JOURNAL OF VIROLOGY, Aug. 2005, p. 10093–10096 Vol. 79, No. 15
0022-538X/05/$08.00+0 doi:10.1128/JVI.79.15.10093–10096.2005

A New Common Integration Site, Int7, for the Mouse Mammary Tumor Virus in Mouse Mammary Tumors Identifies a Gene Whose Product Has Furin-Like and Thrombospondin-Like Sequences

William Lowther, Korah Wiley, Gilbert H. Smith, and Robert Callahan*

Mammary Biology and Tumorigenesis Laboratory, National Cancer Institute, Bethesda, Maryland 20892

Received 12 January 2005/Accepted 22 April 2005

A novel common integration site for the mouse mammary tumor virus (MMTV) was identified (designated Int7) in five independently arising mouse mammary tumors. The insertion sites all cluster within a 1-kb region that is 2 to 3 kb 5′ of the transcription initiation site of a gene, 2610028F08RIK, whose gene product contains furin-like and thrombospondin-like sequences. Expression of Int7 is normally very low or silent during various stages of mammary gland development, but MMTV integration at this site results in the activation of high steady-state levels of expression of the gene. These five tumors were also found to have two or three additional viral insertions, which in each case occurred flanking a member of either the Wnt and/or FGF gene family. Reverse transcriptase PCR results demonstrated that each of the viral insertions led to elevated expression of the presumed target flanking genes.

The mouse mammary tumor virus (MMTV) induces pregnancy-independent mammary adenocarcinomas in the Czech II strain of feral *Mus musculus musculus* with a predictable (20%) frequency after 15 months of age (reviewed in reference 3). These mice also develop preneoplastic hyperplastic nodules that can be transplanted as hyperplastic outgrowth lines. Frequently, mammary tumors will develop from within these outgrowths, and some mice also develop metastatic lesions to the lungs. Unlike other inbred strains of mice, such as C3H and RIII, which have been bred to have a 100% mammary tumor incidence among parous females, with an average latency of 250 days, Czech II mice have not undergone selection for increased tumor frequency or shortened latency. Furthermore, Czech II mice have no endogenous MMTV proviral genomes, while most other inbred strains carry between two and eight. These two aspects of Czech II mice make it an attractive model in which to screen for new common integration sites (CIS) involved with mammary tumorigenesis. Both Int3, the activated intracellular domain of the Notch-4 receptor, and Int6, the p48 component of the eukaryotic translation initiation factor 3 complex (eIF3p48), were originally isolated as retroviral tags in mammary tumors from Czech II mice (5, 10).

We have surveyed additional Czech II mammary tumors in search of new CIS for MMTV. Using the inverse PCR approach (14), we have identified a novel CIS from a panel of 40 independently arising mammary tumors. Tumor DNA was first digested overnight by using a cocktail of BamHI, BglII, and BclI restriction enzymes (Roche, Indianapolis, IN). Digested DNA (20 ng) was then self ligated in a total volume of 200 μl by using T4 DNA ligase (high concentration) (Roche) at 16°C overnight. Template DNA (2 μl) was added to PCRs (92°C for 3 min 20 s, 65°C for 25 s, and 70°C for 2.5 min for 30 cycles and 70°C for 10 min) by using primer set MMTV ltr5-100 (5′CGCGTGCACGCAGACGGGTCGTC CTTGG3′) and Gag-2720 (5′CCTCCTGGAGTTAAAAAG ACTGTATTAGC3′) or MMTV ltr3-9740 (5′CTTGCAAC AGTCCTAACATTCGTCTCTCG3′) and Env-8380 (5′CC AATCTAATGGATTTAACGCCTTCACTCC3′). A portion (2 μl) of this reaction mixture was then reamplified by using the same reaction conditions with primer set MMTV ltr5-40 (5′CCTAAGTGTAGGACACTCTCGGGAGTTC3′) and Gag-2800 (5′CATTTCAAGGCTCGAGGAAGCTGTTTACAG3′) or MMTV ltr3-9780 (5′GCCATCCCGTCTCCGCTCGTCACT TATCC3′) and Env-8360 (5′CACTCCATTGGCAAAGGACT GAGCCAAACC3′). The resulting products were separated on a 1% agarose gel, and individual bands were eluted and their nucleotide sequence directly determined. Nucleotide sequence data were compared by BLASTn with nonredundant and expressed sequence tag (EST) databases to determine genomic sites of integration and relative position to flanking genes.

A new CIS (2610028F08Rik) was identified for MMTV in five independent tumors isolated from various stages of tumor development (Fig. 1). Interestingly, it has not been found to be a CIS target for murine leukemia virus in several large-scale surveys of virus-induced mouse leukemia and lymphomas (1, 4, 6, 13). The preneoplastic hyperplastic outgrowth (HOG) line CZZ26 represented the earliest mammary tumor stage in which the gene was rearranged (Table 1, sample 676). In addition, the gene was also rearranged in a CZZ26-derived mammary tumor and its associated lung metastasis. This suggests that MMTV integration at this site is an early event in the evolution of the tumor. Similarly, the gene was also rearranged in the HOG CZZ28-derived mammary tumor (sample 649) and lung metastasis (sample 641). Unfortunately, CZZ28 HOG DNA was unavailable for testing to determine whether it contained an MMTV genome integrated at this site. All of the integration sites at 2610028F08Rik occurred within a 2- to

* Corresponding author. Mailing address: National Cancer Institute, Building 37/Room 1118A, Bethesda, MD 20892. Phone: (301) 496-9871. Fax: (301) 402-0711. E-mail: rc54d@nih.gov.

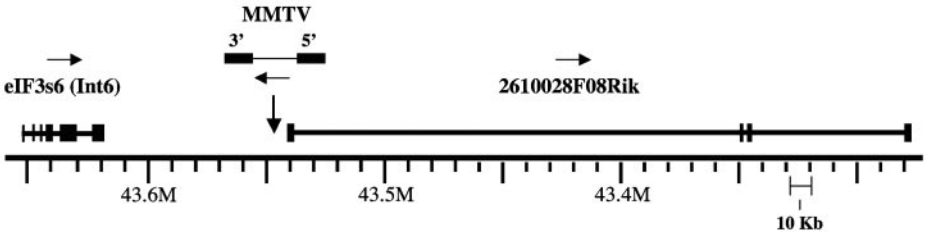

FIG. 1. The 2610028F08Rik (Int7/FLTL) locus. The CIS for MMTV (vertical arrow) in the 2610028F08Rik (Int7/FLTL) locus is adjacent to another CIS for MMTV, eIF3s6 (Int6), on mouse chromosome 15. The transcriptional orientations of the genes are indicated with horizontal arrows. The locations of the genes on chromosome 15 are indicated in millions (M) of nucleotides from the centromere. The boxes indicate exons or clusters of closely linked exons. The distance between each hatch mark is 10,000 nucleotides.

3-kb region of the genome that is 5′ of the transcription initiation site (a TATA-less promoter) of the gene. In each case, the transcriptional orientation of the integrated viral genome was in the opposite direction with respect to the 2610028F08Rik (designated Int7) transcription promoter (Fig. 1A and Table 1).

Int7 contains eight exons spanning over 250 kb of genomic DNA located on chromosome 15 and is positioned next to eIF3e/Int6 in the same transcriptional orientation (Fig. 1). The transcribed message is 3,342 bp. We have cloned and determined the nucleotide sequence of this transcript, and our sequence agrees with the GenBank sequence. The largest open reading frame of Int7 begins with a start codon at position 878 in the second exon and encodes a protein having 243 amino acid residues. Amino acid sequence analysis suggests that the Int7 protein contains a nuclear localization signal near the C′ terminus, as well as a furin-like domain and a thrombospondin-like domain (Fig. 2). We have tentatively renamed the gene Int7/FLTL (furin-like thrombospondin-like) to reflect the presence of these two domains in the protein. At the present time, it is only possible to speculate on the function of the gene product. By use of the Gene Ontology programs (Mouse Genome Informatics [http://www.informatics.Jax.org]), it could be inferred from the sequence or structural similarity that the Int7/FLTL protein may be a transmembrane receptor protein tyrosine kinase, a portion of which could be transferred to the nucleus (11).

Int7/FLTL product is ubiquitously expressed at low levels in several adult tissues (Fig. 3A) and early in mammary gland development (Fig. 3B) but is not detectable in the mammary glands of day 15 pregnant mice (Fig. 3C) or lactating or involuting mice (Fig. 3B). Whereas tumors having a viral insertion at Int7/FLTL express high steady-state levels of Int7/FLTL RNA, tumors in which the gene has not been rearranged by the virus express no detectable levels of Int7/FLTL RNA (Fig. 3B). In addition, the FLTL transcript has been identified in several cDNA libraries, including a library generated from tumors metastasizing to the mammary gland (Unigene; National Center for Biotechnology Information). To begin to assess the function of Int7/FLTL, we have expressed FLTL cDNA in the HC11 mouse mammary epithelial cell line (2). This led to no discernible morphological changes in the cells, and its expression did not confer the capacity for anchorage-independent soft agar growth on the cells (data not shown). This appears not to be a consequence of significantly lower levels of Int7/FLTL RNA in HC11-Int7 compared to a tumor in which Int7/FLTL has been rearranged by MMTV (Fig. 3D, compare lanes 2 and 3).

Members of the FGF and Wnt gene families are the primary

TABLE 1. Characteristics of mammary tumors having MMTV CIS at Int7/FLTL

Sample no.	Tissue	Chromosome	Insertion site	Orientation[a]	Distance (kb)[b]
676	HOG	15	Int7/FLTL	3′←5′-5→3′	3
630	Tumor	15	Int7/FLTL	3′←5′-5→3′	3
		16	CD47	5′→3′-5′→3′	25
		11	WNT3a	3′←5′-5→3′	15
637	Metastasis	15	Int7/FLTL	3′←5′-5→3′	3
		7	FGF3/FGF4	5′→3′-5′→3′-5′→3′	14/7
649	Tumor	15	Int7/FLTL	3′←5′-5′→3′	2.5
641	Metastasis	15	Int7/FLTL	3′←5′-5→3′	2.5
		7	FGF3/FGF4	5′→3′-5′→3′-5′→3′	15/6
4987	Tumor	15	Int7/FLTL	3′←5′-5′→3′	3.5
		7	FGF3	5′→3′-5′→3′	7
4979	Tumor	15	Int7/FLTL	3′←5′-5′→3′	2.5
		11	WNT3a	3′←5′-5′→3′	1
5165	Tumor	15	Int7/FLTL	3′←5′-5′→3′	3
		15	WNT10	5′→3′-5′→3′	2.2

[a] Transcriptional orientation of the provirus relative to the target gene. In the case of samples 637 and 641, the provirus insertion site is located between FGF3 and FGF4.

[b] Distance from the provirus insertion site to the target gene promoter. In cases in which two numbers are given, the first refers to FGF3 and the second to FGF4.

MRFCLFSFAL IILNCMDYSQ CQGNRWRRNK RASYVSNPIC KGCLSCSKDN GCSRCQQKLF

Furin-like

FFLRREGMRQ YGECLHSCPS GYYGHRAPDM NRCARCRI**EN CDSCFSKDFC TKCKVGFYLH**

Thrombospondin-like

RGRCFDECPD GFAPLDETME CVEGCE**VGHW SEWGTCSRNN RTCGFKWGLE TRTRQIVKKP**

NLS

AKDTIPCPTI AESRRCKMAM **RHCPGGKRTP KAKEKRNKKK RRK**LIERAQE QHSVFLATDR

VNQ

Furin-like

ENCD**S**CFSK--DFCTKC**K**VGFYLHRGRCFDECP**D**GFAPLDETMEC**V** 2610028F08Rik

PSCATCTGPGPDQCTSCRHGFYLDGGTCVSECPEGTYADTEGGVCL consensus

Thrombospondin type 1

VGHWSEWGTCSRNNRTCGFKWG**LE**TRTRQIVKKPAKDTIPCPTIA-E**SR**RCKM 2610028F08Rik

WGEWSEWSPCS---VTCGG--GVQTRTRSCCNPPPPGGGPCTGEDPETRACNE consensus

FIG. 2. The amino acid sequence of the 2610028F08Rik (Int7/FLTL) protein. The regions of the protein that are similar to furin and thrombospondin and the nuclear localization (NLS) consensus sequences are underlined and in bold. A comparison of the furin and thrombospondin consensus sequences and Int7/FLTL amino acid sequences is shown in the panel below. Identities are underlined and similarities are in bold. Abbreviations for amino acid residue are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

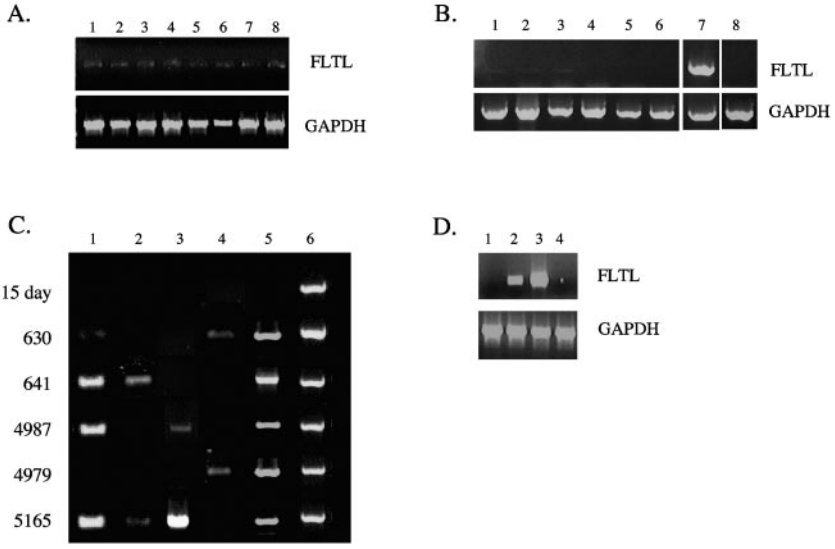

FIG. 3. RT-PCR assay for expression of Int7/FLTL and other CIS genes. (A) Total RNAs were prepared (5) from brain (lane 1), heart (lane 2), kidney (lane 3), liver (lane 4), lung (lane 5), spleen (lane 6), salivary gland (lane 7), and uterus (lane 8) and tested for Int7/FLTL and GAPDH (glyceraldehyde-3-phosphate dehydrogenase) expression. (B) Total RNA was prepared from female mammary glands of virgin (lane 1), day 5 pregnant (lane 2), day 10 pregnant (lane 3), day 5 lactating, day 10 lactating (lane 5), and day of involution (lane 6) mice, tumor 4987 (lane 7), and tumor 5465 (lane 8) and tested for Int7/FLTL protein and GAPDH expression. (C) Total RNAs were prepared from mammary tissue of a day 15 pregnant female, and tumors 630, 641, 4987, 4979, and 5165. They were tested for expression of FGF3 (lane 1), FGF4 (lane 2), Wnt10b (lane 3), Wnt3a (lane 4), Int7/FLTL protein (lane 5), and GAPDH (lane 6). (D) Total RNAs were prepared from HC11 cells (lane 1), HC11-Int7/FLTL cells (lane 2), tumor 630 (lane 3), and tumor 5465 (lane 4) and tested for Int7/FLTL protein and GAPDH expression. The primers used for RT-PCR analysis were as follows: Wnt3a 615f (5′GGAATGGTCTCTCGGGAGTTTG3′) and Wnt3a 977r (5′AGGTTCGCA GAAGTTGGGTGAG3′), Wnt9a 98f (5′TTCGGGCTAACAGGCAGTGAAC3′) and Wnt9a 398r (5′CAGAAGAGATGGCGTAGAGGAAA G3′), FGF3 228f (5′ GGAGATTACTGCGGTGGAAGTG3′) and FGF3 521r (5′ TTTTGTGTGCGGCGGGTCTTGAAG3′), FGF4 247f (5′ TACTGCAACGTGGGCATCGGATTC3′) and FGF4 590r (5′ TGGGTTACCTTCATGGTAGGCGAC3′), Int7/FLTL-f (5′CAATGGTTGCA GCCGATGTCAACAG3′) and Int7/FLTL-r (5′CAGTGCCTCATGGCCATCTTGCATC3′), Wnt10b-f (5′ GTTCTCTCGGGATTTCTTGGAT TC3′) and Wnt10b-r (5′ CATCACACAGCACATAACAGCACC3′), and GADPH F primer (5′CCACCTTCTTGATGTCATCAT) and GADPH R primer (5′ CCCTTCATTGACGTCAACTAG 3′). The Invitrogen (Carlsbad, CA) Superscript one-step RT-PCR with Platinum *Taq* polymerase was used according to the manufacturer's directions; the RT-PCR conditions were 50°C for 30 min and 94°C for 2 min, followed by 32 cycles of 56°C for 30 s and 72°C for 30 s, followed by 72°C for 5 min.

CIS for MMTV in "high-incidence" inbred mouse strains for mammary tumors (reviewed in reference 3). Frequently, tumors are positive for MMTV integrations at both Wnt and FGF genes, leading to the concept that MMTV-induced mutations collaborate in the induction of mammary tumors. This was subsequently confirmed in MMTV-induced mammary tumors in Wnt1 or FGF3 transgenic mice, in which the complementing gene (a member of the FGF or Wnt gene family, respectively) was frequently rearranged by the virus (8, 9, 12). However, since mammary tumors also arise sporadically in bitransgenic (Wnt1 plus FGF3) females, additional mutation(s) must be required for tumor development (7).

Southern blot analysis of the Int7/FLTL-positive tumors shows that each has between three and six proviral genomes (data not shown). In each tumor, at least one of these viral insertions occurred near a member of either the Wnt or FGF gene family (Table 1). To determine whether retroviral integration affected flanking gene expression, we designed primers to the specific gene sequences and performed reverse transcriptase (RT)-PCR to determine relative levels of expression. These genes are not normally expressed or expressed at very low levels in the mammary gland. The results, shown in Fig. 3C, demonstrate that the genes flanking the sites of retroviral integration were selectively expressed in these tumor samples compared to normal day 15 pregnant mammary glands. Interestingly, in metastasis 641 and tumor 5165, expression of both FGF3 and FGF4 was detected. These genes are approximately 20 kb apart. In the case of metastasis 641, the viral insertion occurred between the two genes. At present, we have not detected a viral insertion site around these genes in tumor 5165. Similarly, in tumor 4987, Wnt10b is expressed, although we have not been able to locate the putative viral insertion site in the DNA of this tumor. By Southern blot analysis, the location of the three viral insertions in HOG CZZ26 (sample 676) and tumor 630 were identical, whereas metastasis 637 contained three additional viral insertion sites (data not shown). We suspect that FGF3 expression detected in tumor 630 is a consequence of a viral insertion near this gene in a subpopulation of tumors that subsequent contributed to metastasis 637. Our results (summarized in Table 1), taken together, suggest that virus-induced expression of FLTL represents an early event in mammary tumorigenesis. Although we do not know the molecular consequences of Int7/FLTL expression on mammary gland development, we speculate that tumor progression occurs as a consequence of the collaborative effect of Int7/FLTL expression with the virus-induced expression of members of either the FGF or Wnt family of secreted growth factors.

REFERENCES

1. **Akagi, K., T. Suzuki, R. M. Stephens, N. A. Jenkins, and N. G. Copeland.** 2004. RTCGD: retroviral tagged cancer gene database. Nucleic Acids Res. **32**(Database issue)**:**D523–D527.
2. **Ball, R. K., R. R. Friis, C. A. Schoenenberger, W. Doppler, and B. Groner.** 1988. Prolactin regulation of beta-casein gene expression and of a cytosolic 120-kd protein in a cloned mouse mammary epithelial cell line. EMBO J. **7:**2089–2095.
3. **Callahan, R., and G. H. Smith.** 2000. MMTV-induced mammary tumorigenesis: gene discovery, progression to malignancy and cellular pathways. Oncogene **19:**992–1001.
4. **Erkeland, S. J., M. Valkhof, C. Heijmans-Antonissen, A. van Hoven-Beijen, R. Delwel, M. H. Hermans, and I. P. Touw.** 2004. Large-scale identification of disease genes involved in acute myeloid leukemia. J. Virol. **78:**1971–1980.
5. **Gallahan, D., C. Kozak, and R. Callahan.** 1987. A new common integration region (int-3) for mouse mammary tumor virus on mouse chromosome 17. J. Virol. **61:**218–220.
6. **Joosten, M., Y. Vankan-Berkhoudt, M. Tas, M. Lunghi, Y. Jenniskens, E. Parganas, P. J. Valk, B. Lowenberg, E. van den Akker, and R. Delwel.** 2002. Large-scale identification of novel potential disease loci in mouse leukemia applying an improved strategy for cloning common virus integration sites. Oncogene **21:**7247–7255.
7. **Kwan, H., V. Pecenka, A. Tsukamoto, T. G. Parslow, R. Guzman, T. P. Lin, W. J. Muller, F. S. Lee, P. Leder, and H. E. Varmus.** 1992. Transgenes expressing the Wnt-1 and int-2 proto-oncogenes cooperate during mammary carcinogenesis in doubly transgenic mice. Mol. Cell. Biol. **12:**147–154.
8. **Lee, F. S., T. F. Lane, A. Kuo, G. M. Shackleford, and P. Leder.** 1995. Insertional mutagenesis identifies a member of the Wnt gene family as a candidate oncogene in the mammary epithelium of int-2/Fgf-3 transgenic mice. Proc. Natl. Acad. Sci. USA **92:**2268–2272.
9. **MacArthur, C. A., D. B. Shankar, and G. M. Shackleford.** 1995. Fgf-8, activated by proviral insertion, cooperates with the Wnt-1 transgene in murine mammary tumorigenesis. J. Virol. **69:**2501–2507.
10. **Marchetti, A., F. Buttitta, S. Miyazaki, D. Gallahan, G. H. Smith, and R. Callahan.** 1995. Int-6, a highly conserved, widely expressed gene, is mutated by mouse mammary tumor virus in mammary preneoplasia. J. Virol. **69:** 1932–1938.
11. **Okazaki, Y., M. Furuno, T. Kasukawa, J. Adachi, H. Bono, S. Kondo, I. Nikaido, N. Osato, R. Saito, H. Suzuki, I. Yamanaka, H. Kiyosawa, K. Yagi, Y. Tomaru, Y. Hasegawa, A. Nogami, C. Schonbach, T. Gojobori, R. Baldarelli, D. P. Hill, C. Bult, D. A. Hume, J. Quackenbush, L. M. Schriml, A. Kanapin, H. Matsuda, S. Batalov, K. W. Beisel, J. A. Blake, D. Bradt, V. Brusic, C. Chothia, L. E. Corbani, S. Cousins, E. Dalla, T. A. Dragani, C. F. Fletcher, A. Forrest, K. S. Frazer, T. Gaasterland, M. Gariboldi, C. Gissi, A. Godzik, J. Gough, S. Grimmond, S. Gustincich, N. Hirokawa, I. J. Jackson, E. D. Jarvis, A. Kanai, H. Kawaji, Y. Kawasawa, R. M. Kedzierski, B. L. King, A. Konagaya, I. V. Kurochkin, Y. Lee, B. Lenhard, P. A. Lyons, D. R. Maglott, L. Maltais, L. Marchionni, L. McKenzie, H. Miki, T. Nagashima, K. Numata, T. Okido, W. J. Pavan, G. Pertea, G. Pesole, N. Petrovsky, R. Pillai, J. U. Pontius, D. Qi, S. Ramachandran, T. Ravasi, J. C. Reed, D. J. Reed, J. Reid, B. Z. Ring, M. Ringwald, A. Sandelin, C. Schneider, C. A. Semple, M. Setou, K. Shimada, R. Sultana, Y. Takenaka, M. S. Taylor, R. D. Teasdale, M. Tomita, R. Verardo, L. Wagner, C. Wahlestedt, Y. Wang, Y. Watanabe, C. Wells, L. G. Wilming, A. Wynshaw-Boris, M. Yanagisawa, et al.** 2002. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. Nature **420:**563–573.
12. **Shackleford, G. M., C. A. MacArthur, H. C. Kwan, and H. E. Varmus.** 1993. Mouse mammary tumor virus infection accelerates mammary carcinogenesis in Wnt-1 transgenic mice by insertional activation of int-2/Fgf-3 and hst/Fgf-4. Proc. Natl. Acad. Sci. USA **90:**740–744.
13. **Shin, M. S., T. N. Fredrickson, J. W. Hartley, T. Suzuki, K. Agaki, and H. C. Morse III.** 2004. High-throughput retroviral tagging for identification of genes involved in initiation and progression of mouse splenic marginal zone lymphomas. Cancer Res. **64:**4419–4427.
14. **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.