JVI Accepts, published online ahead of print on 27 May 2009
J. Virol. doi:10.1128/JVI.00427-09

Loss of microRNA-targets in the 3'-untranslated region as a mechanism of retroviral insertional activation of Growth Factor Independence 1

Magdalena Julia Dabrowska[1,2], Karen Dybkær[1,2], Hans Erik Johnsen[1], Bruce Wang[3], Matthias Wabl[4], and Finn Skou Pedersen[2]*

[1]Department of Haematology, Aalborg Hospital, Aarhus University Hospital, Denmark, [2]Department of Molecular Biology, Aarhus University, Denmark, [3]Picobella L.L.C, 863 Mitten Road, Suite 101, Burlingame, CA 94010, USA, [4]Department of Microbiology and Immunology, University of California, San Francisco, CA 94143, USA

Running title: miRNA target loss by retroviral insertions in the *Gfi1* 3'UTR

*Corresponding author:
Finn Skou Pedersen
Department of Molecular Biology, Aarhus University, C.F. Møllers Allé, Bldg.1130, DK-8000, Aarhus C, Denmark.
Tel.: ++4589422614, fax: ++4586196500, e-mail: fsp@mb.au.dk

Word count abstract: 138
Word count text: 5410

ABSTRACT

The non-oncogene bearing retrovirus SL3-3 murine leukemia virus induces strictly T-cell lymphomas with a mean latency of 2-4 months in mice of the NMRI-inbred (i) strain. By high throughput sequencing of retroviral tags, we have identified the genomic region encoding the transcriptional repressor and oncogene, Growth Factor Independence 1 (*Gfi1*) as a frequent target for SL3-3 in the NMRI-i mouse genome. Twenty-four SL3-3 insertions were identified within a 1kb window of the 3'UTR of the *Gfi1* gene - a clustering pattern unique for this lymphoma model. Expression analysis determined that the *Gfi1* gene was transcriptionally activated by SL3-3 insertions, and an upregulation in Gfi1 protein expression was detected for tumors harboring insertions in the *Gfi1* 3'UTR. Here, we provide data in support of a mechanism by which retroviral insertions in the *Gfi1* 3'UTR decouple microRNA mediated posttranscriptional regulation.

INTRODUCTION

The non-oncogene bearing murine leukemia viruses (MLVs) induce leukemias and lymphomas when injected into newborn susceptible mice (1, 21, 75). The major determinant of MLV latency and disease specificity is the retroviral enhancer in the U3 region of the MLV long terminal repeat (LTR) (3, 5, 9, 17, 19, 24, 37, 38, 50, 52, 69, 70, 71). It comprises conserved areas which hold densely packed binding sites for several host transcription factors, including Runx, NF-1, Ets, c-Myb, the glucocorticoid response element (GRE) and basic helix-loop-helix factors. Small nucleotide alterations in the different binding sites influence latency, confer variations in cell-specific expression, and shift disease patterns from lymphoma to plasmacytoma, myeloid leukemia, megakaryoblastic leukemia, erythroleukemia and mixed phenotype. The wild type (wt) SL3-3 is a highly pathogenic ecotropic MLV that induces precursor T-cell lymphomas (pre-TLLs) with a mean latency of 2-4 months and primary manifestations in thymus, spleen and mesenteric lymph nodes

when injected into mice of the NMRI-inbred (i) strain (19, 43, 51). Tumor induction by SL3-3 and other MLVs is a complex process, where the most well defined step involves integration of the viral genome into the host genome and deregulation of nearby proto-oncogenes or tumor suppressors (6, 8, 10, 28, 53, 65, 66, 67). The effect of the provirus depends on its integration position relative to the target gene where the most frequent mechanisms of insertional mutagenesis is enhancement and LTR promotion, both of which result in either upregulation of the wt gene and protein or generation of chimeric transcripts. Another way by which gene expression can be affected by retroviral insertions is by loss of regulatory regions. Early studies of insertional mutagenesis have demonstrated that retroviral integrations in the 3' untranslated region (UTR) of genes may result in generation of prematurely terminated transcripts or transcripts with increased mRNA stability and elevated protein synthesis (6, 8, 10, 67). The 3' UTR may also harbor other regulatory sequences, namely binding sites for microRNAs (miRNAs), which are non-coding 22 nt RNAs encoded from introns or intergenic regions in the genome (36). They act by targeting primarily the 3'UTRs of mRNAs and mediate posttranscriptional downregulation of gene expression by complete complementarity or partial binding of their 5'end nucleotide 2-7 (seed region) to mRNA targets (39). Theoretically, the short seed sequence permits a single miRNA to act on multiple target sites thereby each miRNA is able to recognize an average of 100 different mRNAs (2, 41).

The genomic locus on murine chromosome 5, encoding the transcriptional repressor and oncogene growth factor independence 1 (*Gfi1*) (25), and neuroblastoma 4S oncogene ecotropic viral integration site *5* (*Evi5*) (40) (hereafter also referred to as the *gfi1* locus), is a frequent integration locus in T-cell lymphomas induced by the Moloney murine leukemia virus (MoMLV) (48, 62, 65) and in B-cell lymphomas induced by the Akv MLV (72, 73). Previous studies have demonstrated that retroviral insertions within the *gfi1* locus lead to transcriptional activation of the *Gfi1* gene (62,

65). *Gfi1* is a key regulator of stem cell quiescence (29, 82) and plays a significant role in T-cell development (26, 54, 64, 81) and lineage commitment (80). It further influences maturation of myeloid precursors into granulocytes and monocytes and acts in limiting inflammatory immune response (31). *Gfi1* has a major oncogenic potential and has been associated with both murine and human cancers (15, 32, 59, 68).

In this study we have identified 130 retroviral insertions in the *gfi1* locus and addressed their effect on *Gfi1* mRNA and protein expression. Our results suggest that integrations in the *Gfi1* 3'UTR contribute to increased protein synthesis through a mechanism including loss of potential miRNA binding sites.

MATERIALS AND METHODS

Tumors and isolation of retroviral tags. Tumors originated from previously published (17, 18, 19, 20, 27, 43, 45, 51, 69, 70) and unpublished pathogenicity studies on wild type and enhancer mutated SL3-3, AKV and RFB. Large-scale analysis of integrated retroviruses, performed by a splinklerette-based PCR method described previously (78), was able to identify 120 wt and enhancer mutated SL3-3 integrations in the genomic region encoding *Gfi1* from a total of 790 SL3-3 tags. Seven Akv integrations and three RFB integrations from 2800 Akv tags and 85 RFB tags, respectively, were identified in the *gfi1* locus.

PCR and Sequencing. Total RNA was extracted from snap frozen tissue by use of TRIzol extraction reagent (Invitrogen). Full-genome cDNA was synthesized using First-Strand cDNA synthesis Kit (GE Healthcare) according to the manufacturer's recommendations. PCR for identifying alternative transcripts was performed with a Gfi1 exon 2 forward primer: 5'-ccgactctcagcttaccgag-3', and a Gfi1 exon 5 reverse primer: 5'-ctgtgtggatgaaggtgtgttt-3' (DNA

Technology). PCR for identifying retroviral insertions in the Gfi1 3'UTR was performed with a Gfi1 exon 6 forward primer: 5'-ctcaggaggcaccgagaga-3' and SL3-3 reverse primer: 5'-ccccagaaatagctaaaacaacaacagtttcaa-3' (DNA Technology). PCR fragments were purified on GFX columns (GE Healthcare) and sequenced by use of BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems).

Real Time PCR analysis. Real-Time PCR amplifications for gene mRNA quantification were performed using TaqMan® Expression Assays for *Gfi1* (Mm00515853_m1) and *Ywhaz* (Mm01158417_g1). For miRNA quantification, cDNA was synthesized according to the TaqMan® MicroRNA Assay protocol by use of TaqMan® MicroRNA Reverse Transcription Kit. and TaqMan® MicroRNA Assays for miR-155 (001806), miR-142-3p (001189), miR-330 (001062), miR-133a (002246), miR-34b-3p (002618), miR-879 (002473), miR-466l (002804), miR-10a (002288), and miR-467g (002811). Samples were set up in 20µL reactions with 10µL TaqMan® Universal PCR Master Mix, no AmpErase UNG, 0.5 µL TaqMan® primer-probe and 9µL cDNA. All TaqMan reagents were purchased from Applied Biosystems. To obtain amplification efficiency, samples for gene quantification were run in 4-point dilutions (1:10, 1:50, 1:100 and 1:500) and samples for miRNA quantification were run in 3-point dilutions (1:10, 1:50 and 1:100) . Each measurement was performed in duplicates. No template controls and no RT controls for each tumor sample were included. Samples for *Gfi1* quantification were normalized to *Ywhaz* (the housekeeping genes *Ubc*, *Tfrc*, *B2m*, and *Gapdh* were tested on 10 thymic, 10 splenic and 10 mesenteric lymph node samples where *Ywhaz* showed the most stable expression). MiRNA expression was normalized to snoRNA420 (001239) (Applied Biosystems). Each tumor sample was further normalized to its own tissue control counterpart.

Western Blot Analysis. Protein extraction was performed by homogenization of 60-120 ng snap-frozen tissue in 75mM NaCl, 100mM Tris-HCl pH 8, 5mM EDTA (pH 8), and 1mM PMSF.

Protein concentration was determined by use of a BCA™ Assay Kit (Pierce Biotechnology) according to the manufacturer's recommendations. 5µg protein from each sample was loaded onto Criterion XT 12% Bis-Tris Precast Gels (BioRad) and run in 0.5 × Criterion XT MOPS Running Buffer (BioRad). Proteins were transferred onto a polyvinylidene fluoride (PVDF) membrane (Millipore Corporation) and blocking was performed overnight at 4ºC in TBS-T (20mM Tris-HCl, pH7.6, 200mM NaCl) containing 5% (w/v) fat-free milk (FFM) and 0.05% Tween-20 (Sigma). Gfi1 primary antibody (ab21061) (Abcam) or β-actin primary antibody (sc-1616) (Santa Cruz Biotechnology) was diluted 1:1000 in TBS-T-0.05% Tween 20 and incubated with the membranes for 1 h at room temperature. Secondary horseradish peroxidase (HRP)-conjugated goat anti-rabbit antibody (sc-2004) or rabbit anti-goat (sc-2768) (Santa Cruz Biotechnology) was diluted 1:5000 in blocking solution and incubated with the membranes for 30 min at room temperature. Membranes were washed in TBS-T. All samples were run simultaneously and incubation of the membranes with antibodies was performed in the same solution to ensure sample comparability. The antigen-antibody complexes were visualized by use of ECL Western Blotting Detection Kit (GE Healthcare). The Western blot was repeated for 25 of the tumors with protein from a new round of purification to ensure reproducibility in observed expression patterns (data not shown).

Plasmid constructs and luciferase reporter assay. The SL3-3 LTR, Gfi1 3'UTR and Gfi1 3'UTR-SL3-3 constructs from tumor 2ML, 16T and 25S (integration position indicated in Fig. 1) were amplified using NotI and XhoI containing primers: Gfi1 3'UTR+XhoI forward primer: 5'-CACTCGAGGTACCCTGGCAGCCCGCAA, Gfi1 3'UTR+NotI reverse primer: 5'-CAGCGGCCGCGTAATAATCTTAATACTTTATTAAG-3', SL3-3+XhoI forward primer: 5'-CACTCGAGAATGAAAGACCCCTTCATAAGG-3', SL3-3+NotI reverse primer: 5'-CAGCGGCCGCAATGAAAGACCCCCAGGCTGG-3'. Constructs were ligated into the PsiCheck™-2 Vector (Promega). 293T cells were cultured in 48-well plates with 2×10^4 cells/well

in DMEM containing 10% FBS and maintained at 37ºC and 5% CO_2 for 24 hours prior to transfection. Cells were transfected by use of calcium phosphate in triplicates with 200ng vector and 30nM Pre-miR miRNA precursor (PM13058, PM10398) and Anti-miR miRNA inhibitor (AM13058, AM10398) (Ambion). Transfections were run in miRNA series so that all constructs were simultaneously co-transfected with a particular miRNA. Renilla/firefly activity was measured after 30 hours by use of Dual-Luciferase® Reporter Assay (Promega) on a LUMIstar Optima luminometer. Renilla/firefly values for the construct with the wt *Gfi1* 3'UTR were on average 2.5 fold lower than for the SL3-3 LTR, 2ML, 16T and 25S constructs. Renilla/firefly values for the different constructs were normalized to values for the control transfection with no added miRNAs. The results presented here are representative for at least two independent transfection experiments for each miRNA, meaning that approximately the same downregulation patterns were observed in both experimental sets for each miRNA.

RESULTS

The tumors assayed in this study originated from previously published and unpublished pathogenicity studies involving mainly wt SL3-3 (51) and Akv (43, 69), as well as SL3-3 and Akv mutated in the host transcription factor binding sites; Nuclear factor 1 (NF1) (18), Runx (19, 70), Glucocorticoid response element (GRE) (17, 69) and the basic helix-loop-helix motifs E_{gre} and $E_{a/s}$ (17, 69). A panel of tumors originated from experimental studies on SL3-3 with replaced envelope and integrase sequences from SL3-2 and Akv, respectively (unpublished data). Furthermore, tumors induced by SL3-3 mutated in the upstream conserved region (UCR) (45, and unpublished data) and the two 18-bp deletion variant (SL3-3 Turbo) (20, 51) were included in this study. High throughput sequencing of integrated retroviruses identified 2800 and 790 tags in tumors induced by Akv, and SL3-3, respectively. Additionally, 85 tags were obtained from tumors induced by the RFB (Reilly-

Finkel-Biskis) MLV which causes lymphomas, osteopetrosis, and osteomas when injected into NMRI-i mice (23, 63).

Frequent retroviral insertion in the *gfi1* genomic locus

By comparison of isolated tags within publicly available databases, 130 retroviral integrations from 95 tumors were mapped to a 150 kb genomic region on the murine chromosome 5 encoding *Gfi1* and *Evi5* (the *gfi1* locus) (Fig. 1). The majority of the integrations were mapped to the intergenic region between *Gfi1* and *Evi5* with the provirus oriented mainly in the opposite transcriptional direction of the genes. In a small number of tumors, provirus was positioned downstream of the *Gfi1* gene in the same transcriptional direction and in the 3' end of the *Evi5* gene in the opposite transcriptional direction. Notably, 27 retroviral insertions (24 SL3-3 insertions) were tightly clustered within a 1kb window in the 3'UTR of *Gfi1*, all oriented in the same transcriptional direction as the *Gfi1* gene (Fig.1 and Table 1). Integrations in the *gfi1* locus which were not positioned in the *Gfi1* 3'UTR will be referred to as "integrations outside the *Gfi1* 3'UTR".

The vast majority of the tumors with 3'UTR insertions were induced by the wt or enhancer mutated SL3-3. Akv insertion was only identified in two tumors, and one tumor was found to possess an integration of the RFB MLV. Each integration in the *Gfi1* 3'UTR, was validated by PCR using specific primers positioned in the 6th exon of *Gfi1* and in the SL3-3 LTR. In all cases, sequencing revealed presence of chimeric transcripts containing both *Gfi1* and SL3-3 LTR sequences (data not shown). The exact position of the provirus in respect to the *Gfi1* gene is indicated in Table 1.

This region was the most frequently targeted locus in the NMRI-i mouse genome and contained retroviral insertions in 15 percent (130 of 790) of all SL3-3 induced tumors, of which 90 percent (120 of 130) were integrations of wt SL3-3 or SL3-3 enhancer mutants (data not shown). There was no significant correlation between integration patterns and virus mutants. Our observations

demonstrated that the genomic locus encoding *Gfi1*, and the *Gfi1* 3'UTR in particular, are hotspots for retroviral insertions in the SL3-3/NMRI-i lymphoma model.

Retroviral insertions in the *gfi1* genomic locus generate truncated forms of *Gfi1* mRNA

To study *Gfi1* mRNA expression, 40 tumors were selected for splicing analysis based on accessibility and integration relative to the *Gfi1* gene. Samples included tumor material from thymus, spleen and mesenteric lymph node. PCR was performed using gene specific primers complementary to sequences in the 2nd and 5th exon in the murine *Gfi1* gene. Sequencing revealed the presence of three alternative *Gfi1* transcripts none of which have been previously identified (Fig. 2). The transcripts were characterized by exon 4 skipping (alternative transcript 1) and use of alternative 5' and 3' splice sites in exon 3 and 4, respectively (alternative transcripts 2 and 3, respectively). Moreover, alternative transcripts 1 and 2 had maintained their open reading frames. In the panel "integrations in the Gfi1 3'UTR", transcripts 1 and 3 were detected in SL3-3 (7 of 14) and Akv (1 of 2) induced tumors, while transcript 2 was detected in all tumors from this tumor group. Five of the tumors with insertions in the Gfi1 3'UTR had all three alternative transcripts. In tumors with insertions outside the Gfi1 3'UTR transcripts 1 and 3 were only detected in SL3-3 induced tumors (in 1 of 19 and 7 of 19 cases, respectively) while transcript 2 was identified in both SL3-3 (16 of 19) and Akv (1 of 4) induced tumors. Due to lack of tumor material, it was not possible to include more Akv or RFB induced tumors. Table 2 summarizes the frequency of alternative splicing within these two tumor groups. The alternative transcripts showed relatively equal distribution among the thymus, spleen and mesenteric lymph node tumors with no apparent correlation to either integration position, virus variant, or provirus orientation. Alternative splicing was also detected in MLV induced tumors without known integration on chromosome 5, but not in uninfected tissue nor in either of the control cell lines; L691, MPC11 or NIH-3T3, indicating that the aberrant splicing of

the *Gfi1* gene observed in our study is a general phenomenon of MLV induced lymphomas. We have not determined the relative abundance of the transcripts and the alternative splicing was not investigated further in this study.

Gfi1 is transcriptionally activated by the SL3-3 MLV

Previous small-scale studies have demonstrated that all MoMLV insertions in the genomic region encoding *Gfi1* and *Evi5* activate the *Gfi1* gene leading to a 3- to 6-fold transcriptional upregulation (62, 65). To evaluate the effect of retrovirus integration in the g*fi1* genomic locus on *Gfi1* mRNA expression, 43 tumors were screened by TaqMan Real-time PCR (Fig. 3A). Our data confirmed a general upregulation of *Gfi1* mRNA regardless position of the provirus in this 150 kb region. Notably, an up to 200-fold upregulation in tumor 19T and a 10- to 100-fold mRNA increase in 16 other tumors was observed. The transcription level of *Gfi1* was found to be significantly elevated in nearly all tumors analyzed, regardless of tissue type or provirus orientation. The upregulation was most prominent in SL3-3 induced tumors, but was not observed in Akv induced tumors, possibly indicating that *Gfi1* upregulation primarily takes place in development of T-cell lymphomas. In normal tissue, *Gfi1* was most abundant in spleen, with a somewhat lower expression in mesenteric lymph node and thymus.

Further expression analysis of *Evi5* mRNA (data not shown) revealed significant *Evi5* upregulation in Akv induced tumors without known integrations on chromosome 5, suggesting an oncogenic potential for *Evi5* in B-cell lymphomagenesis. *Evi5* was also activated in the RFB induced tumor 12S harbouring integration in the Gfi1 3'UTR but not in the Akv induced tumor 22S. Expression for tumor 9S was not investigated and no other Akv tumors with insertions in the gfi1 locus were included in the study due to lack of tumor material. In the SL3-3 induced tumors from both tumor groups, *Evi5* expression was varying, with no unambiguous expression pattern.

Decoupling of Gfi1 mRNA and protein expression in tumors harbouring retroviral insertions in the *Gfi1* 3'UTR

To investigate the correlation between *Gfi1* mRNA and protein expression, Western blot analysis was performed with a polyclonal antibody detecting Gfi1 at 50-55 kDa (Fig. 3B). Surprisingly, our results demonstrated major differences in Gfi1 protein expression, which appeared most abundant in tumors possessing insertions in the *Gfi1* 3'UTR. Gfi1 protein was identified at a relative high level in all tumors with Gfi1 3'UTR insertions except in 9S, 11ML, 19T, 25S, and 26S which showed vague or no protein expression. In tumors with integrations elsewhere in the *gfi1* locus, only 38S, 39T, 41T, 48S and 51S expressed the Gfi1 protein at equally high levels. No – or vague – protein expression was observed in the remaining tumors from this panel and no Gfi1 protein was detected in normal tissue controls. Based on the decoupled mRNA and protein expression patterns we hypothesized that integrations in *Gfi1* 3'UTR might disrupt a potential posttranscriptional miRNA regulation of *Gfi1*. Previous studies have indicated that *Gfi1* might be targeted by several miRNAs (7) and numerous predicted miRNA target sites in the *Gfi1* 3'UTR (Fig 4A) can also be found in the miRNA registries (http://microrna.sanger.ac.uk and http://microrna.org). However, experimental validation of whether *Gfi1* is subjected to miRNA regulation has not to our knowledge been presented yet. From alignment of predicted potential target sites for different miRNAs in the *Gfi1* 3'UTRs of various species, miR-142-3p and miR-155 showed the most conservation in the seed-binding region (Fig. 4B). Moreover, miR-142-3p displayed perfect base pairing of the nucleotides 2-8, while a single wobble at position 6 was present in the miR-155 seed sequence. Due to the conservation between species and to the already established expression of miR-142-3p and miR-155 in T- and B-lymphocytes (47), these were selected as main candidates for further analysis and real time qPCR was performed on all 46 samples. Of the remaining miRNAs, which showed no

major conservation between species in the region binding the miRNA seed sequence (alignment not shown), miR-330, miR-133a, miR-34b-3p, miR-10a, miR-879, miR-466l, and miR-467g were selected for expression analysis. Thymus, spleen, mesenteric lymph node and 12 tumors from the two tumor groups with and without Gfi1 protein expression were assayed for miRNA expression. All data were calculated by delta C_T method and the values were normalized to snoRNA420 and tissue controls (Fig. 4C). MiR-879 and miR-467g were not detected in any of the samples. MiR-142-3p and miR-155 expression patterns were varying while expression of miR-330, miR-133a, miR-34b-3p, miR-10a, and miR-466l mainly downregulated in comparison to control tissue. There was no significant difference in expression between the two tumor groups for any of the miRNAs. The expression data indicate that the increase in Gfi1 protein in tumors with integrations in the *Gfi1* 3'UTR was not due to a decrease in miRNA levels.

Downregulation of the *Gfi1* 3'UTR by miR-142-3p, miR-155, miR-10a and miR-133a

To determine if any of the selected miRNAs were able to recognize the 3'UTR and mediate translational regulation of the *Gfi1* transcript, we made a Renilla/luciferase reporter system with different constructs containing the *Gfi1* 3'UTR, SL3-3 LTR, as well as 3'UTR-SL3-3 LTR sequences representing retrovirus integration in the tumors 2ML, 16T and 25S (Fig 5A). The 2ML and 16T constructs contained only the miR-142-3p binding site, while the 25S construct contained the miR-330 and miR 133a binding sites. The *Gfi1* 3'UTR contained all miRNA binding sites.

Our data (Fig 5B) demonstrated that miR-142-3p was capable of downregulating all constructs including the empty psiCheck-2 vector, indicating that the effect was not specific for the *Gfi1* 3'UTR only. In all cases, downregulation by miR-142-3p was rescued by co-transfection with miR-142-3p inhibitor, establishing a specific effect of miR-142-3p on all constructs. Screening of the psiCheck-2 vector sequence detected perfect seed-match between miR-142-3p and Renilla (position

982-987 and 1075-1080, data not shown). Likewise, screening of the SL3-3 LTR sequence identified weak mir-142-3p complementarity (nucleotide 2-6 with one G:U base-pairing, and nucleotides 3-9 with two G:U base-pairings, not shown in this figure).

In contrast, both miR-155 and miR 10a were able to downregulate the *Gfi*1 3'UTR significantly in comparison to the SL3-3 LTR and a full rescue was observed in both cases. MiR-133a downregulated the Gfi1 3'UTR and 25S constructs and a small knockdown was also observed on the SL3-3 LTR. A complementary region to the miR-133a seed sequence was found in the SL3-3 LTR (nucleotides 2-7 with one G:U base-pairing), however, downregulation of the 2ML and 16T constructs was not observed. The miRNAs 34b-3b and 466l did not have a specific effect on any of the constructs (data not shown). Our data suggested that miR-155, miR-10a and miR133a were able to recognize and bind to sequences present in the 3'UTR of *Gfi1* and that the main silencing effect of miR-142-3p was due to recognition of additional binding to complementarity sequences in Renilla and possibly also the SL3-3 LTR.

DISCUSSION

The mechanism of insertional mutagenesis in murine models and identification of retroviral insertion sites by high throughput screening of the mouse genome has been widely used in identification of genes contributing to murine lymphomagenesis (34, 48, 73).

By large-scale analysis of integrated retroviruses in MLV infected NMRI-i mice we have identified the genomic region encoding *Gfi1* as the most frequently targeted locus and addressed the effect of these insertions on expression of the *Gfi1* gene. *Gfi1* has previously been identified as a common integration site for several retroviruses, including MoMLV (22, 30, 62, 65), Akv (72, 73), and MCF (40). Accumulating retroviral insertions identified in various mouse strains have made *Gfi1* a highly targeted gene in MLV induced lymphomas with 82 integrations available from the Retrovirus

Tagged Cancer Gene Database (RTCGD) (http://rtcgd.abcc.ncifcrf.gov/) and many more which have been identified in recent studies (4, 76, 78). We here report on further identification of 130 MLV insertions in and adjacent to *Gfi1*. The majority of the integrations were of wt or enhancer mutated SL3-3. In most of the tumors the provirus was found in the intergenic region between *Gfi1* and *Evi5* in the opposite transcriptional direction, displaying integration patterns similar to those described previously (http://rtcgd.abcc.ncifcrf.gov/). Additionally, a tight cluster of 24 SL3-3 insertions was mapped to a 1kb region in the *Gfi1* 3'UTR. Such clustering in the *Gfi1* gene has not been documented in other virus/host models and appears to be unique for SL3-3 in the NMRI-i mouse strain.

The differences in integration patterns between studies are often a result of different combinations of mouse genetic background and virus strain. For instance, both *Gfi1* and *Myc* are frequently targeted in by MoMLV in *p27kip*- (30) and *Cdkn2a*-deficient (44) mice of the C57Bl6/129 strain, but are rarely targeted in BHX2 mice. Likewise, the SL3-3 Turbo enhancer variant has distinct integration hotspots in the *c-Myc* promoter compared to the wt SL3-3 (51) and the wt SL3-3 has different integration patterns in the *Fos/Jdp2/Batf* locus in comparison to other SL3-3 enhancer variants (55). The variation of targets in different model systems may reflect different but overlapping pathways to lymphoma development (55).

To determine the effect of the provirus on *Gfi1* expression, *Gfi1* mRNA and protein levels were determined in 43 tumors. In agreement with earlier studies (62, 65), we found that MLV integration activated *Gfi1* expression regardless of provirus position. *Gfi1* mRNA upregulation was most profound in SL3-3 induced tumors, supporting the involvement of *Gfi1* in T-cell lymphomas. In Akv induced lymphomas *Gfi1* was downregulated compared to control tissue, strongly indicating that *Gfi1* does not have an oncogenic effect in B-cell lymphomagenesis. Thereby we failed to

support previous reports of frequent Akv integration in this locus (72, 73). Evidence from several studies, point towards a role for *Gfi1* in development of B-cell tumors, a notion supported by findings of plasmacytosis in *Gfi1*-deficient mice (56) and increased *Gfi1* levels in a subset of murine B-cell lymphomas in the marginal zone (68). Furthermore, *Gfi1* expression has been detected in early B-cell progenitors (81) and it has been suggested that Gfi1 control cytokine dependent B-cell differentiation (57). Taken together, these observations point towards a definite role for *Gfi1* in both T-cell and B-cell development and in lymphomagenesis.

Here, we have demonstrated that retroviral insertions in the *Gfi1* 3'UTR result in elevated *Gfi1* mRNA levels in nearly all tumors. However, *Gfi1* protein was primarily detected in tumors with retroviral integrations in the 3'UTR and only in few tumors from the panel "integrations outside the *Gfi1* 3'UTR". The observation that integrations outside the *Gfi1* 3'UTR activate *Gfi1* on the mRNA level but do not have an effect on Gfi1 protein expression may indicate that Gfi1 is posttranscriptionally downregulated.

A stabilizing function of retroviral insertions in the 3'UTR in the same transcriptional orientation as the gene has previously been proposed for several genes, including *Pim-1* (8, 66), *Myc* (6), and *Int-2* (10). Hereby, 3'non-coding sequences that negatively affect the mRNA stability are removed, rendering the normal mRNA unstable, and leading to accumulation of abnormal mRNA and protein levels. Retroviral insertions are further able to facilitate the use of cryptic promoters (6) or destroy important regulatory elements such as A/U rich regions implicated in mRNA destabilization (79). Based on the preferred integration clustering in the 5'end of the *Gfi1* 3'UTR and the Gfi1 protein expression patterns observed here, we speculated whether retroviral integrations decoupled miRNA binding to the *Gfi1* 3'UTR, resulting in an increase in protein synthesis. In this case, it would be possible that the high Gfi1 protein expression observed in some of the tumors from the panel

"integrations outside of the *Gfi1* 3'UTR" (38S, 39T, 41T, 48S, and 51S) could reflect deregulation of other proteins important for miRNA processing, although we did not succeed in identifying such integrations.

The *Gfi1* 3'UTR holds predicted binding sites for several miRNAs, including miR-142-3p and miR-155 which are found with relatively high abundance in most hematopoietic cells (47) and show highly conserved binding sites in the *Gfi1* 3'UTR of various species. Real-time PCR expression analysis for the miRNAs 155, 142-3p showed varying expression patterns while miRNAs 34b-3p, 10a, 466l, 133a, and 330 demonstrated a general downregulation in most of the tumors in comparison to control tissue. Our data indicated that the increase in Gfi1 protein in tumors with integrations in the *Gfi1* 3'UTR was not due to a decrease in miRNA levels for any of the miRNAs investigated here.

In order to determine if any of these miRNAs could be potential downregulators of *Gfi1*, Renilla luciferase reporter assay with different constructs was performed. Our results demonstrated that miR-142-3p was able to downregulate all constructs suggesting that this downregulation results from an interaction with competing target sequences in the psiCheck-2 vector and the SL3-3 LTR. Previous studies have demonstrated that multiple target sites can potentially increase the degree of translational suppression (13) possibly explaining the higher silencing observed here for miR-142-3p on the 3'UTR, 16T and 25S constructs. Only a minor downregulation by miR-142-3p was observed on the 2ML construct although, this contained the full binding site for miR-142-3p. The integration in tumor 2ML is positioned just three nucleotides downstream of the miRNA binding region, possibly influencing the structure of the small *Gfi1* 3'UTR fragment. Several studies have addressed the role of mRNA structure in miRNA target recognition and suggest that binding affinity

of a miRNA to its mRNA target is determined by both the sequence and structure of the mRNA (11, 12, 14, 33, 42). A possible explanation for the variability that we observed in our experiments may simply arise from differences in accessibility imposed by the sequence surrounding the target.

In contrast, miR-155, miR-10a and miR-133a all had a downregulating effect on the *Gfi1* 3'UTR construct, possibly suggesting a role for these miRNAs in posttranscriptional regulation of the *Gfi1* gene. Further more, the 25S construct which was the only chimeric construct containing the miR133a binding site was also downregulated. Together, our results support the hypothesis that Gfi1 may be downregulated by one or more miRNAs. However, we have only assessed the function of a small number of potential miRNAs. Other miRNAs may also have an effect on *Gfi1* regulation. To further validate if any of the miRNAs investigated here target the *Gfi1* gene, additional experiments including miRNA knockdown in different cell lines and following analysis of Gfi1 expression need to be performed.

Our results suggest that retroviral integrations in the Gfi1 3'UTR contribute to Gfi1 activation and possibly T-cell lymphomagenesis through loss of miRNA binding sites. In the majority of the tumors with insertions elsewhere in the *gfi1* locus, no Gfi1 protein expression was observed. It would therefore be unlikely that integrations in this region were the driving force of the tumorigenesis in these tumors. *Ccnd3*, *Myc/Pvt1*, *Ras2* and *RasGrp1* which were previously identified as possible Gfi1 cooperative partners in lymphoma development (76), were found as recurring integrations in several of the tumors with integrations in the *gfi1* locus. Development of T-cell lymphomas in tumors with integrations outside the *Gfi1* 3'UTR that do not express the Gfi1 protein could be a result of activation of these or other oncogenes. Of the 130 insertions, Ccnd3 was a co-target in 10 tumors and was also found as a target in tumor 42S, 45ML, 49ML and 50ML; Myc/Pvt1was targeted 4 times including in 8T and 48S; Ras2 was co-targeted in 9 tumors including 42S and RasGRP1 was co-targeted in 5 tumors including 37M and 43ML. We do not know how

Ccnd3, *Myc/Pvt1*, *Ras2* and *RasGrp1* are expressed in the tumors investigated in this study and the techniques used to identify retroviral tags do not necessarily identify all integrations. Overall, more extensive analyses need to be performed in order to obtain a clearer impression of how these tumors were initiated.

In comparison to tumor development through insertional mutagenesis of proto-oncogenes or tumor suppressors, a recently discovered mode of tumor induction includes retroviral targeting of miRNA loci and deregulation of single miRNAs or miRNA cistrons. The SL3-3 retrovirus has been shown to activate the 17-92 miRNA cistron (78) while the avian leukosis virus targets the BIC gene (the chromosomal region encoding miR-155) (16, 74) and the Radiation MLV frequently integrates into a locus encoding a group of five differentially spliced non-coding RNAs known as *Kis2* (35). Retroviral integration in all these regions caused significant upregulation of the miRNA clusters demonstrating a role for these miRNAs in oncogenesis.

In this study, we have introduced a SL3-3/NMRI-i model with high retroviral integration frequency in the *gfi1* locus and deregulated Gfi1 mRNA and protein expression patterns. Our data indicate that retroviral insertions in the *Gfi1* 3'UTR contribute to activation of *Gfi1* by loss of regulatory regions important for miRNA posttranscriptional downregulation of the gene. It is possible that such loss of regulatory regions in 3'UTR of the human GFI1 gene might likewise contribute to human lymphomagenesis. The human GFI1 gene is encoded by the chromosomal region 1p22 (58), a locus commonly affected in several cancers, including mantle cell lymphoma (60, 61), MALT lymphoma (77) and neuroblastoma (46). Although, there has been no direct correlation between translocations in this region and the effect on the GFI1 gene, our studies support the importance of this genomic region on tumor development. In humans, pre-TLL is a rare disease with poor prognosis, but with a

clear diagnostic parallel to the same type of tumors observed in murine models (49). In term, the results presented here may contribute to understanding of the oncogenic mechanisms by which *Gfi1* is involved in development of both murine and human T-cell lymphomas.

Acknowledgments

The technical assistance of Astrid van der Aa Kühle is gratefully acknowledged. This project was supported by grants from The Karen Elise Jensen Foundation (M.J.D, K.D.), the Danish Cancer Society (H.E.J., F.S.P.), The Danish Agency for Science Technology and Innovation (F.S.P.), and by NIH grant R01AI41570 (M.W).

1 Ben-David, Y., E.B. Giddens, K. Letwin, and A. Bernstein. 1991. Erythroleukemia induction by Friend murine leukemia virus: insertional activation of a new member of the ets gene family, Fli-1, closely linked to c-ets-1. Genes Dev 6:908-18.

2 Brennecke, J., A. Stark, R.B. Russell, and S.M. Cohen. 2005. Principles of MicroRNA-Target Recognition. PLoS Biol 3:e85-e99.

3 Celander, D., and W.A. Haseltine. 1987. Glucocorticoid regulation of murine leukemia virus transcription elements is specified by determinants within the viral enhancer region. J Virol 61:269-75.

4 Chakraborty, J., H. Okonta, H. Bagalb, S.J. Lee, B. Fink, R. Changanamkandat, and J. Duggan 2008 Retroviral gene insertion in breast milk mediated lymphomagenesis. Virology 377:100-9.

5 Chatis, P. A., C. A. Holland, J. W. Hartley, W. P. Rowe, and N. Hopkins. 1983. Role for the 3′ end of the genome in determining disease specificity of Friend and Moloney murine leukemia viruses. Proc. Natl. Acad. Sci. USA 80:4408-11.

6 Corcoran, L.M., J.M. Adams , A.R. Dunn , and S. Cory . Murine T lymphomas in which the cellular myc oncogene has been activated by retroviral insertion. Cell 1984 37:113-22.

7 Costa, I.G.,S. Roepcke, and A. Schliep. 2007. Gene expression trees in lymphoid development. BMC Immunol 8:25-44.

8 Cuypers, H.T., G. Selten , W. Quint, M. Zijlstra , E.R. Maandag , W. Boelens , P. van Wezenbeek , C. Melief , and A. Berns . 1984. Murine leukemia virus-induced T-cell lymphomagenesis: integration of proviruses in a distinct chromosomal region. Cell 37:141-50.

9 DesGroseillers, L., and P. Jolicoeur. 1984. The tandem direct repeats within the long terminal repeat of murine leukemia viruses are primary determinant of their leukemogenic potential. J Virol 52:945-52.

10 Dickson, C., R. Smith , S. Brookes , and G. Peters . 1990. Proviral insertions within the int-2 gene can generate multiple anomalous transcripts but leave the protein-coding domain intact. J Virol 64:784-93

11 Didiano, D., and O. Hobert. 2006. Perfect seed pairing is not a generally reliable predictor for miRNA-target interactions. Nat Struct Mol Biol 13:849-51.

12 Didiano, D., and O. Hobert. 2008. Molecular architecture of a miRNA-regulated 3'UTR. RNA 14: 1297-317.

13 Doench, J.G., C.P. Petersen, and P.A. Sharp. 2003. siRNAs can function as miRNAs. Genes Dev 17:438–42.

14 Doench, J.G., and P.A. Sharp. 2004. Specificity of microRNA target selection in translational repression. Genes Dev 18:504-11.

15 Dwivedi, P.P., P.H. Anderson , J.L. Omdahl , H.L. Grimes , H.A. Morris , and B.K. May . 2005. Identification of growth factor independent-1 (GFI1) as a repressor of 25-hydroxyvitamin D 1-alpha hydroxylase (CYP27B1) gene expression in human prostate cancer cells. Endocr Relat Cancer 12:351-65.

16 Eis, P.S., W. Tam , L. Sun , A. Chadburn , Z. Li , M.F. Gomez , E. Lund , and J.E. Dahlberg. 2005. Accumulation of miR-155 and BIC RNA in human B cell lymphomas. Proc Natl Acad Sci U S A. 102:3627-32.

17 Ejegod, D., K.D. Sørensen, I. Mo□brugger, L. Quintanilla-Martinez, J. Schmidt, and F.S. Pedersen. 2009. Control of pathogenicity and disease specificity of a T-cell lymphomagenic gammaretrovirus by E-box motifs but not by an overlapping glucocorticoid response element. J Virol 83:336-46.

18 Ethelberg, S., B. Hallberg, J. Lovmand, J. Schmidt, A. Luz, T. Grundström, and F.S. Pedersen. 1997. Second-site proviral enhancer alterations in lymphomas induced by enhancer mutants of SL3-3 murine leukemia virus: Negative effects of Nuclear Factor 1 binding site. J Virol 71:1196-206.

19 Ethelberg, S., J. Lovmand, J. Schmidt, A. Luz, and F. S. Pedersen. 1997. Increased lymphomagenicity and restored disease specificity of AML1 site (core) mutant SL3-3

murine leukemia virus by a second-site enhancer variant evolved in vivo. J Virol 71:7273-80.

20 Ethelberg, S., A.B. Sørensen, J. Schmidt, A. Luz, and F.S. Pedersen. 1997. An SL3-3 murine leukemia virus enhancer variant more pathogenic than the wild type obtained by assisted molecular evolution in vivo. J Virol 71:9796-99.

21 Fan, H., B.K. Brightman, B.R. Davis, and Q.X. Li. 1991. Leukemogenesis by Moloney murine leukemia virus. *In* Viruses that affect the immune system (ed. H.Y. Fan et al.), pp. 155–74. American Society for Microbiology, Washington, D.C.

22 Gilks, C.B., S.E. Bear , H.L. Grimes , and P.N. Tsichlis. 1993. Progression of interleukin-2 (IL-2)-dependent rat T cell lymphoma lines to IL-2-independent growth following activation of a gene (Gfi-1) encoding a novel zinc finger protein. Mol Cell Biol 13:1759-68.

23 Gimbel, W., J. Schmidt, J. Barack-Werner, A. Luz, P.G. Strauss, V. Erfle, and T. Werner. 1996. Molecular and pathogenic characterization of the RFB osteoma virus: Lack of oncogene and induction of osteoma, osteopetrosis and lymphoma. Virology 224:533-8.

24 Golemis, E.A., N.A. Speck, and N. Hopkins. 1990. Alignment of U3 region sequences of mammalian type C viruses: identification of highly conserved motifs and implications for enhancer design. J Virol 64:534-42.

25 Grimes, H.L., T.O. Chan, P.A. Zweidler-McKay, B. Tong, and P.N. Tsichlis. 1996. The Gfi-1 proto-oncoprotein contains a novel transcriptional repressor domain, SNAG, and inhibits G1 arrest induced by interleukin-2 withdrawal. Mol Cell Biol 16:6263-72.

26 Grimes, H.L., C.B.Gilks, T.O. Chan, S. Porter, and P.N. Tsichlis. 1996. The Gfi-1 protooncoprotein represses Bax expression and inhibits T-cell death. Proc Natl Acad Sci USA 93:14569-73.

27 Hallberg, B., J. Schmidt, A. Luz, F.S. Pedersen, and T. Grundström. 1991. SL3-3 enhancer factor I transcriptional activators are required for tumor formation by SL3-3 murine leukemia virus. J Virol 65:4177-81.

28 Hayward, W.S., B.G. Neel, and S.M. Astrin. 1981. Activation of a cellular onc gene by promoter insertion in ALV-induced lymphoid leukosis. Nature 290:475-80.

29 Hock, H., M.J. Hamblen, H.M. Rooke, J.W. Schindler, S. Saleque, Y. Fujiwara and S.H. Orkin. 2004. Gfi-1 restricts proliferation and preserves functional integrity of haematopoietic stem cells. Nature 431:1002-7.

30 Hwang, H.C., C.P. Martins, Y. Bronkhorst, E. Randel, A. Berns, M. Fero, and B.E. Clurman, 2002. Identification of oncogenes collaborating with p27kip1 loss by insertional mutagenesis and high-throughput insertion site analysis. Proc Natl Acad Sci U S A 99:11293-8.

31 Karsunky, H., H. Zeng, T. Schmidt, B. Zevnik, R. Kluge, K.W. Schmid, U. Dührsen, and T. Möröy. 2002. Inflamatory reactions and severe neutropenia in mice lacking the transcriptional repressor Gfi1. Nat Genet 259-300.

32 Kazanjian, A., D. Wallis, N. Au, R. Nigam, K.J.T. Venken, P.T. Cagle, B.F. Dickey, H.J. Bellen, C.B. Gilks, and H.L. Grimes. 2004. Growth Factor Independence-1 is expressed in primary human neuroendocrine lung carcinomas and mediates the differentiation of murine pulmonary neuroendocrine cells. Cancer Res 64:6874-82.

33 Kertesz M., N. Iovino, U. Unnerstall, U. Gaul, and E. Segal. 2007. The role of site accessibility in miRNA target recognition. Nat Genet 39:1278-84.

34 Kim, R., A. Trubetskoy, T. Suzuki, N.A. Jenkins, N.G. Copeland, and J. Lenz. 2003. Genome-based identification of cancer genes by proviral tagging in mouse retrovirus-induced T-cell lymphomas. J Virol 77:2056-62.

35 Landais, S., S. Landry, P. Legault, and E. Rassart. 2007. Oncogenic potential of the miR-106-363 cluster and its implication in human T-cell leukemia. Cancer Res 67: 5699-707.

36 Lee, R.C., R.L. Feinbaum, and V. Ambros. 1993. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 75: 843–54.

37 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 308:467-70.

38 Lewis, A. F., T. Stacy, W. R. Green, L. Taddesse-Heath, J. W. Hartley, and N. A. Speck. 1999. Core-binding factor influences the disease specificity of Moloney murine leukemia virus. J. Virol. 73:5535-47.

39 Lewis, B.P., C.B. Burge, and D.P. Bartel. 2005. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 120:15-20

40 Liao, X., A.M. Buchberg, N.A. Jenkins, and N.G. Copeland. 1995. Evi-5, a common site of retroviral integration in AKXD T-cell lymphomas, maps near Gfi1 on mouse chromosome 5. J Virol 69:7132-7.

41 Lim, L.P., N.C. Lau, P. Garrett-Engele, A. Grimson, J.M. Schelter, J. Castle, D.P. Bartel, P.S. Linsley, and J.M. Johnson. 2005. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. Nature 433:769–73.

42 Long, D. R. Lee, P. Williams, C.Y. Chan, V. Ambros and Y. Ding. 2007. Potent effect of target structure on microRNA function. Nat Struct Mol Biol 14:287-94.

43 Lovmand, J., A.B. Sørensen, J. Schmidt, M. Østergaard, A. Luz, and F.S. Pedersen. 1998. B-cell lymphoma induction by Akv murine leukemia viruses harboring one or both copies of the tandem repeat in the U3 enhancer. J Virol 72:5745-56.

44 Lund, A.H., G. Turner, A. Trubetskoy , E. Verhoeven , E. Wientjens , D. Hulsman , R. Russell , R.A. DePinho, J. Lenz , and M. van Lohuizen. 2002. Genome-wide retroviral insertional tagging of genes involved in cancer in Cdkn2a-deficient mice. Nat Genet 32:160-5.

45 Ma, S.L., J. Lovmand, A.B. Sørensen, A. Luz, J. Schmidt, and F.S. Pedersen. 2003. Triple basepair changes within and adjacent to the conserved YY1 motif upstream of the U3 enhancer repeats of SL3-3 murine leukemia virus cause a small but significant shortening of latency of T-lymphoma induction. Virology 313:638-44.

46 Mead, R.S., and J.K. Cowell. 1995. Molecular characterization of a (1;10)(p22;q21) constitutional translocation from a patient with neuroblastoma. Cancer Genet Cytogenet 81:151-7.

47 Merkerova, M., M. Belickova, and H. Bruchova. 2008. Differential expression of microRNA in hematopoietic cell lineages. Eur J Haematol 81:304-310

48 Mikkers, H., J. Allen, P. Knipscheer, L. Romeijn, A. Hart, E. Vink, and A. Berns. 2002. High-throughput retroviral tagging to identify components of specific signaling pathways in cancer. Nat Genet 32:153-9.

49 Morse, H.C. 3rd, M.R. Anver, T.N. Fredrickson, D.C. Haines, A.W. Harris, N.L. Harris, E.S. Jaffe, S.C. Kogan, I.C. MacLennan, P.K. Pattengale, and J.M. Ward. 2002. Hematopathology subcommittee of the Mouse Models of Human Cancers Consortium. Bethesda proposals for classification of lymphoid neoplasms in mice. Blood 100:246-58.

50 Nielsen, A.L., P.L. Nørby, F.S. Pedersen, and P. Jørgensen. 1996. Various models of basic helix-loop-helix protein-mediated regulation of murine leukemia virus transcription in lymphoid cells. J Virol 70:5893-901.

51 Nielsen, A.A., A.B. Sørensen, J. Schmidt, and F.S. Pedersen. 2005 Analysis of wild-type and mutant SL3-3 murine leukemia virus insertions in the c-myc promoter during lymphomagenesis reveals target site hot spots, virus dependent patterns, and frequent erroe-prone gap repair. J Virol 79:67-78.

52 Nieves, A., L.S. Levy, and J. Lenz. 1997. Improtance of a c-Myb binding site for lymphomagenesis by the retrovirus SL3-3. J Virol 71:1213-9.

53 Nusse, R., and H.E. Varmus. 1982. Many tumors induced by the mouse mammary tumor virus contain a provirus integrated in the same region of the host genome. Cell 31:99-109.

54 Pargmann, D., R. Yücel, C. Kosan, I. Saba, L. Klein-Hitpass, S. Schimmer, F. Heyd, U. Dittmer, and T. Möröy. 2007. Differential impact of the transcriptional repressor *Gfi1* on mature $CD4^+$ and $CD8^+$ T lymphocyte function. Eur J Immunol 37:3551-63.

55 Rasmussen M.H., A.B. Sørensen, D.W. Morris, J.C. Dutra, E.K. Engelhard, C.L. Wang, J. Schmidt, and F.S. Pedersen. 2005. Tumor model-specific proviral insertional mutagenesis of the Fos/Jdp2/Batf locus. Virology 337:353-64.

56 Rathinam, C., H. Lassmann, M. Mengel, and C. Klein. 2008. Transcription factor Gfi1 restricts B-cell mediated autoimmunity. J Immunol 181: 6222-9.

57 Rathinam, C., and C. Klein. 2007. Transcriptional repressor Gfi1 integrates cytokine-receptor signals controlling B-cell differentiation. PLoS ONE 2:e306-e317.

58 Roberts, T., and J.K. Cowell. 1997. Cloning of the human Gfi-1 gene and its mapping to chromosome region 1p22. Oncogene 14:1003-5.

59 Sakai, I., H. Yamauchi, M. Yasukawa, H. Kohno, and S. Fujita. 2001. Expression of the Gfi-1 gene in HTLV-I-transformed T cells. Int J Hematol 73:507-16.

60 Salaverria, I., B. Espinet , A. Carrió , D, Costa , L, Astier , J. Slotta-Huspenina , L, Quintanilla-Martinez, F. Fend , F. Solé , D. Colomer , S. Serrano , R. Miró , S. Beà , and E. Campo . 2008. Multiple recurrent chromosomal breakpoints in mantle cell lymphoma revealed by a combination of molecular cytogenetic techniques. Genes Chromosomes Cancer 47:1086-97.

61 Sander, S., L. Bullinger , E. Leupolt , A. Benner , D. Kienle , T. Katzenberger , J. Kalla , G. Ott , H.K. Müller-Hermelink , T.F. Barth , P. Möller , P. Lichter , H. Döhner , and S. Stilgenbauer . 2008. Genomic aberrations in mantle cell lymphoma detected by interphase fluorescence in situ hybridization. Incidence and clinicopathological correlations. Haematologica 93:680-7.

62 Scheijen, B., J. Jonkers, D. Acton, and A. Berns. 1997. Characterization of pal-1, a common proviral insertion site in murine leukemia virus-induced lymphomas of c-myc and pim-1 transgenic mice. J Virol 71:9-16.

63 Schmidt, J., K. Lumniczky, B.D. Tzschaschel, H.L. Guenther, A. Luz, S. Riemann, W. Gimbel, V. Erfle, and R.G. Erben. 1999. Onset and Dynamics of Osteosclerosis in Mice Induced by Reilly-Finkel-Biskis (RFB) Murine Leukemia Virus. Am J Pathol 155:557–70

64 Schmidt, T., H. Karsunky, B. Rödel, B. Zevnik, H-P. Elsässer, and T. Möröy. 1998. Evidence implicating Gfi-1 and Pim-1 in pre-T-cell differentiation steps associated with β-selection. EMBO J 17:5349-59.

65 Schmidt, T., M. Zörnig, R. Benke, and T. Möröy. 1996. MoMuLV proviral integrations identified by Sup-F selection in tumors from infected myc/pim bitransgenic mice correlate with activation of the gfi-1 gene. Nucleic Acids Res 24:2528-34.

66 Selten, G.H, H.T Cuypers, and A. Berns. 1985. Proviral activation of the putative oncogene Pim-1 in MuLV induced T-cell lymphomas. EMBO J 4:1793-8.

67 Selten, G.H, H.T Cuypers, M. Zijlstra, C. Melief, and A. Berns. 1984. Involvement of c-Myc in T cell lymphomas in mice: frequency and mechanism of activation. EMBO J 3: 3215-22.

68 Shin, M.S., T.N. Fredrickson , J.W. Hartley , T. Suzuki , K. Akagi , and H.C. Morse, 3rd. 2004. High-throughput retroviral tagging for identification of genes involved in initiation and progression of mouse splenic marginal zone lymphomas. Cancer Res 64:4419-27.

69 Sørensen, K. D., S. Kunder, L. Quintanilla-Martinez, J. Sørensen, J. Schmidt, and F.S. Pedersen. 2007. Enhancer mutations of Akv murine leukemia virus inhibit the induction of mature B-cell lymphomas and shift disease specificity towards the more differentiated plasma cell stage. Virology 362:179-91.

70 Sørensen, K.D., L. Quintanilla-Martinez, S. Kunder, J. Schmidt, and F.S. Pedersen. 2004. Mutation of all Runx (AML1/core) sites in the enhancer of T-lymphomagenic SL3-3 murine leukemia virus unmasks a significant potential for myeloid leukemia induction and favors enhancer evolution toward induction of other disease patterns. J Virol 78:13216-31.

71 Speck, N.A., B. Renjifo, E. Golemis, T.N. Fredrickson, J.W. Hartley, and N. Hopkins. 1990. Mutation of the core or adjacent LVb elements of the Moloney murine leukemia virus enhancer alters disease specificity. Genes Dev 4:233-42.

72 Suzuki, T., K. Minehata, K. Agaki, N.A. Jenkins, and N.G. Copeland. 2006. Tumor suppressor gene identification using retroviral insertional mutagenesis in Blm-deficient mice. EMBO J 25:3422-31.

73 Suzuki T., H. Shen, K. Akagi, H.C. Morse, J.D. Malley, D.Q. Naiman, N.A. Jenkins, and N.G. Copeland. 2002. New genes involved in cancer identified by retroviral tagging. Nat Genet 32:166-74.

74 Tam, W., D. Ben-Yehuda, and W.S. Hayward. 1997. bic, a novel gene activated by proviral insertions in avian leukosis virus-induced lymphomas, is likely to function through its noncoding RNA. Mol Cell Biol 17:1490-502.

75 Tsichlis, P.N., and P.A. Lazo. 1991. Virus-host interactions and the pathogenesis of murine and human oncogenic retroviruses. Curr Top Microbiol Immunol 171:95-173.

76 Uren, A.G., J. Kool, K. Matentzoglu, J. de Ridder, J. Mattison, M. van Uitert, W. Lagcher, D. Sie, E. Tanger, T. Cox, M. Reinders, T.J. Hubbard, J. Rogers, J. Jonkers, L. Wessels, D.J. Adams, M. van Lohuizen, and A. Berns. 1998. Large-scale mutagenesis in p19ARF-and p53-deficient mice identifies cancer genes and their collaborative networks. Cell 133:727-41.

77 Vega, F., and L.J. Medeiros. 2001. Marginal-zone B-cell lymphoma of extranodal mucosa-associated lymphoid tissue type: molecular genetics provides new insights into pathogenesis. Adv Anat Pathol 8:313-26.

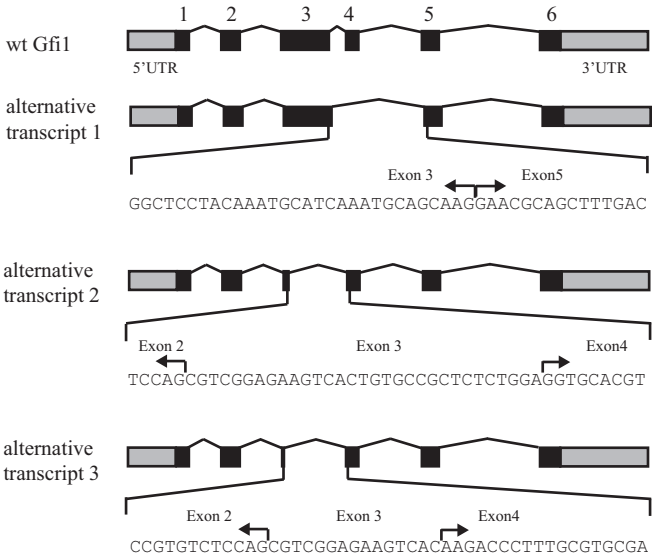
78 Wang, C.L., B.B. Wang, G. Bartha, L. Li, N. Channa, M. Klinger, N. 2006. Killeen, and M. Wabl. Activation of an oncogenic microRNA cistron by provirus integration. Proc Natl Acad Sci USA. 103: 18680-84.

79 Wingett, D., R. Reeves, and N.S. Magnuson. 1992. Characterization of the testes-specific pim-1 transcript in rat. Nucleic Acids Res. 20: 3183–9.

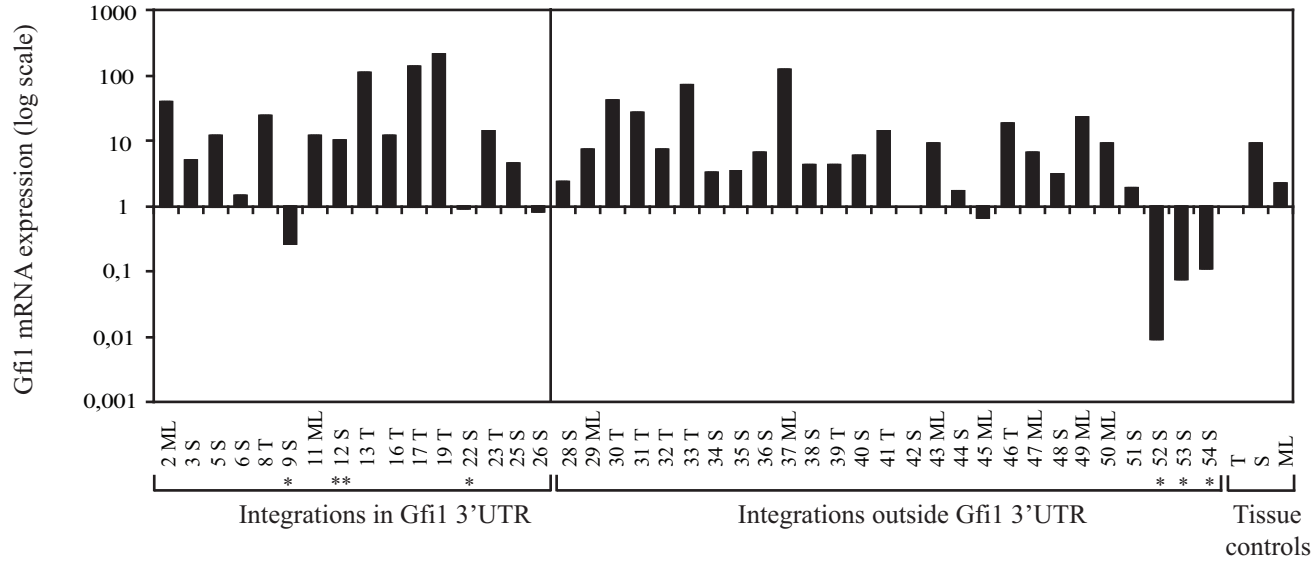
80 Yücel, R., H. Karsunky, L. Klein-Hitpass, and T. Möröy. 2003. The transcriptional repressor Gfi1 affects development of early uncommitted c-kit+ T cell progenitors and CD4/CD8 lineage decision in the thymus. J Exp Med 197:831-44.

81 Yücel, R., C. Kosan, F. Heyd, and T. Möröy. 2004. Gfi1:Green fluorescent protein knock-in mutant reveals different expression and autoregulation of the Growth Factor Independence 1 (Gfi1) gene during lymphocyte development. J Biol Chem 279:40906-17.

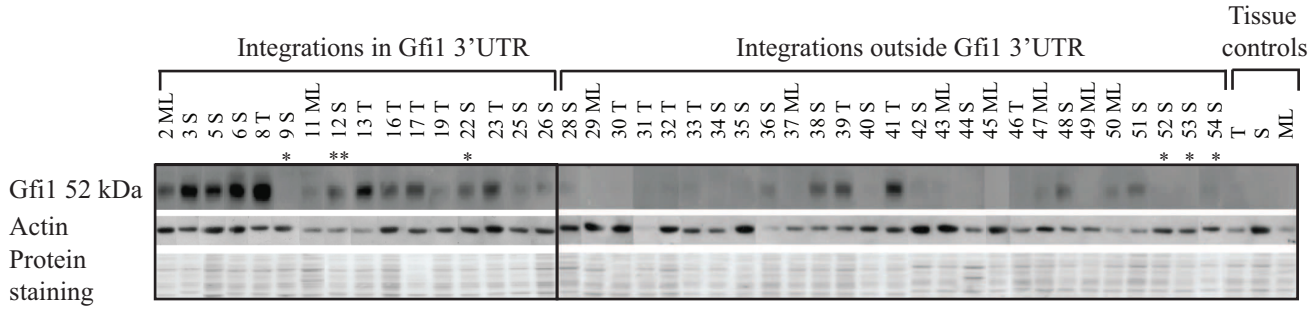
82 Zeng, H., R. Yücel, C. Kosan, L. Klein-Hitpass, and T. Möröy. 2004. Transcription factor Gfi1 regulates self-renewal and engraftment of hematopoietic stem cells. EMBO J 23:4116-25.

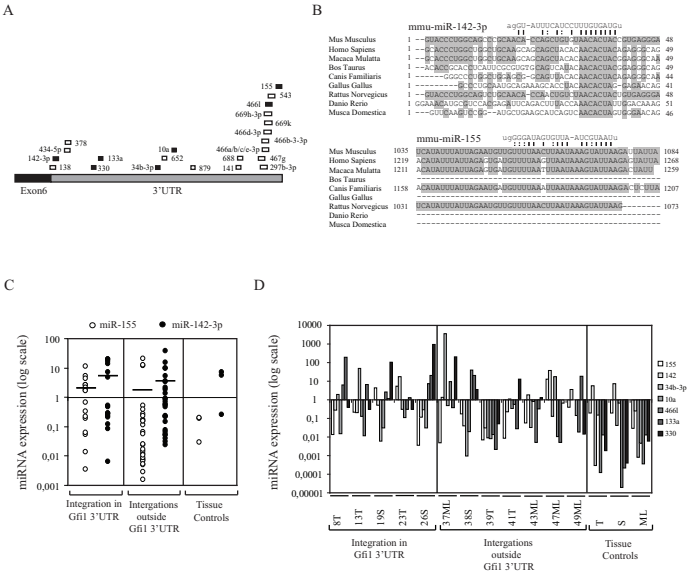


A



B





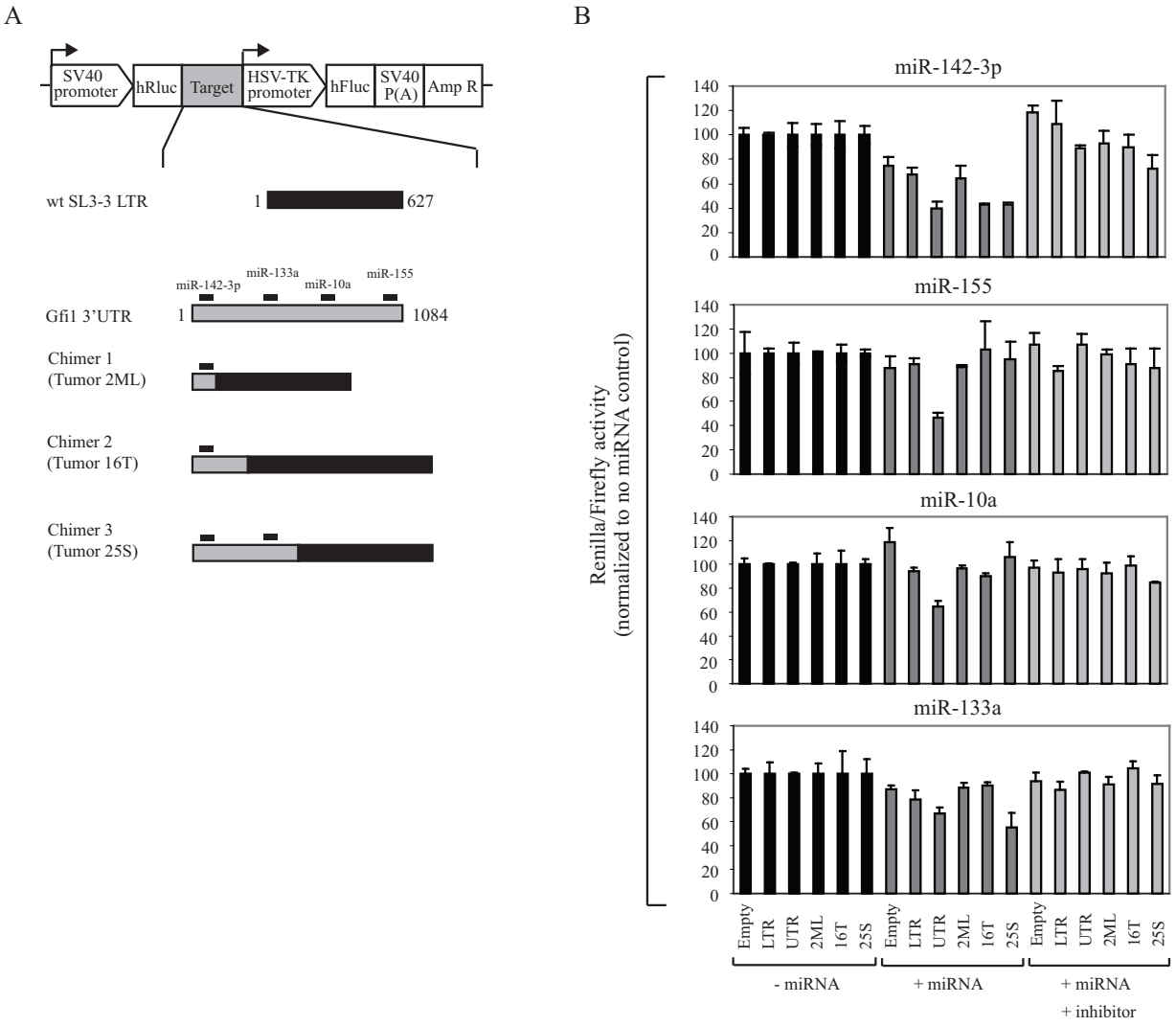


FIG. 1. MLV integrations identified in the genomic region encoding *Gfi1* and *Evi5*. *Gfi1* and *Evi5* gene structures are shown above with coding sequences (black) and UTRs (grey). The transcriptional direction of both genes is indicated by arrows. The number of integrations is indicated by bars. Grey bars represent retroviral integrations in the same transcriptional orientation; white bars represent integrations in the opposite transcriptional direction. Each bar represents 1084 nucleotides, corresponding to the size of the *Gfi1* 3'UTR. Integrations in the *Gfi1* 3'UTR are shown below the graph. Position and transcriptional orientation of the provirus is indicated by arrowheads.

FIG. 2. *Gfi1* alternative transcripts identified by sequencing. The wt *Gfi1* is shown at the top. Coding exons; black, UTRs; grey. Three alternative transcripts were detected in SL3-3 and Akv induced tumors, here referred to as alternative transcript 1, 2 and 3. Sequencing revealed exon 4 skipping (alternative transcript 1) and use of alternative 5' and 3' splice sites in exon 3 and 4, respectively (alternative transcripts 2 and 3)

FIG. 3. (A) *Gfi1* expression in MLV induced lymphomas. TaqMan real-time PCR was performed on 43 tumors harbouring integrations in the *Gfi1* 3'UTR and elsewhere in the *gfi1* locus as well as on the three control tissues; thymus (T), spleen (S), and mesenteric lymph node (ML). Tumors are indicated by numbers.*Akv integrations, **RFB integrations. *Gfi1* mRNA expression was normalized to the expression of the housekeeping gene, tyrosine 3-monooxygenase, Ywhaz. Thymic, spleenic and mesenteric lymph node tumors were further normalized to thymus, spleen and mesenteric lymph node tissue controls, respectively. All control tissue was normalized to the thymus control. (B) *Gfi1* protein expression in tumors with integrations in the Gfi1 3'UTR and outside of the Gfi1 gene. Tumors are indicated by numbers. T; thymus, S; spleen, ML; mesenteric

lymph node. *Akv integrations, **RFB integrations. *Gfi1* was detected at 50-55kDa. β-actin and AmidoBlack protein staining were used as a loading controls.

FIG 4. (A) Predicted miRNA binding sites in the *Gfi1* 3' UTR from miRNA registries (http://microrna.sanger.ac.uk and http://microrna.org, 04/2009). Position of the different miRNA binding sites in the *Gfi*1 3'UTR are indicated by boxes and miRNA IDs are indicated by numbers. MiRNAs indicated by black boxes were selected for expression analysis. (B) Alignment of miR-142-3p and miR-155 binding sites in different species. Conserved nucleotides are boxed. Presumed binding sites and nucleotide sequences for miR-142-3p and miR-155 are shown in the figure and complementary nucleotides are connected by lines. G:U base pairing is indicated by dashed lines. (C) MiR-142-3p and miR-155 expression in 43 tumors with integrations in the *Gfi1* 3'UTR and outside the *Gfi1* gene and three control tissue; thymus (T), spleen (S) and mesenteric lymph node (ML) assayed by TaqMan real-time PCR. The data were calculated by delta C_T method and the values were normalized to snoRNA420. Thymic, spleenic and mesenteric lymph node tumors were further normalized to thymus, spleen and mesenteric lymph node tissue controls, respectively. Expression for tissue controls is shown as normalized to snoRNA420. Lines indicate mean values for miRNA-142-3p and miR-155 expression. (D) Real time PCR expression analysis for miR-155, miR-142, miR-34b-3p, miR-10a, miR-466l, miR-133a, and miR-330 on 12 tumors with integrations in the Gfi1 3'UTR and elsewhere in the *gfi1* locus and three control tissue; thymus (T), spleen (S) and mesenteric lymph node (ML). The data was processed as described above.

FIG 5. Downregulation of *Gfi1* 3'UTR determined by Renilla luciferase Assay. (A) Constructs containing the SL3-3 LTR, *Gfi1* 3'UTR and *Gfi1* 3'UTR-SL3-3 chimeric sequences representing integrations of tumor 2ML, 16T and 25S were ligated into the psiCheck-2 vector. MiR-142-3p,

miR-155, miR-10a and miR-133a binding sites in the *Gfi1* 3'UTR and in the chimeric constructs are indicated. (B) Constructs were co-transfected into 293T cells with miRNA precursors and their respective anti-miRNA inhibitors. Single-transfections with the different constructs (minus (-) miRNA) and the empty psiCheck-2 vector were used as controls. Renilla/firefly activity for each co-transfection was normalized to the activity for the control trnasfections (minus miRNA control). The results presented here are representative for at least two independent transfection experiments for each miRNA, meaning that approximately the same downregulation patterns were observed in both experimental sets for each miRNA.

TABLE 1. Retroviral integrations in the *Gfi1* 3'UTR

Integration[a]	Tissue[b]	Virus[c]	Provirus orientation[d]	Provirus position[e]	Reference
1	S	SL3-3 UCR	+	6641	45, and unpublished data
2*	ML	SL3-3 wt	+	6646	51
3	S	SL3-3 $E_{a/s}$	+	6653	17
4*	ML	SL3-3 wt	+	6654	51
5	S	SL3-3 wt	+	6654	51
6***	S	SL3-3 UCR	+	6656	45, and unpublished data
7	S	SL3-3 wt	+	6658	51
8	T	SL3-3 (SL3-2Env)	+	6694	Unpublished data
9	S	Akv1-99 E_{gre}+$E_{a/s}$	+	6715	70
10	T	SL3-3 wt	+	6734	51
11	ML	SL3-3 wt	+	6734	51
12	S	RFB wt	+	6734	Unpublished data
13**	T	SL3-3 Turbo	+	6787	20, 51
14	S	SL3-3 GR+$E_{a/s}$	+	6791	17
15**	T	SL3-3 Turbo	+	6819	20, 51
16	T	SL3-3 Turbo	+	6819	20, 51
17	T	SL3-3 wt	+	6826	51
18	T	SL3-3 wt	+	6842	51
19	T	SL3-3 wt	+	6919	51
20***	S	SL3-3 UCR	+	6024	45, and unpublished data
21	S	SL3-3 wt	+	6932	51
22	S	Akv1-99 Runx	+	7024	27, 71
23	T	SL3-3 Turbo	+	7065	20, 51
24****	S	SL3-3 GR+$E_{a/s}$	+	7068	17
25	S	SL3-3 wt	+	7070	51
26****	S	SL3-3 GR+$E_{a/s}$	+	7086	17
27	S	SL3-3 (AkvIN)	+	7267	Unpublished data

[a] Integrations that have been identified at more than one position in the Gfi1 3'UTR from different purification rounds are here considered as independent integration events and indicated by asterisks, where the number indicate that the integrations are derived from the same tumor.
[b] T, thymus; S, spleen; ML, mesenteric lymph node
[c] NMRI-I mice were infected with wt SL3-3 (51), Akv (43, 70), and RFB (unpublished data) MLVs as well as several SL3-3 and Akv mutated in host transcription factor binding sites; Runx (26, 69), UCR (43, and unpublished data), E_{gre} and $E_{a/s}$ (17, 70), Turbo (2Δ18-3) (20, 51), GRE (glucocorticoid response element) (17, 70), SL3-2Env (SL3-3 envelope replaced with SL3-2 envelope) (unpublished data), AkvIN (SL3-3 integrase replaced with Akv integrase) (unpublished data), Akv1-99 (single enhancer repeat variant of Akv) (43).
[d] Integrated virus position in the same (+) transcriptional direction as *Gfi1*
[e] Retrovirus integration position from *Gfi1* transcriptional start site. The *Gfi1* 3'UTR is positioned at 6606 to 7690 from the *Gfi1* transcriptional start site.

TABLE 2. Frequency of Gfi1 alternative splicing[a]

Alternative transcript	Integrations in Gfi1 3'UTR (No. Of tumors with alternative transcript/ No. of tumors analysed)			Integrations outside Gfi1 3'UTR (No. Of tumors with alternative transcript/ No. of tumors analysed)		
	SL3-3	Akv	RFB	SL3-3	Akv	RFB
1	7/14	1/2	0/1	1/19	0/4	0/0
2	14/14	2/2	1/1	16/19	1/4	0/0
3	7/14	1/2	0/1	7/19	0/4	0/0

[a] Number of tumors containing the different alternative transcripts form a total of 40 tumors analysed.