Retroviral activation of the mir-106a microRNA cistron in T lymphoma

Amy M Lum¹, Bruce B Wang¹, Lauri Li¹, Namitha Channa¹, Gabor Bartha² and Matthias Wabl*³

Address: ¹Picobella, L.L.C., 863 Mitten Road, Suite 101, Burlingame, CA, 94010, USA, ²Synergenics, L.L.C., 863 Mitten Road, Suite 101, Burlingame, CA, 94010, USA and ³Department of Microbiology and Immunology, University of California, San Francisco, CA, 94143-0414, USA

Email: Amy M Lum - amy@picobella.com; Bruce B Wang - bruce@picobella.com; Lauri Li - lauri@picobella.com; Namitha Channa - namitha@picobella.com; Gabor Bartha - gabor@synergenics.net; Matthias Wabl* - mutator@ucsf.edu

* Corresponding author

Abstract

Retroviral insertion into a host genome is a powerful tool not only for the discovery of cancer genes, but also for the discovery of potential oncogenic noncoding RNAs. In a large-scale mouse T lymphocyte tumor screen we found a high density of integrations upstream of the mir-106a microRNA cistron. In tumors containing an integration, the primary transcript encoding the mir-106a cistron was overexpressed five to 20-fold compared with that of control tumors; concomitantly, the mature mir-106a and mir-363 microRNAs were highly overexpressed as well. These findings suggest the mir-106a cistron plays an important role in T cell tumorigenesis.

Findings

Retroviral insertions into the genome of a host can induce tumor formation by altering gene expression or function. Integration of a retrovirus near a gene can induce overexpression of the gene through the viral promoter or enhancer, while insertion of a retrovirus into a gene can cause both activation and inactivation. If the affected genes are proto-oncogenes or tumor suppressor genes, the insertion events may lead to tumor formation [1]. Consequently, retroviral mutagenesis has been used to search entire genomes for genes involved in cancer development [2-4], including oncogenic microRNAs (miRNAs) [5]. MiRNAs are short (~22 bp) noncoding RNAs that are implicated in gene regulation and cancer [6-10]. In a large-scale retroviral insertion mutagenesis screen, we used the murine leukemia virus (MLV) strain SL3-3, which causes T lymphomas [11], and identified several miRNAs that are potentially involved in tumorigenesis.

We previously demonstrated that a group of these retroviral insertions induces overexpression of the oncogenic mmu-mir-17 miRNA cistron in mouse tumors [5]. Here we build on our validation of the retrovirus insertional mutagenesis method to identify oncogenic miRNA and present another potentially oncogenic miRNA cistron, mmu-mir-106a. In this screen, male BALB/c mice were treated with ethyl-nitroso-urea (ENU) and bred to normal female mice. ENU treatment was conducted to increase the recovery of tumor suppressors in the F1 progeny through mutagenesis of the paternal allele. Newborn offspring mice were then injected with MLV strain SL3-3. After becoming moribund due to tumor development, mice were euthanized and thymus and spleen tissues were collected and stored at -80°C. Locations of the SL3-3 virus integration sites were identified as previously described using a splinkerette based PCR method [3] that amplifies genomic DNA flanking the 5' LTR of the virus.
We identified 6234 integration sites in 2199 tumors; of these tumors, 76 sites were located on chromosome X upstream of a miRNA cluster containing mmu-mir-106a, mmu-mir-20b, mmu-mir-19b-2, mmu-mir-92-2, and mmu-mir-363. The locations of the integrations ranged from 1.5 kb to 22 kb upstream of the miRNA cluster (Fig. 1), with proviral inserts in both sense and anti-sense orientations with respect to the primary RNA transcript encoding the miRNA cistron. The Mouse Retroviral Tagged Cancer Gene Database [12], which compiles retroviral insertions into the genomic DNA from various non-T cell derived mouse tumors, also lists 10 integrations located upstream of the mmu-mir-106a cluster. Furthermore, Hwang et al. found that EST AI464896, which maps to the same location as mmu-mir-363, was overexpressed in tumors with proviral MLV integrations into this region [13]. The radiation leukemia virus (RadLV) also frequently integrates at this locus and a group of five differentially expressed noncoding RNAs known as Kis2 (GenBank Accession numbers AY940614-AY940618) are overexpressed in these tumors [14]. Because the Kis2 transcripts lie directly upstream of the mir-106a miRNA cluster (mmu-mir-106a overlaps these transcripts by four bases), they likely are part of the primary transcripts containing the miRNA cluster.

To determine whether the retroviral integrations in this region affected the expression of the mir-106a cistron, we used quantitative PCR (qPCR) to measure expression levels of the primary transcript (Kis2) and the mature miRNAs (mmu-mir-106a and mmu-mir-363) in tumors containing mir-106a cistron integrations as well as in control tumors lacking such integrations. To measure primary transcript (Kis2) expression levels, a probe and primer set was designed to AY940616, which is a common exon to three of the alternatively spliced forms of Kis2. The probe and primers for AY940616 were as follows: 5’TGCAGTCTGCCAGGGTGTCGGCAGGCT-3’, 5’GGTACCGGCTCGCCCTGG-3’, and 5’GGTACCGGCTCGCCCTGG-3’. The mature species of mmu-mir-106a and mmu-mir-363 were respectively 67% and 69%. Concentrations of the mature species were calculated using the SOFTMIX software released by Applied Biosystems.

Representative tumors with integration sites spanning the upstream region of mir-106a were measured for expression of the miRNA primary transcript (Fig. 1 and Table 1). In 16 of the 21 tumors assayed, expression of AY940616 was elevated five to 20 fold as compared to the average expression of tumors with no integrations at this locus (Fig. 2A). This confirms the previous report that proviral integrations in this region can increase expression of the Kis2 locus [14].

The mature species of mmu-mir-106a and mmu-mir-363 were then measured by RT-qPCR using a stem-loop RT primer specific for each miRNA [15]. Accordingly, 50 ng of each tumor miRNA preparation was reverse transcribed with the SuperScript First-Strand Synthesis System for RT-PCR using the following stem loop RT primers (50 nM final concentration): 5’GGTACCGGCTCGCCCTGG-3’. The reverse transcription reactions were diluted 1:200 and 5 μl of these dilutions were used in the 25 μl qPCR reactions. The annealing step was 50°C for 60s. The qPCR probes and primers were as follows: mmu-mir-106a: 5’CCGCAAAGTCTGAACAGTG-3’, 5’-6-FAM-5’-TCACGGAATACGGTGCCGTCGAGGCT-3’, mmu-mir-363: 5’-6-FAM-5’-ATCCGAGGTATTTACGGCAATACAGT-3’. Synthetic RNA oligos (IDT) were used to generate a calibration curve for each miRNA: 5’-CGGCAAAGTCTGAACAGTG-3’ and 5’-6-FAM-5’-TCACGGAATACGGTGCCGTCGAGGCT-3’.

Fifteen tumors with integrations in this region were assayed by qPCR for the mature species of mmu-mir-106a and mmu-mir-363. Approximately 70% of these tumors had increased expression levels of mmu-mir-106a by two to six fold, and of mmu-mir-363 by four to 12 fold over the average expression of tumors with no integrations in this region (Fig. 2B). The mature miRNA expression differences between tumors with integrations in this region and the tumor controls was statistically significant (p < 0.00001 (mmu-mir-106a) and p < 0.00001 (mmu-mir-363)) by a two sample unequal variance Student’s t test. From these data we conclude that retroviral integrations
Map of integration sites upstream of the mmu-mir-106a cistron. A map of the SL3-3 retroviral sites upstream of the mir-106a cluster is displayed using the UCSC genome website browser (February 2006 version of the mm8 genome assembly). Insertion sites are depicted as vertical handlebars. Tumors assayed by quantitative PCR are numbered and noted in black text.
Expression of the primary transcript and mature species of the mmu-mir-106a cistron. Quantitative PCR data for tumors with integrations upstream of the mir-106a cistron. (A) Relative expression of AY940616 in tumors containing integration sites near the mir-106a miRNA cluster. Control tumors contain integration sites at locations in the genome other than the mir-106a region. "N" is cDNA generated from normal mouse spleen RNA (Ambion). (B) Relative expression of the mature species of mmu-mir-106a and mmu-mir-363 in tumors containing integration sites near the mir-106a cistron. Tumors are numbered as in Figure 1.
tion from cells of mice overexpressing c-Myc [7]. The mir-17 cluster is also overexpressed in human lung cancer [18]. However, in breast cancer cells, mir-17-5p expression is decreased; there it acts as a translational repressor of the oncogene AIB1 (amplified in breast cancer 1) [19], and in this context may formally act as a tumor suppressor.

It is well established that tumorgenesis is the result of accumulating several cooperating mutations that drive relentless proliferation and aid in metastases. Viral insertion mutagenesis, though perhaps not providing all the mutations necessary for a full-blown tumor, follows this multistep scenario. Although in general the superinfection barrier largely prevents multiple proviral integrations within the same cell, re-infection does happen over time. Because it is a rare event, such cells are selected over the others only when these integrations also give a growth advantage. As a consequence, in general, most viral insertions (“co-mutations”) in a single tumor are thought to be causative in its formation. With the caveats of potential passenger genes and potential oligoclonality of tumors, co-mutation analysis may be a powerful way to find cooperating signaling pathways in tumorigenesis.

We detected multiple insertion sites in all of the tumor samples we assayed from the mir-106a cluster. Genes near common co-integration sites for these tumors include Ahi1, Evi5, and Gfi1, candidates previously appearing in retroviral screens [12], as well as PVT1, a noncoding RNA frequently amplified with myc [20]. A summary of all integration sites in the assayed tumors is listed in Table 2.

Through retroviral insertion in the mouse, we have discovered another potentially oncogenic microRNA cluster, mir-106a-363. Retroviral insertion caused significant overexpression of this microRNA cluster indicating its role in tumor development. This study further demonstrates the power of retrovirus insertion as a tool to discover new oncogenic noncoding RNAs.

### Competing interests
The authors declare a financial interest in Picobella, LLC.

### Authors’ contributions
AML carried out the RNA isolation, quantitative PCR, expression data analysis, and drafted the manuscript. GB, LL, NC, and BBW carried out the tag recovery and identification. BBW and MW planned and directed the execution of the retroviral screen, the design of the study, and drafted the manuscript. GB, MB, ML, and NC performed the experiments. MB and NC contributed with valuable comments on the manuscript.

### Acknowledgements
This work was supported by NIH grant CA100266 to MW, and Synergenics, LLC. We thank Dr. Clifford Wang for his technical advice and for his comments on the manuscript.
Table 2: Summary of integrations in tumors assayed for the mmu-mir-106a cistron

| Tumor # | Tumor Name | Location | Orientation | Gene located near insertion site |
|---------|------------|----------|-------------|---------------------------------|
| 1       | 1759S-158-18 | chr18:78192824 | G+T+ | XM_973419 |
| 1       | 1759S-158-18 | chr7:73480330 | T-G+ | XM_978127.1 |
| 1       | 1759S-158-18 | chr5:107965444 | G-T+ | Gfi1 Growth factor independent 1 |
| 2       | 1890S-76-1 | chr10:20756607 | G+T- | Ahi1 Jouberin |
| 2       | 1890S-76-1 | chr5:61868694 | G+T+ | PVT1 Plasmacytoma variant translocation 1 |
| 3       | 718S-68-23 | chr17:29126185 | G+T+ | Fgd2 FYVE, RhoGEF and PH domain containing 2 |
| 3       | 718S-68-23 | chr11:5819233 | G-T- | Gck Glucokinase |
| 3       | 718S-68-23 | chr11:66037536 | G-T- | Gm879 Gene model 879 |
| 4       | 1890S-76-1 | chr10:20779381 | T+G- | Ev5 Ecotropic viral integration site 5 |
| 5       | 818S-36-8 | chr18:5348771 | G-T- | Zfp438 Zinc finger protein 438 |
| 5       | 818S-36-8 | chr15:5380826 | T-G+ | Cacna2d1 Calcium channel, voltage-dependent, alpha2/delta subunit 1 |
| 6       | 535S-122-16 | chr10:207975997 | G-T- | Gfi1 Growth factor independent 1 |
| 6       | 535S-122-16 | chr15:62006727 | G+T+ | PVT1 Plasmacytoma variant translocation 1 |
| 7       | 1759S-158-17 | chr5:115421225 | G-T+ | Hypothetical protein LOC109154 |
| 7       | 1759S-158-17 | chr10:20798359 | G-T+ | Gfi1 Growth factor independent 1 |
| 7       | 1759S-158-17 | chr17:46990647 | G-T- | Tbn Taube nuss |
| 8       | 1890S-187-12 | chr5:107970686 | G-T+ | Gfi1 Growth factor independent 1 |
| 8       | 1890S-187-12 | chr9:20808975 | G-T+ | Tyk2 Tyrosine kinase 2 |
| 8       | 1890S-187-12 | chr15:63293889 | T+G- | XM_139402 |
| 9       | 3390T-128-37 | chr2:117124415 | G-T- | Acsl5 Acyl-CoA synthetase long-chain family member 5 |
| 9       | 3390T-128-37 | chr17:29125847 | G+T+ | Rasgrpl RAS guanyl releasing protein 1 |
| 10      | 49S-41-2 | chr2:72016508 | G-T+ | Rras2 Related RAS viral (+ras) oncogene homolog 2 |
| 10      | 49S-41-2 | chr17:29125847 | G+T+ | Spon1 Spondin 1, (f-spondin) extracellular matrix protein |
| 10      | 49S-41-2 | chr17:29125847 | G+T- | Ahi1 Jouberin |
| 11      | 544S-136-23 | chr3:6980695 | G-T+ | Jundm2 Jun dimerization protein 2 |
| 11      | 544S-136-23 | chr10:120923128 | T+G- | Notch1 Notch gene homolog 1 (Drosophila) |
| 11      | 544S-136-23 | chr10:20779700 | G-T+ | XM_988509 |
| 12      | 49S-41-2 | chr2:72016508 | G-T+ | Rapgef4 Rap guanine nucleotide exchange factor (GEF) 4 |
| 12      | 49S-41-2 | chr17:29125847 | G-T+ | Tbn Taube nuss |
| 13      | 49S-41-2 | chr3:6980695 | G-T+ | Sics372 Solute carrier family 37 (glycerol-3-phosphate transporter), member 2 |
| 14      | 49S-41-2 | chr2:72016508 | G-T+ | Jundm2 Jun dimerization protein 2 |
| 15      | 544S-136-23 | chr10:120923128 | T+G- | Notch1 Notch gene homolog 1 (Drosophila) |
| 15      | 544S-136-23 | chr10:20779700 | G-T+ | XM_988509 |
| 16      | 49S-41-2 | chr2:72016508 | G-T+ | Acsl5 Acyl-CoA synthetase long-chain family member 5 |
| 16      | 49S-41-2 | chr17:29125847 | G-T+ | Rasgrpl RAS guanyl releasing protein 1 |
| 17      | 49S-41-2 | chr17:29125847 | G-T+ | Spon1 Spondin 1, (f-spondin) extracellular matrix protein |
| 17      | 49S-41-2 | chr17:29125847 | G-T+ | Ahi1 Jouberin |
| 18      | 1469S-150-34 | chr3:6980695 | G-T+ | Jundm2 Jun dimerization protein 2 |
| 18      | 1469S-150-34 | chr10:20779700 | G-T+ | XM_988509 |

Additional retroviral insertion sites recovered from tumors containing an insertion site upstream of the mir-106a cistron. Retroviral insertion site locations and orientations are notated as in Table 1. Nearby genes to the insertion sites are also listed.
References

1. Uren AG, Kool J, Berns A, van Lohuizen M: Retroviral insertional mutagenesis: past, present and future. Oncogene 2005, 24:7656-7672.

2. Suzuki T, Shen H, Akagi K, Morse HC, Malley JD, Naiman DQ, Jenkins NA, Copeland NG: New genes involved in cancer identified by retroviral tagging. Nat Genet 2002, 32:166-174.

3. Mikkers H, Allen J, Knipscheer P, Romeijn L, Hart A, Vink E, Berns A: High-throughput retroviral tagging to identify components of specific signaling pathways in the P53 tumor suppressor gene. Nat Genet 2002, 32:153-159.

4. Li J, Shen H, Himmel KL, Dupuy AJ, Largaespada DA, Nakamura T, Shaughnessy JD Jr, Jenkins NA, Copeland NG: Leukemia disease genes: large-scale cloning and pathway predictions. Nat Genet 1999, 23:348-353.

5. Wang CL, Wang BB, Bartha G, Li L, Channa N, Klinger M, Killeen N, Wab M: Activation of an oncogenic microRNA cluster by provirus integration. Proc Natl Acad Sci U S A 2006, 103:18680-18684.

6. Shamovsky I, Ivanikov M, Kandel ES, Gershon D, Nuñez R: RNA-mediated response to heat shock in mammalian cells. Nature 2006, 440:556-560.

7. He L, Thomson JM, Hemann M, Hernandez-Monge E, Lu D, Goodson S, Powers S, Carda C, Lowe SW, Hannon GJ, Hammond SM: A microRNA polycistron as a potential human oncogene. Nature 2005, 435:828-833.

8. Calin GA, Sevignani C, Duhit CD, Hyslop T, Noch E, Yendamuri S, Shimizu M, Rattan S, Bullrich F, Negrini M, Croce CM: Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. Proc Natl Acad Sci U S A 2004, 101:2999-3004.

9. O’Donnell KA, Wenzel EA, Zeller KL, Dang CV, Mendell JT: c-Myc-mediated microRNAs modulate E2F1 expression. Nature 2003, 435:839-843.

10. Costinean S, Zanesi N, Pekarsky Y, Tili E, Volinia S, Amin S, Peifer M, Clurman BE: Identification of oncogenes collaborating with p27kip1 loss by insertional mutagenesis and high-throughput insertion site analysis. Proc Natl Acad Sci U S A 2002, 99:11293-11298.

11. Landais S, Quantin R, Rassart E: Activation of an oncogenic microRNA cistron by provirus integration. Proc Natl Acad Sci U S A 2002, 103:7024-7029.

12. Chen C, Ridzon DA, Broome AJ, Zhou Z, Lee DH, Nguyen JT, Barbisin M, Xu NL, Mahuvakar VR, Andersen MR, Lao KQ, Livak KJ, Guegler KJ: Real-time quantification of microRNAs by stem-loop RT-PCR. Nucleic Acids Res 2005, 33:e179.

13. Tanzer A, Stadler PF: Molecular evolution of a microRNA cluster. J Mol Biol 2004, 339:327-335.

14. Volinia S, Calin GA, Liu CG, Ambis S, Cimmino A, Petrocca F, Vison R, Iorio M, Roldo C, Ferracin M, Prueitt RL, Yandahara N, Lanza G, Scarp A, Vecchione AN, Negri M, Harris CC, Croce CM: A microRNA expression signature of human solid tumors defines cancer gene targets. Proc Natl Acad Sci U S A 2006, 103:2257-2261.

15. Hayashita Y, Osada H, Tatematsu Y, Yamada H, Yanagisawa K, Tomida S, Yatabe T, Kawahara K, Sekido Y, Takahashi T: A polycistronic microRNA cluster, miR-17-92 is overexpressed in human lung cancers and enhances cell proliferation. Cancer Res 2005, 65:9628-9632.

16. Hossain A, Kuo MT, Saunders GF: Mir-17-5p Regulates Breast Cancer Cell Proliferation by Inhibiting Translation of AIB1 mRNA. Mol Cell Biol 2006, 26:8191-8201.

17. Shibelman E, Bishop JM: The PVT gene frequently amplifies with MYC in tumor cells. Mol Cell Biol 1989, 9:1148-1154.