Use of the Gene Trap Resource for Cancer-related lncRNAs to Study the Role of Malat1 in Pancreatic Cancer.

Andrei Golovko[1], Huiping Guo[1], Amy Gonzales[1], Stephen H. Safe[2], Indira Jutooru[2], Parisa Imanirad[2], and Benjamin Morpurgo[1]

[1]Texas A&M Institute for Genomic Medicine; [2]Texas A&M College of Veterinary Medicine, Dept. of Physiology & Pharmacology. College Station, TX, USA.

Abstract

Texas A&M Institute for Genomic Medicine (TIGM) houses the world's largest library of knockout C57BL/6N ES cells and provides transgenic mice for researchers worldwide. Among multitude of genetic targets inactivated in these cells, there is a significant group of long non-coding RNAs (lncRNA), the non-protein coding transcripts longer than 200 nucleotides implicated in a variety of disease states and demonstrated their involvement in oncogenesis. Our screening of more than 18,000 clones has identified over 1,000 inactivated ncRNAs including a number of lncRNAs. One such clone, IST14461G11, was used to establish a colony of homozygous mutant Metastasis-Associated-Lung-Adenocarcinoma-TranScript-1 (Malat1) mice in pure C57BL/6N genetic background. Malat1 is an lncRNA that is overexpressed in multiple cancer cell lines and tumors. The highly conserved mouse homologue of Malat1 was found to be highly expressed in hepatocellular carcinoma and now we have demonstrated that Malat1 is also pro-oncogenic in pancreatic cancer cells. Homozygous mutants are viable and don't display any gross phenotype. The effects of loss of Malat1 expression on pancreatic tumor formation will be determined.

Introduction

With the advent of high throughput genome sequencing and annotation technologies, the increasing significance of non-protein encoding RNAs (ncRNAs) in cellular homeostasis and disease has become apparent [1]. MicroRNAs (miRs) are small non-coding RNAs containing 21 – 23 nucleotides and the >100 miRs primarily repress gene expression, and it has been estimated that miRs regulate up to 50% of all mRNAs. Long non-coding RNAs (lncRNAs) contain >200 nucleotides and recent reports from the ENCODE Project Consortium have identified 9277 lncRNA genes that produce 14,880 transcripts [2].

Metastasis-Associated-Lung-Adenocarcinoma-TranScript-1(Malat1) is a lncRNA that is overexpressed in multiple cancer cell lines and tumors, and Malat1 expression is a prognostic factor for decreased survival of stage 1 non-small cell lung cancer (NSCLC). Malat1 expression is also associated with metastasis in NSCLC patients and Malat1 expression is correlated with• poor prognosis (survival/recurrent/metastasis) in squamous cell carcinoma of the lung, hepatocellular carcinoma, bladder and colorectal cancer. Moreover, functional studies determined by Malat1 knockdown or overexpression indicate that Malat1 enhances cell and tumor growth, migration, invasion and epithelial-to-mesenchymal transition. Moreover, we have now demonstrated that Malat1 is also pro-oncogenic in pancreatic cancer cells. The highly conserved mouse homologue of Malat1 was found to be highly expressed in hepatocellular carcinoma [3] and we have identified Malat1 expression and function in pancreatic cancer and our proposed studies will investigate generation of lncRNA knockout mice and Malet1 mice will be crossed with transgenic mouse model expressing KRASG12D with a p53 mutation in the pancreas and 100% of these develop pancreatic tumors.

Technology

The basic gene trap vectors we have used include a reporter gene downstream of a splice acceptor sequence (Fig.1).

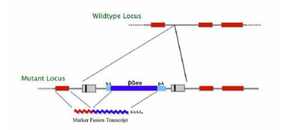


Figure 1. Gene trap vectors used in C57BL/6 library (β-geo version shown). LTR, long terminal repeat; SA, splice acceptor sequence; β-geo, galactosidase/neomycin phosphotransferase fusion gene; pA, polyadenylation sequence .

They are designed to function when inserted in an intron, to produce incorrect splicing of the target gene such that all exons downstream of the insertion site are not expressed. The gene trap cassette is inserted in a retroviral vector. Retroviruses insert as a single copy per locus, with no rearrangement of flanking sequences. They have a preference for insertions at the 5' end of genes, often upstream of the initiator ATG, and the splice acceptor sequence we use does not appear to be bypassed by the RNA-splicing machinery. As a result, the majority of the mutations generated using our gene trap vectors are predicted to lead to null alleles.

References

1. Rinn JL, Chang HY. Genome regulation by long noncoding RNAs. Annu Rev Biochem 81:145-166, 2012.
2. Derrien T. Johnson R. Bussotti G, Tanzer A, Djebali S. Tilgner H. et al. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure evolution and expression. Genome Res 22:1775-1789, 2012. PMCID: PMC3431493.
3. Lin R, Maeda S, Liu C, Karin M, Edgington TS (February 2007). "A large noncoding RNA is a marker for murine hepatocellular carcinomas and a spectrum of human carcinomas". Oncogene 26 (6): 851- 8.

Results

Expression of MALAT-1 in pancreatic cancer cells

Our studies show that MALAT-1 is expressed in pancreatic cancer cells and can be detected in serum from pancreatic cancer patients (Fig. 2) and this correlates with the reported high expression of MALAT-1 in pancreatic tumors compared to non-tumor tissues [3].

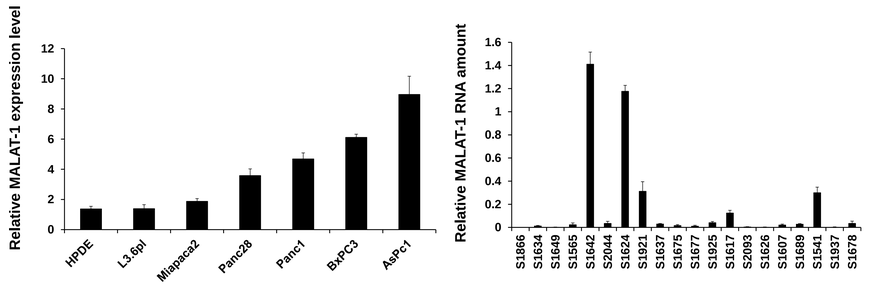


Figure 2. Expression of MALAT-1 in pancreatic cancer cells (A) a

We also investigated the effects of MALAT-1 knock down by RNA interference using two oligonucleo-tides (siMALAT-1a and siMALAT-1b), and the results (Fig. 3A) show that loss of MALAT-1 in Panc1 and Panc28 cells resulted in significant inhibition of cell proliferation compared to cells transfected with a control oligonucleotide (siLamin). Thus, MALAT-1 and MALAT-1-regulated genes play a role in pancreatic cancer cell growth. MALAT-1 knockdown by RNAi also significantly decreased transwell migration in Panc28 and Panc1 cells using a Boyden chamber assay (Fig. 3B).

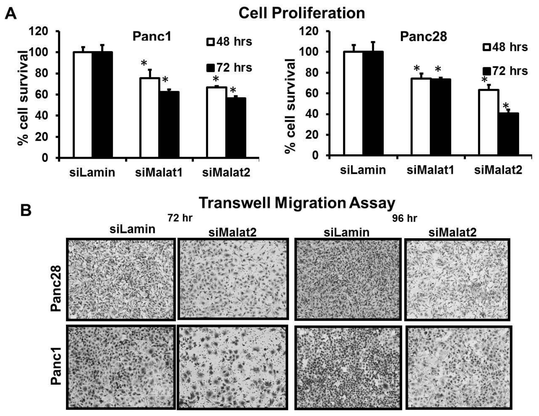


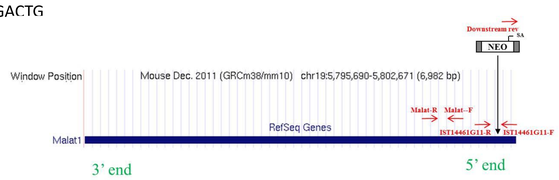
Figure 3. Knockdown of MALAT-1 (siMALAT-1/siMALAT-2) significantly decreases Panc1 and Panc28 cell proliferation (A) and migration (B) in an Boyden Chamber assay. (* Significant inhibition; p<0.005)

Generation of Malat1 knockout

The bioinformatics efforts have revealed that more than 18,000 ES cell clones from our collection have insertions which appear to inactivate 1,000 unique long ncRNAs (those greater than 200bp). Included in these are lncRNAs such as Malat1, Tsix and Air. TIGM C57BL/6N ES cell clone IST14461G11 was found to carry mutation in Malat1 (Figure 4) and was chosen for further work. The clone was expanded and the genomic sequence surrounding the gene trap insertion site was determined as follows (the insertion site is denoted with an asterisk *):

CAGGCATTCAGGCAGCGAGAGCAGAGCAGCGTAGAGCAGCACAGCTGAGCTCGTGAGGCAGGAGACTCAGCCCGAGGAAATCGCAGATA
AGTTTTTAATTAAAAAGATTGAGCAGTAAAAAGAATTAGAACTCTAAACTTAAGCTAATAGAGTAGCTTATCGAAATATTACTTAGTCTTAATAA
TCTAAGAAGATCTTAAGAGATAACATGAAGGCTTATTTAAACAGTTTGAAAAAGGAAATGAGGAGA*AAAGTATTTGTACTGTATAATGGAG
GCTGACCAGAGCAGTTTAGGAGATTGTAAAGGGAGGTTTTGTGAAGTTCTAAAAGGTTCTAGTTTGAAGGTCGGCCTTGTAGATTAAAACG
AAGGTTACCTAAATAGAATCTAAGTGGCATTTAAAACAGTAAAGTTGTAGAGAATAGTTTGAAAATGAGGTGTAGTTTTAAAAGATTGAGAA
AAGTAGGTTAAGTTGACGGCCGTTATAAAAATCCTTCGACTG

Figure 4 Schematic representation of the mutated locus Malat1. IST14461G11-F, IST14461G11-R and Downstream rev are the primers used in genotyping. Malat1-F and Malat1-R are primers used to confirm knockout via RT-PCR.



The IST14461G11 clone was used to establish a colony of homozygous mutant Malat1 mice in pure C57BL/6N genetic background. Mutant mice were generated using standard procedures. In short, a mutant ES cell clone was expanded and microinjected into albino B6 host blastocysts to generate germline chimeras. Those were bred to C57BL/6 females for germline transmission of the mutant Malat1 allele. The correct mutation was confirmed using PCR-based genotyping protocol (Table 2) using primers specific for genomic insertion site and for the vector (Figure 2).

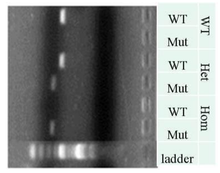
Table 2 Schematic representation of the mutated locus Malat1. IST14461G11-F, IST14461G11-R and Downstream rev are the primers used in genotyping. Malat1-F and Malat1-R are primers used to confirm knockout via RT-PCR.

Primer Sequences (5' to 3'):	
Mutant Oligos: IST14461G11-F + Downstream Rev (337bp)	
Wt Rxn Oligos: IST14461G11-F + IST14461G11-R (468bp)	
IST14461G11-F	AGAGCAGAGCAGCGTAGAGC
IST14461G11-R	TAACGGCCGTCAACTTAACC
Downstream rev	CCAATAAACCCTCTTGCAGTTGC

Generation of Malat1 knockout (Continued)

The heterozygous mice were subsequently bred to obtain homozygous mutants. As shown in Figure 5, the wild type Malat1 amplicon (468 bp) was only detected in Wt and Het animals, whereas the mutant product (270 bp) was amplified in both Null and Het animals.

Figure 5. Genotyping results showing wild type (WT), heterozygous (Het) and homozygous (Hom) mutants using primer sets for wild type (WT) and mutant (Mut) alleles in TIGM clone IST14461G11 for Malat1.



Malat1 knockout confirmation

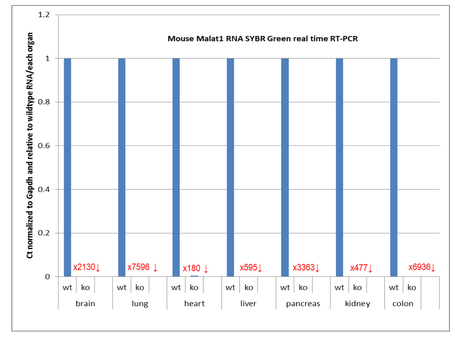
To confirm that gene trap-based inactivation resulted in the anticipated reduction of the gene expression, Malat1 mRNA expression levels in mouse brain, lung, heart, liver, pancreas, kidney and colon were determined by real-time PCR on the Step one plus (Applied Biosystems) instrument. SYBR Green one step Real-time RT-PCR was performed with 20ng total RNA using iScript one-step RT-PCR and SYBR Green mix (BIO-RAD). The assay was performed using a primer pair amplifying a 170 nt fragment of the transcript downstream of the gene trap insertion (Figure 3). Expression levels were normalized to Gapdh expression using primers specific to this gene.

Primer sequences used in this study were as follows:
Malat1-F ggcagaatgcctttgaagag
Malat1-R ggtcagctgccaatgctagt
Gapdh-F ggcattgctctcaatgacaac
Gapdh-R gccatgtaggccatgaggt

As shown in Figure 6, Malat1 expression was completely eliminated in all tissues in homozygous mice compared with wild type siblings. These findings indicate that gene trap insertions in long non-coding RNAs can effectively knock out genes and create true mutant alleles.

Homozygous mutants are viable and don't display any gross phenotype.

Figure 6. Fold change in Malat1 gene expression in knockout mice as compared to the wild type.



Overall, our preliminary data indicates that genomic localization of the gene trap insertions from TIGM library along with mapping of the non-coding transcriptome can serve as a very effective tool to segment out a collection of functional mutations in thousands of murine lncRNAs that, in turn, can be used to produce a repository of novel mutant mouse models for various research areas.

Conclusions and Future Directions

TIGM maintains the world's largest library of stable mouse knockout embryonic stem (ES) cells in the C57BL/6 background, with a total of over 350,000 clones representing more than 10,000 unique protein-coding genes. The library has been analyzed to identify over 1,000 inactivated ncRNAs including a number of lncRNAs. As a proof of concept, we have successfully established a mutant mouse line carrying a homozygous mutation in Malat1, an lncRNA involved in several different types of cancer, including pancreatic. Future efforts will include continuing analysis of the TIGM ES cell collection to discover and verify additional inactivated lncRNAs. Clones with these mutations will be made easily accessible to the scientific community and can be found on the TIGM website using text or sequence searching tools. TIGM is also planning to produce more mutant mouse lines with disrupted lncRNAs of the highest scientific value, most of which will include potential cancer targets, and make them available to the scientific community. Also, we have demonstrated that Malat1 is pro-oncogenic in pancreatic cancer cells. Homozygous mutants are viable and don't display any gross phenotype; therefore, in order to investigate the role of Malat1 in pancreatic cancer, these mutant mice are now being crossed with a transgenic mouse model expressing $KRAS^{G12D}$ and carrying the p53 mutation that effectively develops pancreatic tumors. The effects of loss of Malat1 expression on pancreatic tumor formation will be determined.

For more information please visit www.tigm.org or contact us at info@tigm.org.