routes: retroviral infection and electroporation (another technique with an potential for therapeutic applications; see ref *24*).

The retroviral expression plasmid pM5βgeo (Figure 3a) was used to establish two series of authentic single copy clones for which the expression profile has been determined before (*19*). The vector contains the long terminal repeats (LTRs) and the packaging signal of the murine myeloproliferative sarkoma virus; retroviral coding sequences have been replaced by selection marker (neo) and reporter gene (β-galactosidase) functions. As such the vector has no affinity for an in vitro preparation of the nuclear matrix. For the electroporation series dispensable parts of plasmid DNA have been removed in order to approach the structure of an authentic provirus in the case of the electroporants.

The entire series contained 17 infectants and 21 electroporants (*19*). Although the three highest expressing clones originated from infection and the two lowest expressers were electroporants, average levels differed by only a factor 1.3. Within each group the behavior of members was identical, which led us to investigate eight representative clones from each of the two populations, infectants (solid black bars) and electroporants (open bars in Figure 3b). Each population contained three high-expressing and three low-expressing clones in addition to two unstable ones. The latter lose expression after relieving the selection pressure (hatched bars in Figure 3b; see also Table 2).

In a first step maintenance of the reported transcriptional activities after freezing and thawing had to be verified by the β-galactosidase reporter assay. Figure 3b shows the selected members arranged according to transcription level and stability. For the unstable clones the inactivation characteristics clearly depend on the gene transfer route as reported before (*19*): While unstable electroporants were exclusively found among the low expressers and shutoff was complete, the overall expression of infectants underwent little change even if cultivated without selective pressure. Still, with regard to single cell observations, the gene is completely silenced in some cells (see Figure 6a below), meaning that, in principle, a variegated expression can occur for both electroporated and infected clones. Together, these findings indicate that the process of retroviral infection places the provirus into a genomic environment which is mostly permissive for stable expression whereas electroporation also populates sites with a reduced intrinsic stability.

Halo-FISH Analysis of Stable Expressers. To examine the location of the transgene relative to the nuclear matrix, a halo-FISH procedure according to Figure 1 (right-hand track) was applied. The extracted nuclei of an individual clone were hybridized with the pM5βgeo transgene, and after visualization by a red fluorescent marker, the loop and matrix portions

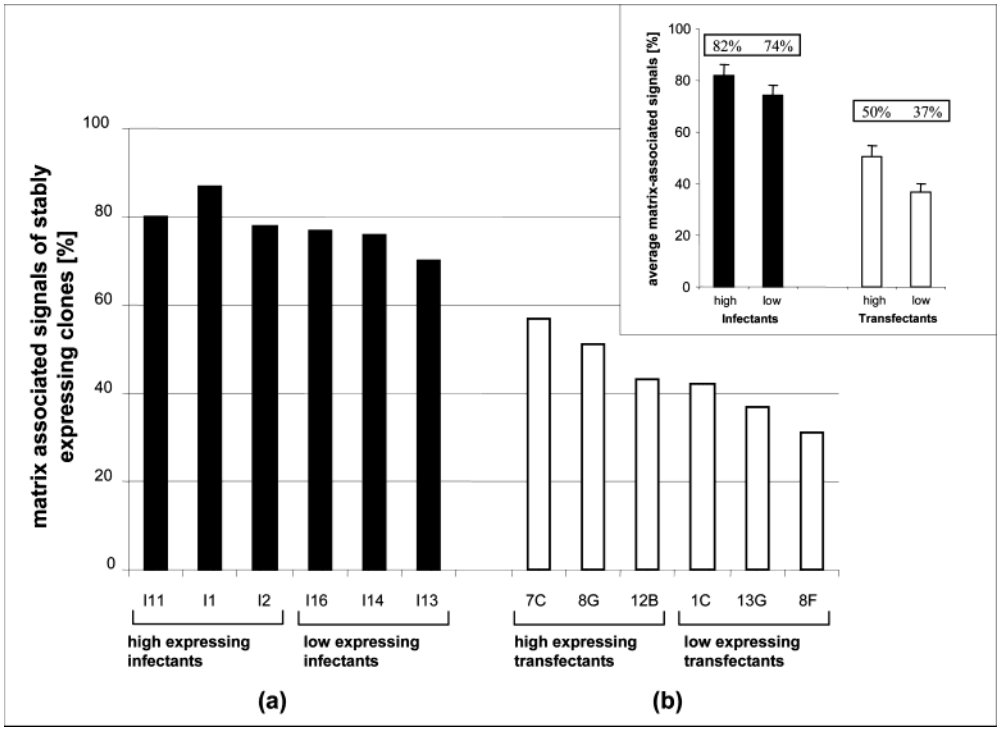

FIGURE 4: Matrix-association potential of infectants and electroporants (raw data without the area corrections in Table 2). Note that, in contrast to expression levels, there is a clear-cut separation between the groups of infectants and electroporants (sections a and b). Each bar averages the information from 40 to 50 pictures that have been taken per individual clone as exemplified in Figure 5. The average uncorrected association values of the different clone populations are depicted in the inset. Key: (■) infectants; (□) electroporated clones.

Table 2: Distribution of the Single Copy Proviruses over Loop and Matrix Portions of a Nuclear Halo[a]

	loop				loop		
infectants	visual part (%)	includes surface correction (%)	matrix attached (%)	electroporants	visual part (%)	includes surface correction (%)	matrix attached (%)
I11	20	30	70	7C	43	65	35
I1	13	19	81	8G	49	74	26
I2	22	33	67	12B	57	86	14
I16	23	35	65	1C	58	88	12
I14	24	37	63	13G	63	95	5
I13	30	45	55	8F	69	100	0
I10 (−G418)	*24 (26)*	*36 (39)*	*64 (61)*	*3H (−G418)*	*57 (80)*	*86 (100)*	*14 (0)*
I18 (−G418)	*37 (42)*	*56 (64)*	*44 (36)*	*4H (−G418)*	*67 (62)*	*100 (94)*	*0 (6)*
			64 ± 11				13 ± 12

[a] Columns mark direct counts and data after surface correction as described in Materials and Methods. Data for stable clones (rows 2−7) are indifferent whether they were maintained in the continued presence of G418 (see text). Rows 7 and 8 (in italics) are for unstable clones kept in the presence (absence) of G418 (data in parentheses were left out for calculating standard deviations). Finally, the overall attachment values including standard deviations are presented (last row).

were inspected by fluorescence microscopy. For each clone 40−50 single cell hybridizations were evaluated in at least two independent hybridization experiments. It was felt that this number of inspected specimens is adequate to derive statistical information about the proportion of free versus attached gene copies as described in Materials and Methods and summarized in Figure 1. The results are depicted in Figure 4.

Figures 4a and 5a confirm that all infectants show a high propensity for matrix association whether or not they belong to the class of high expressers: for this group, on average, 78% of the integrated proviruses is found to colocalize with the matrix portion of the nuclear halo. There is hardly any difference between low and high expressers, the average deviation between the investigated populations being about 8% (see inset to Figure 4). Minor differences for the two groups may be explained by the characteristics of the individual integration sites; i.e., highly expressed transgenes might have targeted stronger S/MARs and, in the postulated dynamic system of chromatin domains, attached to the nuclear matrix at a higher frequency.

A strikingly different association pattern was observed for the group of electroporants which in Figure 4 form a coherent group (panel b). Here, the majority of transgenes does not overlap the nuclear matrix portion but rather appears in the visible loop portion of the nuclear halo (also see Figure 5b). Whereas a maximum association (57%) is found for clone 7C, the minimum is associated with clone 8F (31%). For this group of clones the differences can readily be explained by the transcriptional activity which is affected by the integration locus: to be transcribed, genes on the loop have to move toward the nuclear matrix, a process that is facilitated by S/MAR sequences located in close proximity to the transgene (*5*, *13*; model in ref *25*). Therefore, differences in matrix association also reflect the transcriptional status in the case of the electroporants.

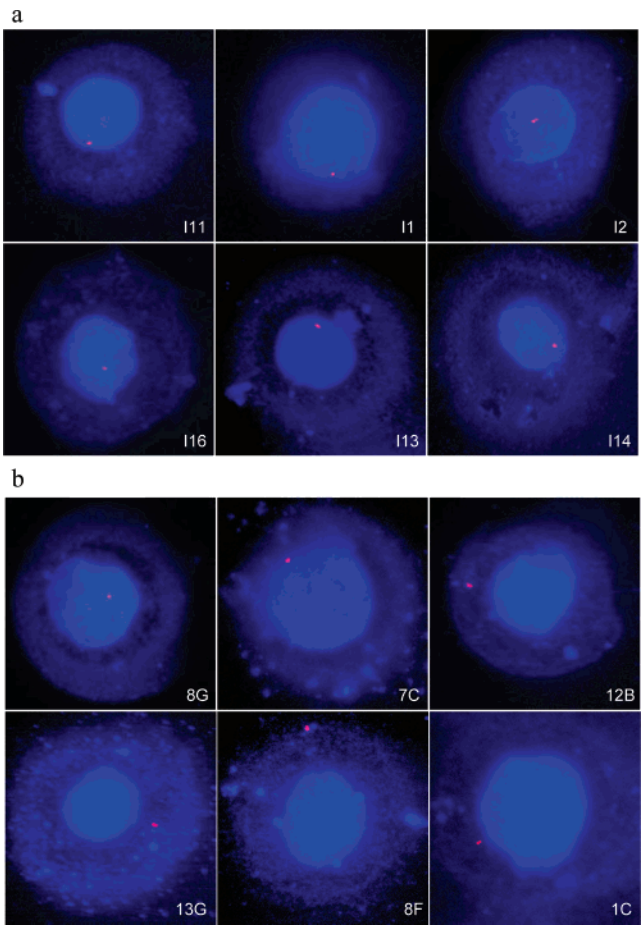

FIGURE 5: Matrix attachment of infectants and electroporants: halo-FISH analyses. The FISH signal of the pM5βgeo transgene is visualized as red spot. (a) Representative views of six infectants. (b) Representative views of six electroporants.

The overall conclusion that can be drawn from these experiments is that integration targets are significantly different depending on the gene transfer mode. These differences are even larger than immediately obvious from the imaging procedure: since the nuclear matrix portion covers about one-third in the center of the 2D top view image, signals located within this area are not necessarily associated with the nuclear matrix but rather parts of the halo fraction above the nucleus. We have compiled our evidence for this assumption in Materials and Methods and in Figure 2 and taken it into consideration as shown in Table 2. Here we list the statistical occurrence of signals in the visual part of the loop portion which accounts for two-thirds of the total. If surface correction is applied, the proportion of nonattached signals on top of the matrix can be extrapolated and actual matrix attachment can be calculated. From these data it can be concluded that for infectants 64 $\pm$ 11% of copies are attached at any given moment while this is true for only 13 $\pm$ 12% of electroporated copies.

So far, our results emphasize the importance of the integration route with regard to matrix attachment. Retroviral integration appears as a specific process with a higher propensity to integrate into destabilized DNA elements, which at the same time provide matrix attachment (*5*) and contribute to transcriptional potential (*13*). An additional feature emerging from Figure 4 (inset) is the fact that with 82% or 74%, respectively, there is barely any difference in matrix attachment between high- and low-expressing infectants (data after surface correction: 73% or 61%, i.e., a 1.2-fold difference). The same is not true for the electroporants for which high expressers show 50% and low expressers 37% attachment (after surface correction: 25% and 6%, 4.2-fold difference). These observations are readily explained by the requirement for electroporated copies to move to the nuclear matrix in order to be transcribed. Since infected proviruses show a higher basic matrix attachment already, such differences tend to be masked for this class of clones.

Matrix Association and Gene Silencing. The nuclear environments impose position effects on the expression of a transgene. If a gene is integrated in a heterochromatic environment, such a position might provoke silencing. Gene silencing is one of best studied epigenetic modifications and is often accompanied by DNA methylation, predominantly at the promoter. Owing to a tight coupling with the histone deacetylation apparatus (*26*), the same sequences will also have a low level of acetylated lysines. In this work we investigated several single copy clones which showed shutoff for some or all cells after several months in the absence of G418 (*19*). In all cases Southern blot analyses showed that the integrated transgene copy had remained intact (data not shown). The loss of expression can be ascribed to variegation since the transgene was completely silenced in a portion of cells (see X-Gal staining in Figure 6a which distinguishes this situation from one in which all cells show a decreased expression level). While in the presence of the selective drug most of the cells remained active as expected for a fusion protein with linked reporter and selector functions, clones kept in its absence exhibited marked differences, depending on the gene transfer route. Figure 6a indicates that under these conditions there was a complete shutoff for all electroporants whereas the infectants still contained ~47% of blue and thereby expressing cells.

Halo-FISH analysis revealed further differences for these clones (Figure 6b). As far as infectant I10 is concerned, the association frequency did not change significantly as a consequence of gene silencing. Therefore, by this criterion it was still comparable with the stable expressers (also see Table 2). For I18, on the other hand, the association frequency was reduced from 63% to 58%, respectively (Figure 6b). Both of these values are 10−20% below that of the stable infectants. Unstable electroporants are consistently clones with a low initial level of transgene expression (see Figure 3b) which, regarding their attachment, barely differs from the group of low but stable expressers. This is also reflected by the matrix attachment activities summarized in Table 2.

DISCUSSION

Retroviral vectors are still the most important tool for gene therapeutic applications. This is due, in part, to the fact that the integration machinery targets highly expressed genomic sites by integrating close to or within preexisting expression units. According to a conventional protocol (trace b in Figure 2) these targets are S/MARs by definition (Table 1). Until recently, the possible dangers of interfering with preexisting expression patterns has remained mere speculation, but this has definitely changed with reports on a case where a leukemia-like disease had to be ascribed to the integration event (*23*). For a better appreciation of the associated risks a detailed knowledge about the underlying principles is urgently required.

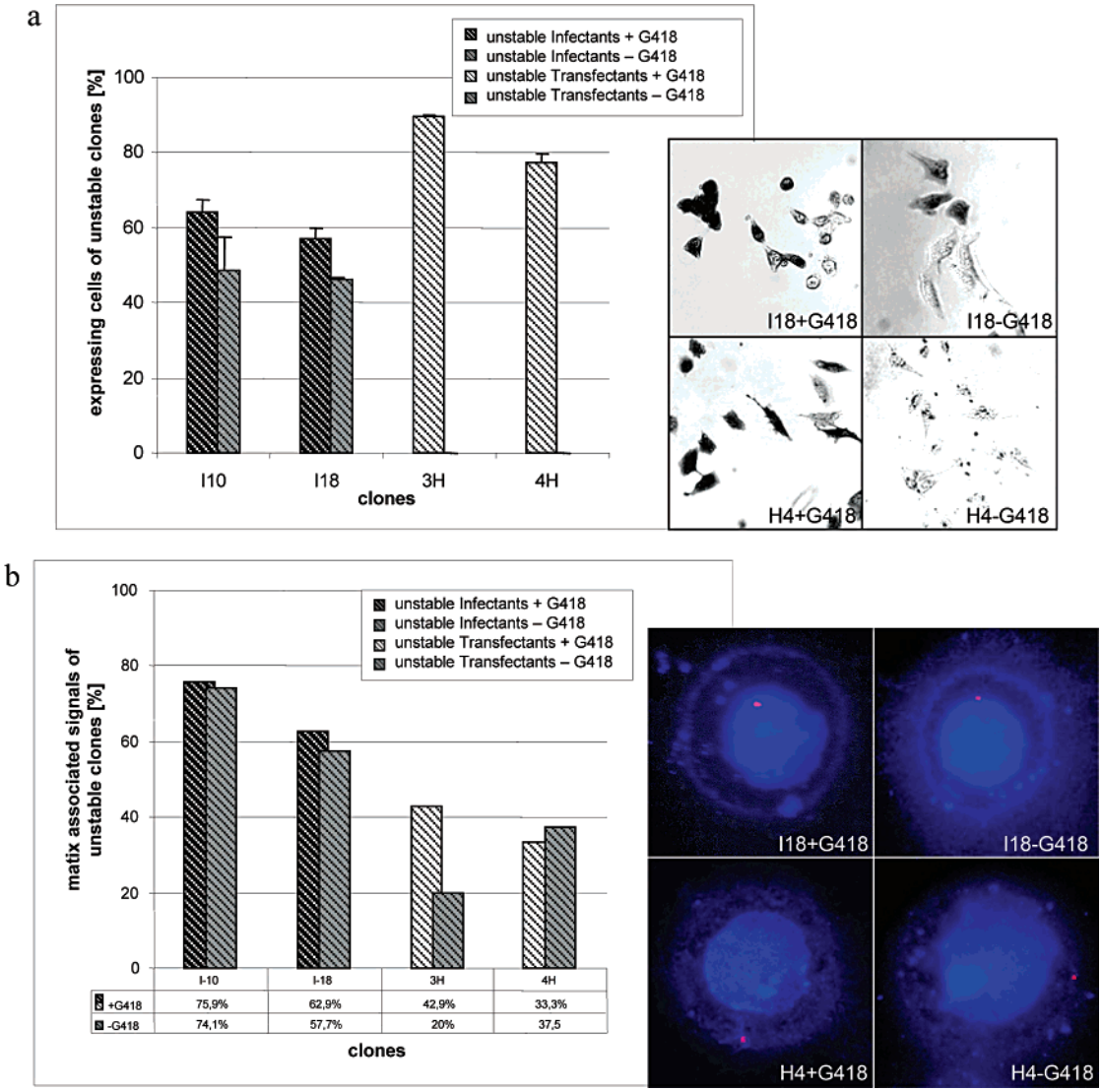

FIGURE 6: Unstable members among infectants and electroporants: shutoff and matrix association. (a) X-Gal blue/white staining of single cells shows a variegated expression for a proportion of cells that show a complete transgene silencing. Especially striking is the complete shutoff of electroporants cultivated without G418. The associated figures show, in extracts, the X-Gal stained cells. (b) Matrix-association behavior of the same clones. Each bar averages the information from 40 to 50 pictures that have been taken per individual clone. A representative clone appearance is depicted in the associated halo-FISH pictures.

In this study we have compared the integration status of a proviral transgene with regard to nuclear matrix association. Proviral DNA was transferred either by retroviral infection or by electroporation. Halo-FISH analysis reveals striking differences between these two populations. While electroporation generates clones with the provirus mainly located in the halo portion, matrix attachment predominates for the clones generated by retroviral infection (Figures 4 and 5 and Table 2). These differences indicate distinct integration preferences. While all transfection procedures depend on the occurrence of preexisting breaks that might well be generated at a class of fragile sites that arises adjacent to S/MARs (*6*, *27*) and while these are often associated with unpredictable rearrangements (*12*), retroviral infection is a precise enzymatically catalyzed process. However, even in this case functions of the cellular repair machinery are required which provide for the duplication of the terminal bases of cellular DNA (*5*). Since this system resides at the nuclear matrix (*28*), both routes of nonhomologous recombination might have been expected to share a similar integration preference.

For retroviral integration sites there is now firm evidence that prior to insertion the site is consistently matrix-associated (ref *5* and introduction). However, all previous studies had to leave unanswered the question whether matrix anchorage is altered or lost after the insertion of the provirus. We have established systems by which S/MAR activities can be treated in a quantitative manner (*14*), and during these studies we found that binding strength can be severely perturbed by extra DNA sequences that have been either inserted into a S/MAR or added at its borders (*29*). If these forces were also valid during provirus integration, its initial contact with the transcription machinery could be lost, since the proviral DNA itself is devoid of matrix-association potential. One of the major outcomes of this study is therefore the demonstration that proviruses do not perturb the attachment of their genomic surroundings. In this position they experience a selective advantage which may explain that this situation is found in a wide variety of neoplastic cells (*9*). Following a similar kind of reasoning it has been argued that the matrix-attached group of clones survives as a consequence of selection pres-

sure. While we cannot rule out that there might exist a proportion of subthreshold expressers, we have shown that the sensitive procedure of G418 selection permits survival of cells with an at least 8-fold difference of expression levels (*19*). By reference to the electroporants it has become evident that the required expression level can be met in the absence of matrix attachment. Nevertheless, a selective advantage may arise during long-term cultivation.

After the discovery of S/MARs, authors were initially in favor of a fixed S/MAR-anchoring model. This original, static model was primarily based on observations during *Drosophila* development indicating that enhancer-associated S/MARs of three developmentally regulated genes (Adh, ftz, Sgs-4) were attached in both embryonic and adult cells [Gasser and Laemmli (*30*)]. Similar results were reported for the hsp70 gene cluster (*31*). For *Xenopus* genes, on the other hand, data support a dynamic nuclear matrix anchorage model. For this system it was demonstrated that, for instance, the keratin gene is anchored at times when the chromatin domain is actively transcribed (*32*). In such a setup, matrix contacts are supported at times of transcriptional activity owing to superhelical changes which are transmitted to the S/MARs, the structure of which changes according to topological stress.

We have shown in this work that, according to Table 2, at any given time the majority (64%) of the infected proviruses is matrix-associated while 36% is released into the loop. Considering the fact that S/MAR elements undergo a dynamic association with the matrix (*16*; Heng et al., submitted for publication) and that the number of sites is limited (*21*), loop localization may arise at instances where neighboring sequences are transcribed (*33*). Comparing the high- and low-expressing members of this group, we observe an only moderately higher nuclear matrix attachment rate for the high expressers, which reflects the fact that the majority of provirus integration sites is predominantly matrix-associated at any time. For the electroporants transcription-dependent differences are more prominent as these have integrated in a more or less random mode.

Regarding the mentioned incidence of a leukemia-like disease after a gene therapy protocol with a retroviral vector, efforts have been initiated to create a map of all known human integration sites to gain a better appreciation of the risk (*34*). The findings reported here will not only provide a better understanding of these sites' function but also suggest protocols by which establishment of such a functional map may be supported. All of these efforts can now be accompanied by novel computational methods in order to establish or refine the principles that have already emerged (*29*).

While certain DNA structures may be the primary signal to guide integration of retrotransposons and retroviruses, interactions with yet unknown chromosomal proteins may modulate these preferences. If these preferences could be modified as demonstrated for a yeast model (*35*), new retroviral vectors might become available that avoid classes of critical targets.

ACKNOWLEDGMENT

The authors express their gratitude to Henry Heng for enthusiasm and support and for hosting Sandra Goetze in his laboratory.

REFERENCES

1. Scangos, G., and Ruddle, F. H. (1981) *Gene 14*, 1−10.
2. Withers-Ward, E. S., Kitamura, Y., Barnes, J. P., and Coffin, J. M. (1994) *Genes Dev. 8*, 1473−1487.
3. Jin, Y. F., Ishibashi, T., Nomoto, A., and Masuda, M. (2002) *J. Virol. 76*, 5540−5547.
4. Rohdewohld, H., Weiher, H., Reik, W., Jaenisch, R., and Breindl, M. (1987) *J. Virol. 61*, 336−343.
5. Mielke, C., Maass, K., Tümmler, M., and Bode, J. (1996) *Biochemistry 35*, 2239−2252.
6. Bode, J., Benham, C., Ernst, E., Knopp, A., Marschalek, R., Strick, R., and Strissel, P. (2000) *J. Cell. Biochem. 35* (Suppl.), 3−22.
7. Katz, R. A., Gravuer, K., and Skalka, A. M. (1998) *J. Biol. Chem. 273*, 24190−24195.
8. Bode, J., Bartsch, J., Boulikas, T., Iber, M., Mielke, C., Schübeler, D., Seibler, J., and Benham, C. (1998) *Gene Ther. Mol. Biol. 1*, 551−880 (http://www.gtmb.org/volume1/29_Bode.htm).
9. Shera, K. A., Shera, C. A., and McDougall, J. K. (2001) *J. Virol. 75*, 12339−12346.
10. Stevens, S. W., and Griffith, J. D. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 5557−5561.
11. Tikhonov, A. P., Lavie, L., Tatout, C., Bennetzen, J. L., Avramova, Z., and Deragon, J. M. (2001) *Chromosome Res. 9*, 325−337.
12. Kato, S., Anderson, R. A., and Camerini-Otero, D. (1986) *Mol. Cell. Biol. 6*, 1787−1795.
13. Bode, J., Goetze, S., Ernst, E., Hüsemann, Y., Baer, A., Seibler, J., and Mielke, C. (2003) in *New Comprehensive Biochemistry—Gene Transfer and Expression in Mammalian Cells* (Makrides, S., Ed.) Elsevier, Amsterdam (in press).
14. Kay, V., and Bode, J. (1995) in *Methods in Molecular and Cellular Biology: Methods for studying DNA-protein interactions—an overview* (Papavassiliou, A. G., and King, S. L., Eds.) Vol. 5, pp 186−194, Wiley-Liss, New York.
15. Heng, H. H., Krawetz, S. A., Lu, W., Breer, S., Liu, G., and Ye, C. J. (2001) *Cytogenet. Cell Genet. 93*, 155−161.
16. Kramer, J. A., and Krawetz, S. A. (1996) *J. Biol. Chem. 271*, 11619−11622.
17. Heng, H. H., and Tsui, L. C. (1998) *J. Chromatogr., A 806*, 219−229.
18. Raap, A. K. (1998) *Mutat. Res. 400*, 287−298.
19. Baer, A., Schubeler, D., and Bode, J. (2000) *Biochemistry 39*, 7041−7049.
20. Karreman, C., Karreman, S., and Hauser, H. (1996) *Virology 220*, 46−50.
21. Mielke, C., Kohwi, Y., Kohwi-Shigematsu, T., and Bode, J. (1990) *Biochemistry 29*, 7475−7485.
22. Raetsch, A., Joos, S., Kioschis, P., and Lichter, P. (2002) *Exp. Cell Res. 273*, 12−20.
23. Marshall, E. (2002) *Science 298*, 510−511.
24. Lipps, H. J., Jenke, A., Nehlsen, K., Scinteie, M., Stehle, I. M., and Bode, J. (2002) *Gene 304*, 23−33.
25. Bode, J., Goetze, S., Heng, H., Krawetz, S. A., and Benham, C. (2003) *J. Chromosome Res.* (in press).
26. Burgers, W. A., Fuks, F., and Kouzarides, T. (2002) *Trends Genet. 6*, 275−277.
27. Sawasaki, T., Takahashi, M., Goshima, N., and Morikawa, H. (1998) *Gene 218*, 27−35.
28. Boulikas, T. (1995) in *Structural and Functional Organization of the Nuclear Matrix—International Review of Cytology* (Jeon, K. W., and Berezney, R., Eds.) pp 279−388, Academic Press, Orlando, FL.
29. Goetze, S., Gluch, A., Benham, C., and Bode, J. (2003) *Biochemistry 42*, 154−166.
30. Gasser, S. M., and Laemmli, U. K. (1986) *Cell 46*, 521−530.
31. Mirkovitch, J., Mirault, M. E., and Laemmli, U. K. (1984) *Cell 39*, 223−232.
32. Vassetzky, Y., Hair, A., and Mechali, N. (2000) *Genes Dev. 14*, 1541−1552.
33. Berezney, R., and Wei, X. Y. (1998) *J. Cell. Biochem. 238*, 238−242.
34. Marshall, E. (2002) *Science 298*, 34−35.
35. Gai, X. W., and Voytas, D. F. (1998) *Mol. Cell 1*, 1051−1055.

BI0340907