JOURNAL OF VIROLOGY, Dec. 1993, p. 7118–7124
0022-538X/93/127118-07$02.00/0

Vol. 67, No. 12

Amplification and Sequence Analysis of DNA Flanking Integrated Proviruses by a Simple Two-Step Polymerase Chain Reaction Method

ANNETTE B. SØRENSEN, MOGENS DUCH, POUL JØRGENSEN, AND FINN S. PEDERSEN*

Department of Molecular Biology, University of Aarhus, C.F. Møllers Allé, Building 130, DK-8000 Aarhus C, Denmark

Received 1 June 1993/Accepted 8 September 1993

We describe a two-step polymerase chain reaction method that can be used for the amplification of cellular DNA sequences adjacent to an integrated retroviral provirus. The technique involves a partly degenerate, arbitrary primer that will hybridize in the provirus-flanking cellular DNA. By using this primer in combination with a biotinylated provirus-specific primer, a provirus-cellular DNA junction fragment can be isolated from the nonspecific amplification products by using streptavidin-coated magnetic beads. A second amplification employing a nested provirus-specific primer and a biotinylated nondegenerate primer derived from the partly degenerate primer followed by purification with streptavidin-coated beads enhances the specificity and the efficiency of recovery of a fragment(s) containing the unknown flanking sequences. In addition to being relevant in studies of viral integration sites, the method should be generally useful to analyze DNA sequences either upstream or downstream from a known sequence.

In studies of integration sites of retroviral proviruses, for example, in the search for common integration sites in tumor DNAs, the proviral 5′- and 3′-flanking sequences are indispensable, e.g., as probes for further analyses. In these cases, ordinary cloning techniques to purify flanking cellular DNA are very laborious, as no standard or common library can be constructed and used in the analyses. Hence, for studies of proviral integration sites other and faster approaches are desirable.

The polymerase chain reaction (PCR) is a method which permits the amplification of specific DNA sequences without cloning (16, 20). One of the limitations of this technique, however, is the need for knowing the sequence of two specific primers that flank the region to be amplified. Several strategies have been developed to overcome this limitation in situations in which one wants to amplify and sequence unknown DNA flanking a known sequence. These strategies include steps such as ligation and circularization (10, 17, 19), ligation of adaptors to the ends of the amplified sequence (14, 15), PCR with mixed random primers combined with oligomer extension (13), or tailing of specific deoxynucleotides to produce a site to which a primer can hybridize (7, 9, 11). Still, none has found widespread applications in the mapping of integration sites of retroviral proviruses.

This paper describes a simple method without additional or intervening procedures, such as critical ligation or circularization steps, that permits amplification and sequencing of unknown cellular DNA flanking integrated proviruses. The method consists of the following steps: (i) a PCR using a biotinylated provirus-specific primer and a partly degenerate, arbitrary primer (flanking primer) with a fixed 3′ end, which will hybridize within a statistically defined range in the cellular DNA flanking the provirus; (ii) purification of the PCR products that originate from the provirus primer by employing streptavidin-coated magnetic beads; (iii) a second PCR using the purified products from the first PCR as templates and a nested provirus-specific primer together with a biotinylated nondegenerate primer derived from the flanking primer; and (iv) direct sequencing of the purified (on streptavidin-coated magnetic beads) PCR products (5). The sizes of fragments amplified by this technique have been up to approximately 1,500 bp.

We have documented the validity of the method by amplifying and sequencing DNA flanking an inherited ecotropic provirus in BALB/c mice, single-copy integrated retroviral vector proviruses, and finally, integrated SL3-3 proviruses in tumor DNAs from SL3-3-infected NMRI mice.

MATERIALS AND METHODS

PCRs. All PCRs were done in a 50-μl reaction volume containing 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 0.01% (wt/vol) gelatin, 0.2 mM (each) deoxynucleoside triphosphates, 25 pmol of the specific primer and 75 pmol of the random primer (first PCR [Fig. 1]) or 25 pmol of each primer (second PCR [Fig. 1] and shifting PCR [as defined in Results and Discussion]), and 1.2 U of *Taq* DNA polymerase (Stratagene). Genomic or plasmid DNA (100 to 1,000 ng) was used as a template in the first PCR. For the second PCR (or for the shifting PCR) 2 μl of the isolated fragments from the first PCR was used as a template (see below). The amplification mixture was overlaid with a drop of paraffin oil and underwent 40 cycles of PCR in a water-based thermal cycler (Microlab, Aarhus, Denmark) with denaturation at 94°C for 1.2 min (first cycle, 94°C for 4 min), annealing at 62°C for 1.2 min, and extension at 73°C for 3 min (last cycle, 73°C for 10 min). For analysis, 10 μl of the amplification reaction mixture was electrophoresed on a 2% (wt/vol) agarose gel and stained with ethidium bromide.

Dynabead-streptavidin purification. After the first PCR the amplified biotinylated fragments were isolated (by using a magnet to sediment the beads in the tubes during supernatant removal and washing procedures) by mixing 40 μl of the PCR mixture with 40 μl of 200-μg prewashed (in 2× B&W buffer [10 mM Tris-HCl, pH 7.5, 1 mM EDTA, 2.0 M NaCl])

* Corresponding author. Electronic mail address: fsp@mbio.aau.dk.

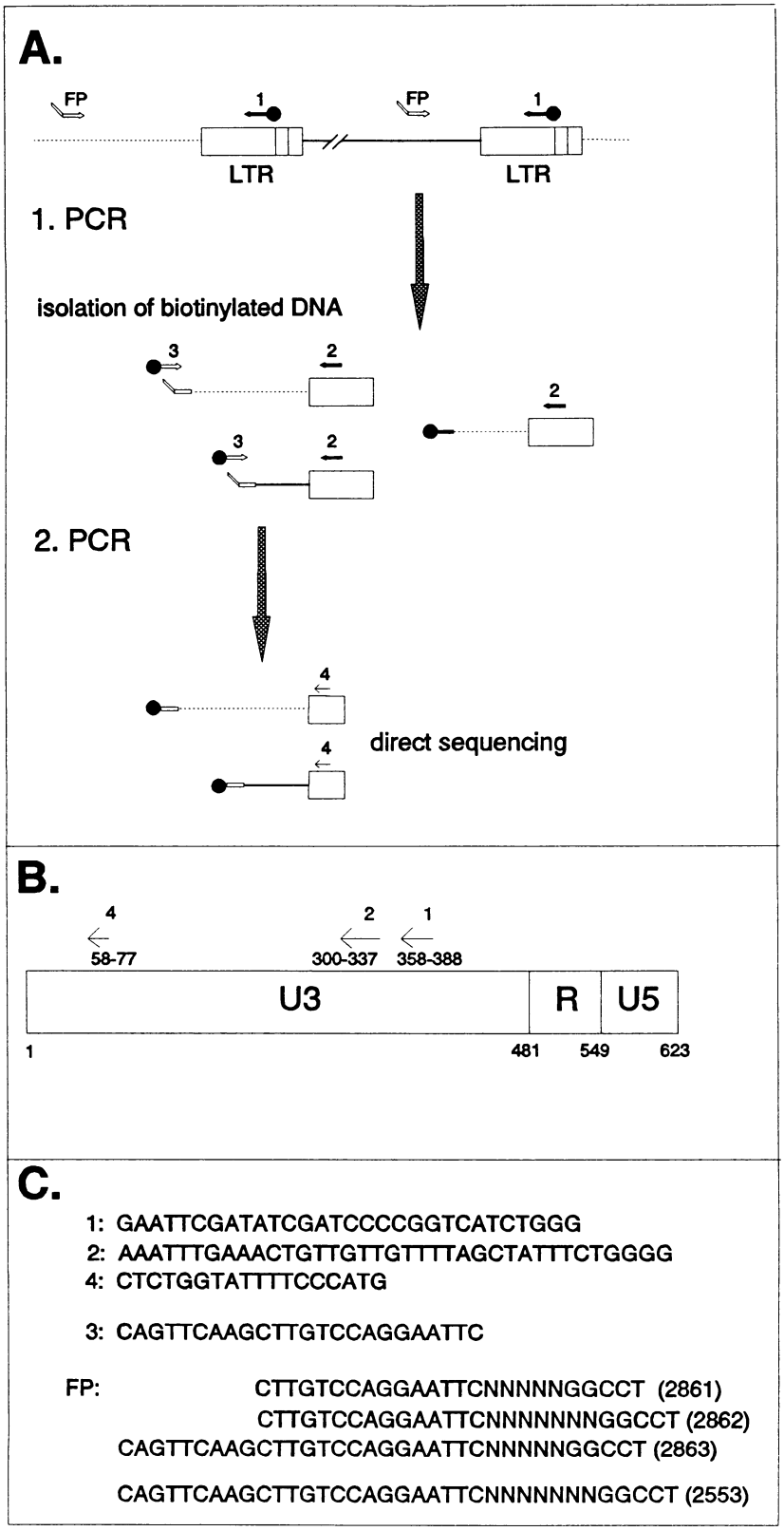

FIG. 1. Strategy for amplifying unknown DNA flanking integrated proviruses. (A) The integrated provirus is indicated by the two boxes (the LTRs) and the solid line. The dotted lines represent the flanking DNA. The arrows denote the oligonucleotide primers (FP, flanking primer; 1, 2, and 3, specific PCR primers; 4, sequencing primer). A filled circle indicates biotin coupled to the 5′ end of the DNA. (B) SL3-3 LTR region (8) showing the positions of the provirus-specific primers used in panel A. (C) Sequences of the oligonucleotide primers used in panel A. Primer 1 is an ecotropic specific primer (15). N, AGCT.

Dynabead M280-streptavidin (Dynal AS, Oslo, Norway). Following incubation for 15 min and washing in 1× B&W buffer the Dynabead-bound fragments were melted in 8 μl of 0.1 M NaOH for 10 min. The supernatant containing the nonbiotinylated strands was neutralized with 4 μl of 0.2 M HCl and 1 μl of 1 M Tris-HCl, pH 8.0, and the volume was adjusted with water to 30 μl, of which 2 μl was used as a template in the second PCR.

Sequencing of the PCR fragments. Isolation of the biotinylated fragments from the second PCR followed the same procedure as the isolation of fragments from the first PCR, except that after separating of the strands the immobilized biotinylated strands were washed once with 50 μl of 0.1 M NaOH, once with 40 μl of 1× B&W buffer, and once with 50 μl of 10 mM Tris-HCl, pH 7.5–1 mM EDTA buffer. The biotinylated strands were resuspended in 7 μl of H_2O and sequenced with primer 4 (Fig. 1) by using either (i) a Sequenase kit per the manufacturer's recommendations (United States Biochemical Corp., Cleveland, Ohio) followed by sedimentation of the Dynabeads and loading the supernatant onto the sequencing gel or (ii) a *Taq* DyeDeoxy Terminator Cycle Sequencing Kit (Applied Biosystems, Inc., Foster City, Calif.); the manufacturer's recommendations were followed except for employing a Hybaid OmniGene thermal cycler and denaturing at 96°C for 35 s, annealing at 42°C for 15 s, and polymerization at 58°C for 4 min (25 cycles total). The phenol-chloroform protocol was used for extractions of excess DyeDeoxy Terminators from the completed sequencing reactions before loading the samples onto an Applied Biosystems 373A DNA sequencer according to supplied instructions.

Preparation of oligonucleotides. Oligonucleotides were synthesized (DNA Technology ApS, Aarhus, Denmark) on an Applied Biosystems 381B DNA synthesizer and purified on Oligo Purification Cartridges.

RESULTS AND DISCUSSION

The overall strategy for amplifying unknown flanking DNA is illustrated in Fig. 1. The procedure allows direct amplification of cellular DNA adjacent to an integrated provirus without a priori knowledge of sequence information beyond the provirus priming sites. Cellular DNA containing the provirus is amplified in two steps, using in the first PCR step a flanking primer. This primer was chosen to consist of a 3′ end of 5 fixed nucleotides, which on average should exist every 4^5 (i.e., 1,024) bp in the genome; in this way, the amplification product size will generally be limited within the range for the PCR to take place. In order to increase the temperature of annealing in the PCRs the fixed 3′ end should be rich with G and/or C, but at the same time, to avoid unspecific hybridization to GC-rich regions the very 3′ nucleotide must be A or T. GGCCT was chosen because this sequence is located at a suitable distance in the recombinant plasmid which was used in preliminary testing experiments. Next to the fixed 5 nucleotides are 5 or 7 random nucleotides which again are followed by 15 or 24 nucleotides corresponding to different restriction sites. This flanking primer is used together with a provirus-specific primer that is complementary to a portion of the U3 region of the long terminal repeat (LTR) in which the sequence of ecotropic viruses differs from that of other known endogenous retroviruses (17). This ecotropic specific primer is biotinylated in order to eliminate all or most of the fragments amplified by the flanking primer alone, by using the system of streptavidin-coated magnetic beads. The resulting purified fragments are used as templates in the second PCR with a nested provirus-specific primer and a biotinylated primer corresponding to the nondegenerate 5′ part of the flanking primer. The isolated biotinylated fragments resulting from this PCR are directly sequenced.

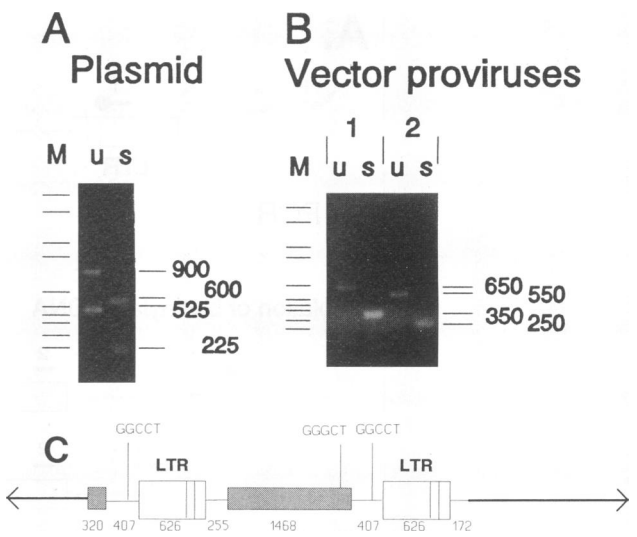

FIG. 2. Amplification of sequences flanking proviral 5′ LTRs. Ethidium bromide-stained agarose gels with PCR products from the second (unshifted) PCR (lanes u) or from the shifting PCR (lanes s). M, molecular size markers (2,176, 1,766, 1,230, 1,033, 653, 517, 453, 394, 298, and 234 or 220 bp, respectively). (A) Amplification of plasmid DNA, ptvAkv-neo. (Flanking primer 2553 was used in the first PCR.) (B) Amplification of genomic DNA from tvAkv-neo-transduced L691 cells. 1 and 2, two different cell clones. (A mixture of flanking primers 2861, 2862, and 2863 [25 pmol of each] was used in the first PCR.) (C) Structure of ptvAkv-neo. Heavy lines denote the pBR322 plasmid. Open boxes represent Akv LTR sequences (626 bp). Hatched boxes (320 and 1,468 bp) represent the *neo* gene with surrounding regions of Tn*5* (left box [320 bp], only Tn*5* sequences). The thin lines denote Akv sequences (407, a 407-bp fragment immediately upstream of the Akv 3′ LTR; 255 [or 172], a 255-bp [or a 172-bp] fragment immediately downstream of the Akv 5′ LTR). Indicated by the vertical lines are the sites to which the flanking primer apparently hybridizes.

In initial studies recombinant plasmid DNA (ptvAkv-neo; Fig. 2C) (6, 12) was amplified and purified as shown in Fig. 1. After the two rounds of PCR with the intervening Dynabead-streptavidin purification, two fragments were specifically amplified (Fig. 2A, lane u): one of the predicted size of 526 bp arising from the flanking primer binding to GGCCT located 159 to 155 bp upstream both LTRs and one of approximately 900 bp. This band may correspond to a fragment amplified from the flanking primer binding to the GC-rich region CGCCCACCCCGGGCT, with the 3′ end, GGCCT, of the flanking primer presumably binding to GGGCT located 557 to 553 bp upstream the 3′ LTR.

In order to exclude the possibility that the bands seen after the second PCR correspond to fragments amplified from the flanking primer alone, another amplification was performed with another ecotropic specific primer (17) complementary to the sequence just before the inverted repeat at the beginning of the LTR (bases 12 to 23; 1 = the first base in U3). By using this primer instead of primer 2 (Fig. 1) in an otherwise similar second PCR, the fragments from such a PCR (shifting PCR) should be 301 bp shorter. Besides, the shifting PCR further strengthens the supposition of the ecotropic character of the amplified fragments. Figure 2A, lane s, shows that the two bands shifted correctly.

We next tried to amplify sequences adjacent to the ecotropic

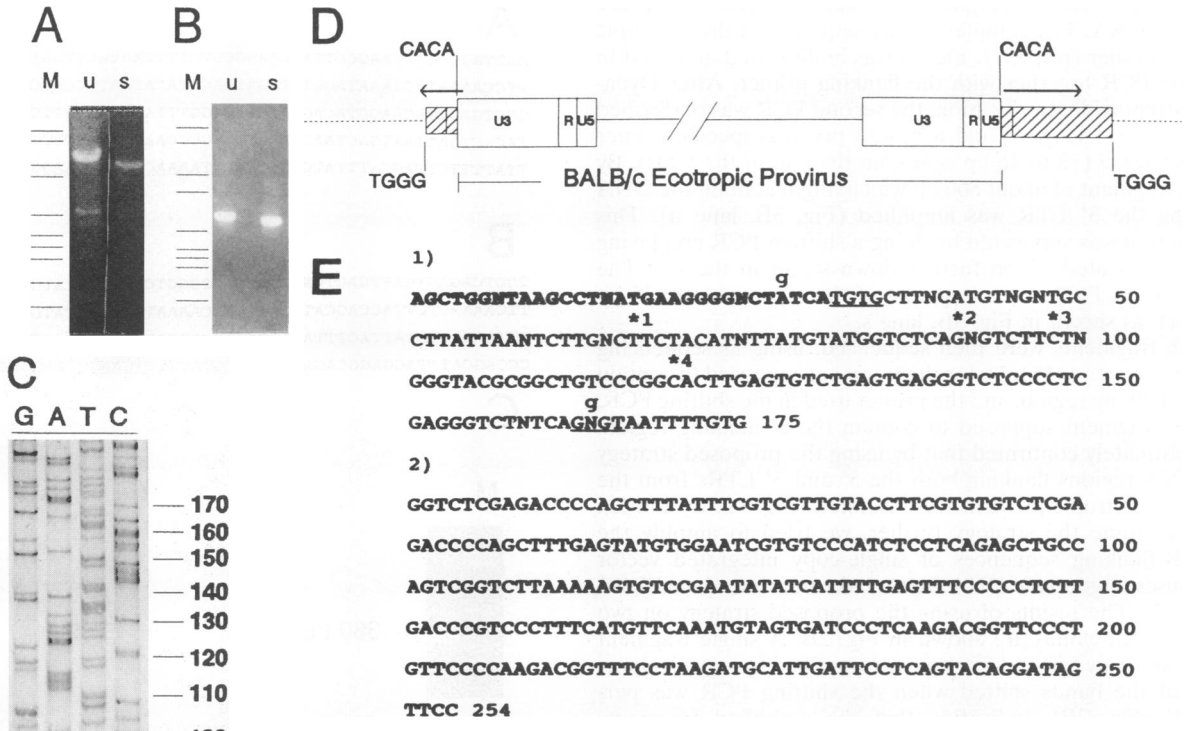

FIG. 3. Amplification of the sequences flanking the ecotropic proviral LTRs in BALB/c DNA. (A and B) Ethidium bromide-stained agarose gels showing the fragments amplified in the second PCR (lanes u) and in the shifting PCR (lanes s). M, molecular size markers (2,176, 1,766, 1,230, 1,033, 653, 517, 453, 394, 298, and 234 or 220 bp, respectively). (The first PCRs were done with a mixture of flanking primers 2861, 2862, and 2863 [25 pmol of each]). (A) Fragments amplified from the 5′ LTR. Before the first PCR the BALB/c DNA was cut with *Afl*II to eliminate amplification from the 3′ LTR (see text). (B) Fragments amplified from the 3′ LTR. (C) Results of direct sequencing of the PCR fragment shown in panel B, lane u (3′-LTR-flanking region), with the shifting primer as the sequencing primer. Numbers correspond to numbers in the sequence shown in panel E, sequence 2. (D) Schematic illustration of the integration site of BALB/c endogenous ecotropic provirus showing the extent of sequence determination, represented by the arrows, of the amplified 5′- and 3′-LTR-flanking regions. The open boxes indicate BALB/c ecotropic LTRs (U3, R, and U5), while the hatched boxes represent the flanking VL30 LTR-related sequences. The dotted lines denote flanking cellular DNA. The repeated sequences at the junctions of ecotropic and VL30 sequences and at the junctions of VL30 and cellular sequences are indicated (3). (E) Sequences of the 5′- and 3′-LTR-flanking regions. (Sequence 1) The sequence of the 5′-LTR-flanking region was determined by automatic sequencing with a *Taq* DyeDeoxy Terminator Cycle Sequencing Kit, and the sequence shown here is a computer interpretation of the peaks of the chromatogram. The boldfaced sequence denotes the ecotropic LTR sequence upstream the sequencing primer. The repeated sequences at the junctions (ecotropic-VL30 and VL30-cellular sequences) are underlined. An asterisk above the sequence indicates a difference from the sequence published by Horowitz et al. (3) (*1, TT/T–; *2, G/–; *3, C–T/CAT; *4, GGC/GGG; the bases following the slash are from the sequence of Horowitz et al. [3], and dashes indicate no base). A "g" above the sequence indicates that a G (instead of A and N, respectively) may be interpreted from the chromatogram. The undetermined bases, represented by N, may in all cases be replaced by the bases from the sequence determined by Horowitz and al. (3) if manual interpreting of the peaks is performed. (Sequence 2) Sequence of the 3′-LTR-flanking region, of which a part is shown in panel C. This sequence is identical to the one of Horowitz et al. (3).

proviral LTRs in BALB/c genomic DNA. BALB/c is a mouse strain that contains a single ecotropic provirus located on mouse chromosome 5; the 5′- and 3′-flanking sequences of this provirus have been determined (3, 4). Figure 3D shows a schematic illustration of the integration site. The BALB/c ecotropic provirus is integrated within the R region of a VL30 LTR-related sequence, in opposite transcriptional orientation. In this case we wanted to amplify the sequences adjoining both the 5′ and the 3′ LTRs. The 5′-flanking sequence was amplified as shown in Fig. 1. However, in order to avoid amplifying internal proviral sequences from the 3′ LTR (see below), the DNA was digested with *Afl*II before the first PCR was performed. Since *Afl*II cuts immediately upstream the 3′ LTR, this digestion should eliminate amplification of internal provirus fragments. Figure 3A, lane u, shows that a fragment of about 1,500 bp was amplified. The shifting PCR, which was done as described above, resulted in a fragment of the expected size of about 1,200 bp (Fig. 3A, lane s), indicating that the 1,500-bp fragment contained the DNA region (about 1,200 bp) adjacent to the 5′ LTR.

In principle, the strategy shown in Fig. 1 should also be applicable to isolate unknown DNA flanking the 3′ LTR by using the complementary sequence of the primers shown in Fig. 1. The amplification will then be directed from the 3′ LTR

out into the 3′-flanking regions. This was examined on BALB/c genomic DNA. The complementary sequence of the ecotropic specific primer (primer 1, Fig. 1) was biotinylated and used in the first PCR together with the flanking primer. After Dynabead-streptavidin purification, the second PCR was performed with primer 3 (Fig. 1) and a nested provirus-specific primer located in U5 (73 to 46 bp upstream the end of the LTR). By this, a fragment of about 800 bp which might contain the DNA flanking the 3′ LTR was amplified (Fig. 3B, lane u). This assumption was supported by doing a shifting PCR employing a primer located 35 bp further downstream in the U5. The result of this PCR was a fragment of the expected size (35 bp shorter), as shown in Fig. 3B, lane s.

Both fragments were then sequenced, using as sequencing primers primer 4 (Fig. 1), for the fragment supposed to contain the 5′-flanking region, and the primer used in the shifting PCR, for the fragment supposed to contain the 3′-flanking region. This ultimately confirmed that by using the proposed strategy the DNA regions flanking both the 5′ and 3′ LTRs from the BALB/c ecotropic provirus were isolated (Fig. 3C, D, and E).

To examine the strategy further, we tried to amplify the 5′-LTR-flanking sequences of single-copy integrated vector proviruses in genomic DNA from tvAkv-neo-transduced L691 cells (12). The results of using the proposed strategy on two different cell clones are shown in Fig. 2B. A single fragment (approximately 650 and 550 bp) was amplified from each clone. Both of the bands shifted when the shifting PCR was performed (Fig. 2B), indicating that the amplified fragments contained the flanking sequences. This was confirmed by sequencing the fragments. The PCR fragments contained the sequence upstream of the sequencing primer (primer 4, Fig. 1) followed by unknown flanking sequences. Moreover, 3′-LTR-flanking regions from the same integrated proviral sequences have been isolated, and analyses of the 5′- and 3′-flanking regions belonging together showed the expected 4-bp repeat sequence of the sites of integrations (Fig. 4A and B). In addition, PCR amplifications of preintegration sites in tvAkv-neo-infected L691 cells with primers of sequences obtained from the flanking DNA showed fragments of the expected sizes, strongly signifying that the amplification products from the tvAkv-neo-transduced cells did not originate from endogenous LTRs (Fig. 4C).

We are currently using the method to identify the integration sites in murine tumors induced by the SL3-3 retrovirus and in cell lines with single-copy integrated retroviral vector proviruses. We have isolated and sequenced about 30 LTR 5′-flanking DNAs from different SL3-3-induced tumors (2, 18). Genomic DNAs from the tumors were isolated, and sequences flanking the integrated proviruses were amplified by the proposed strategy. Figure 5 shows an example of the resulting fragments following the two rounds of PCR, employed with seven different tumor DNAs. Shifting PCR was performed, and as seen in Fig. 5, most of the bands shifted correctly, indicating that the amplification products contained flanking DNA. The products from the second PCR were directly sequenced, and the fragments contained true flanking sequences only in four cases (lanes 1, 2, 5, and 7 in Fig. 5). Direct sequencing of the remaining fragments (lanes 3, 4, and 6) revealed that they contained internal proviral sequences. Using a provirus-specific primer located in the LTR region in the first round of PCR means that the specific primer will bind in the other LTR as well, and if the flanking primer binds to a sequence internal in the provirus, internal fragments will be amplified (Fig. 1). Bands corresponding to any such resulting fragments will shift, as will the bands corresponding to fragments containing true flanking sequences, in the shifting PCR.

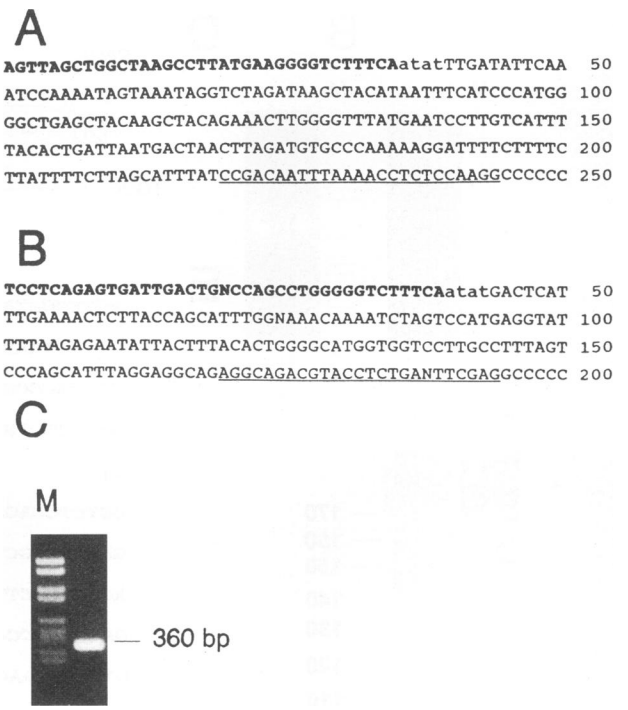

FIG. 4. Amplification of a preintegration site. (A and B) Sequences of 5′- and 3′-LTR-flanking DNA, respectively, related to the same Akv-*neo* integrated provirus. The boldfaced sequences denote provirus sequences, U3 in panel A and U5 in panel B. Underlined are the sequences to which the primers used for amplification of the preintegration site hybridize. Lowercase letters indicate the 4-bp repeats. (C) Ethidium bromide-stained agarose gel showing the amplification of the preintegration site in ptvAkv-neo-infected cells by using primers of sequences obtained from the flanking DNA shown in panels A and B. Lane M, molecular size markers (2,176, 1,766, 1,230, 1,033, 653, 517, 453, 394, 298, and 234 or 220 bp, respectively).

In fact, of the above-mentioned 30 LTR 5′-flanking DNA sequences, 14 turned out to be not true flanking sequences but internal proviral sequences. These sequences, however, have illustrated some important aspects of the strategy, the first of which seems to be how the flanking primer compares to the genomic DNA: does it recognize the 5-nucleotide match? All the internal fragments except one are of sizes that fit with a perfect GGCCT match. Three sizes of fragments, estimated to about 550 bp (two fragments), 800 bp (8 fragments), and 1,300 to 1,350 bp (three fragments), have been observed. These sizes are consistent with the occurrence of exactly three GGCCT sequences located in fitting positions in the region of interest, the *env* gene (using the available sequence data on SL3-3 [reference 8 and unpublished results] and the closely related Akv [1]). The size of the one fragment that does not fit with a GGCCT site is about 450 bp. The sequence of this fragment shows that the flanking primer in this case had bound to a GGCCCT site located 114 to 119 bp upstream from the 3′ LTR. One explanation could be that in this particular provirus the site mutated from GGCCCT to GGCCT (or GGGCCT), thereby creating the correct flanking primer binding site. Alternatively, a mismatch in the hybridization of the 3′ end of the flanking primer to GGCCCT might have been allowed in this one case. Anyway, these observations indicate that during the first PCR the flanking primer generally recognizes a perfect 5-nucleotide match.

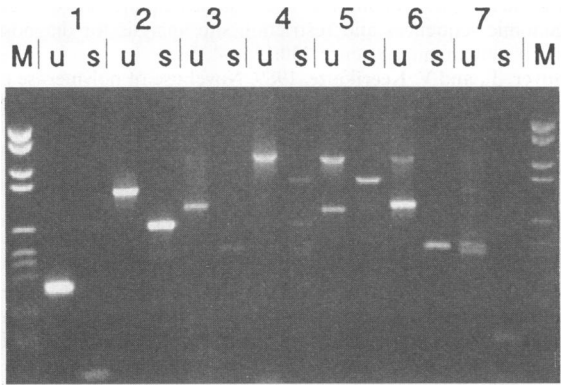

FIG. 5. Amplification of proviral LTR 5′-flanking sequences in tumor DNAs from SL3-3-infected mice. Ethidium bromide-stained agarose gel with PCR products from the second PCR (lanes u, unshifted) or the shifting PCR (lanes s). The first PCRs were done with a mixture of flanking primers 2861, 2862, and 2863 (25 pmol of each). 1 to 7 correspond to DNA from seven different tumors. Lanes M, molecular size markers (2,176, 1,766, 1,230, 1,033, 653, 517, 453, 394, 298, and 234 or 220 bp, respectively).

Sometimes more than one fragment is amplified, which may be due to more than one integrated provirus. Nevertheless, the method described here still works in these cases, with some minor modifications, such as separating the different amplified fragments after the second PCR before sequencing. However, we have observed that this may not always be necessary, because the different bands seen after the second PCR may as well be amplification products originating from the same integrated provirus, with the flanking primer binding more than one site in the flanking DNA, and hence, more than one fragment can be amplified from the same provirus. This will of course not influence the direct sequencing.

Examination of the internal sequences also indicates competition in the PCRs. Only one (in a few cases, two) internal fragment is amplified to significant amounts in the same PCR, suggesting competition, but what fragment(s) will initially be selected for amplification cannot be predicted. It seems that there is a slight preference for the fragment of about 800 bp, but the overall homology between the flanking primer and the proviral DNA in this region does not suggest such a preference. Perhaps the size of the fragment to be amplified plays a role. The implication of competitive PCR also explains why the expected two fragments (one from either LTR) are not generally observed. The amplification of either one (the flanking DNA) or the other (the internal DNA) is inhibited in the PCR. Yet this problem should be eliminated by cutting the DNA with a suitable restriction enzyme before the first PCR to remove either one of the two LTRs, provided that the viral sequence is known and bearing in mind the risk of cutting off the desired flanking DNA. We did such a digestion when the DNA adjacent to the ecotropic 5′ LTR in BALB/c DNA was to be isolated, because we had had problems in previous experiments that otherwise only internal sequences could be amplified in this case. Alternatively, in order to avoid the undesirable internal sequences, a provirus-specific primer located outside the LTR could be chosen for the first PCR. Yet this would result in a larger fragment to be amplified before reaching the unknown, flanking region.

Occasionally, we are not able to amplify provirus-flanking cellular coherent fragments using our flanking primer that contains GGCCT in its 3′ end. This may in some cases be caused by structural features of the flanking DNA making it difficult to amplify, but it may also be due to a distant location of the nearest GGCCT site; that is, the flanking primer does not hybridize to the flanking region within a suitable range for the PCR to take place. Generally, the average product size is limited to under 4 kb. In theory, this problem should be overcome by using flanking primers containing other nucleotides in their 3′ end. In the case of analyzing proviral integration sites in tumors it will be of importance to amplify flanking cellular DNA from all integrated proviruses in a given tumor as detected in Southern blots. For this purpose, it may also be necessary to employ flanking primers with different 3′ ends.

Competitive PCR may be considered a problem when amplifying tumor DNA with more than one provirus, but by employing the above-mentioned set of flanking primers with different 3′ ends, this problem should be overcome.

We find that our method in general can be successfully used in studies of integration sites of retroviruses, which otherwise often are very laborious. Besides, the proposed strategy could be applied in many other situations. For example, we have amplified genomic regulatory regions adjacent to cDNA (data not shown). Also, the method should be relevant in exon-intron junctions analyses and in small-scale chromosome walking.

ACKNOWLEDGMENTS

This work was supported by CEC-BRIDGE contract BIOT-CT91-0286, by contract FI3P-CT920051 of the CEC Radiation Protection Research Action, and by the Danish Biotechnology Programme, the Danish Cancer Society, the Novo Foundation, the Ib Henriksen Foundation, and the Leo Nielsen Foundation.

REFERENCES

1. **Etzerodt, M., T. Mikkelsen, F. S. Pedersen, N. O. Kjeldgaard, and P. Jørgensen.** 1984. The nucleotide sequence of the Akv murine leukemia virus genome. Virology **134:**196–207.
2. **Hallberg, B., J. Schmidt, A. Luz, F. S. Pedersen, and T. Grundström.** 1991. SL3-3 enhancer factor 1 transcriptional activators are required for tumor formation by SL3-3 murine leukemia virus. J. Virol. **65:**4177–4181.
3. **Horowitz, J. M., G. D. Holland, S. R. King, and R. Risser.** 1987. Germ line integration of a murine leukemia provirus into a retroviruslike sequence. J. Virol. **61:**701–707.
4. **Horowitz, J. M., and R. Risser.** 1985. Molecular and biological characterization of the endogenous ecotropic provirus of BALB/c mice. J. Virol. **56:**798–806.
5. **Hultman, T., S. Ståhl, E. Hornes, and M. Uhlén.** 1989. Direct solid phase sequencing of genomic and plasmid DNA using magnetic beads as solid support. Nucleic Acids Res. **17:**4937–4946.
6. **Jensen, N. A., P. Jørgensen, N. O. Kjeldgaard, and F. S. Pedersen.** 1986. Mammalian expression-and-transmission vector derived from Akv murine leukemia virus. Gene **41:**59–65.
7. **Jones, D. H., and S. C. Winistorfer.** 1992. Sequence specific generation of a DNA panhandle permits amplification of unknown flanking DNA. Nucleic Acids Res. **20:**595–600.
8. **Lenz, J., D. Celander, R. L. Crowther, R. Patarca, D. W. Perkins, and W. A. Haseltine.** 1984. Determination of the leukaemogenicity of a murine retrovirus by sequences within the long terminal repeat. Nature (London) **308:**467–470.
9. **Loh, E. Y., J. F. Elliott, S. Cwirla, L. L. Lanier, and M. M. Davis.** 1989. Polymerase chain reaction with single-sided specificity: analysis of T cell receptor delta chain. Science **243:**217–220.
10. **Ochman, H., M. M. Medhora, D. Garza, and D. L. Hartl.** 1990. Amplification of flanking sequences by inverse PCR, p. 219–227. *In* M. Innis et al. (ed.), PCR protocols: A guide to methods and application. Academic Press, Inc., New York.
11. **Ohara, O., R. L. Dorit, and W. Gilbert.** 1989. One-sided polymerase chain reaction: the amplification of cDNA. Proc. Natl. Acad. Sci. USA **86:**5673–5677.

12. **Paludan, K., H. Y. Dai, M. Duch, P. Jørgensen, N. O. Kjeldgaard, and F. S. Pedersen.** 1989. Different relative expression from two murine leukemia virus long terminal repeats in unintegrated transfected DNA and in integrated retroviral vector proviruses. J. Virol. **63:**5201–5207.
13. **Parker, J. D., P. S. Rabinovitch, and G. C. Burmer.** 1991. Targeted gene walking polymerase chain reaction. Nucleic Acids Res. **19:**3055–3060.
14. **Riley, J., R. Butler, D. Ogilvie, R. Finniear, D. Jenner, S. Powell, R. Anand, J. C. Smith, and A. F. Markham.** 1990. A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones. Nucleic Acids Res. **18:**2887–2890.
15. **Rosenthal, A., and D. S. C. Jones.** 1990. Genomic walking and sequencing by oligo-cassette mediated polymerase chain reaction. Nucleic Acids Res. **18:**3095–3096.
16. **Saiki, R. K., S. Scharf, F. Faloona, K. B. Mullis, G. T. Horn, H. A. Erlich, and N. Arnheim.** 1985. Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science **230:**1350–1354.
17. **Silver, J., and V. Keerikatte.** 1989. Novel use of polymerase chain reaction to amplify cellular DNA adjacent to an integrated provirus. J. Virol. **63:**1924–1928.
18. **Thornell, A., B. Hallberg, and T. Grundström.** 1991. Binding of SL3-3 enhancer factor 1 transcriptional activators to viral and chromosomal enhancer sequences. J. Virol. **65:**42–50.
19. **Triglia, T., M. G. Peterson, and D. J. Kemp.** 1988. A procedure for *in vitro* amplification of DNA segments that lie outside the boundaries of known sequences. Nucleic Acids Res. **16:**8186.
20. **Wong, C., C. E. Dowling, R. K. Saiki, R. G. Higuchi, H. A. Erlich, and H. H. Kazazian, Jr.** 1987. Characterization of β-thalassaemia mutations using direct genomic sequencing of amplified single copy DNA. Nature (London) **330:**384–386.