THE WISTAR INSTITUTE

Qihong Huang, M.D., Ph.D.
Associate Professor

t: +1-215-495-6835

qhuang@wistar.org

REPORT OF HOST INSTITUTE

1. Name of Professor (under whom training was carried out) : **Dr. Qihong Huang M.D, PhD** Associate Professor

2. Name and address of host institute : **The Wistar Institute,** 3601 spruce street, Philadelphia, PA.

3. Duration of fellowship : 6 months

4. Brief highlights of the achievements :

Dr. Suresh Kumar was involved in the project of noncoding RNA and RNA binding proteins in metastasis process. He made plasmid constructs of noncoding RNAs and cell lines with knockdown of RNA binding proteins. He confirmed the downregulation of RNA binding proteins by immunoblots. These cell lines can potentially be used for investigation of the functions of these RNA binding proteins in tumor metastasis.

5. Your assessment of the ICMR-IF :

Dr. Suresh Kumar is a hard-working scientist. He tried to learn technologies and techniques in my laboratory and core facilities at our institute as much as possible from day one he joined my laboratories. He was also getting familiar with the literatures of noncoding RNA and RNA binding proteins and expanded his general knowledge in this area. Hopefully his training in my laboratory will help his independent research and his future career.

6. Any other comments :

Qihong Huang
Associate Professor
The Wistar Institute
3601 Spruce Street
Philadelphia, PA 19104
U.S.A.

Certificate of Completion

The University of Pennsylvania certifies that

suresh kumar ramakrishnan

has completed the online course/curriculum titled

Species-Specific Hands-on Training -- Mouse - ULAR
(4391315)

on this date: **August 16, 2012**

This course/curriculum satisfies specific training and education requirements as set forth in the policies of the University of Pennsylvania.

Penn
UNIVERSITY of PENNSYLVANIA

Dr. Diane Gaertner
Director
University Laboratory Animal Resources
215-898-2434

Certificate of Completion

The University of Pennsylvania certifies that

suresh kumar ramakrishnan

has completed the online course/curriculum titled

Animal Facility Orientation-ULAR-Hill (4366935)

on this date: **August 1, 2012**

This course/curriculum satisfies specific training and education requirements as set forth in the policies of the University of Pennsylvania.

Penn
UNIVERSITY of PENNSYLVANIA

Dr. Diane Gaertner
Director
University Laboratory Animal Resources
215-898-2434

Certificate of Completion

The University of Pennsylvania certifies that

suresh kumar ramakrishnan

has completed the online course/curriculum titled

Occupational Health and Safety Animal Users Risk Assessment - UNIV
(4367411)

on this date: **July 25, 2012**

This course/curriculum satisfies specific training and education requirements as set forth in the policies of the University of Pennsylvania.

Penn
UNIVERSITY of PENNSYLVANIA

Troy M. Hallman , MS, VMD, DCLAM
IACUC - Director of Animal Welfare
Office of Regulatory Affairs
215-573-2540

Certificate of Completion

The University of Pennsylvania certifies that

suresh kumar ramakrishnan

has completed the online course/curriculum titled

Mouse Training Curriculum - IACUC-ULAR (4366962)

on this date: **August 16, 2012**

This course/curriculum satisfies specific training and education requirements as set forth in the policies of the University of Pennsylvania.

Penn
UNIVERSITY of PENNSYLVANIA

Dr. Diane Gaertner
Director
University Laboratory Animal Resources
215-898-2434

Certificate of Completion

The University of Pennsylvania certifies that

suresh kumar ramakrishnan

has completed the online course/curriculum titled

Mouse Training Pre-requisite - ULAR (4366963)

on this date: **July 25, 2012**

This course/curriculum satisfies specific training and education requirements as set forth in the policies of the University of Pennsylvania.

Penn
UNIVERSITY of PENNSYLVANIA

Dr. Diane Gaertner
Director
University Laboratory Animal Resources
215-898-2434

Introduction

Since 50 years, in practice the term 'gene' was used to denote the sequences that codes for mRNAs that are translated in to protein. Earlier perception of molecular biology focused with undoubted belief on "Central Dogma", i.e. DNA to mRNA to protein, heavily concentrated on protein coding genes and their regulation, functions , establishing their network etc: The central dogma that were developed mainly from E coli, has established with bias that RNA functions, mainly as a informational intermediate between DNA and protein. The well known fact is, the part of genome that has protein coding potential is approximately 1.5% and rest of them are considered as junk DNAs(Kevin wang and Howard chang 2011).

The advent of modern technologies brought great surprises in understanding the complex biological problems. One among them is, the under-estimated, unrevealed, so called "Dark matter/junk DNA " which gave lots of surprises today than previously thought.

From whole genome arrays and sequencing efforts, it was found that 90% of genome get transcribed, which also includes the non coding Junk component of genome (Birney et al 2007) . Recently The HUGO gene nomenclature committee has assigned gene symbol to the genes and out of 30000 genes, 19000 of them were protein coding genes(Seal RL et al 2011, wright ,Bruford 2011). The so called Darlk matter now consist of non protein coding genes (RNA elements) which are transcribed and are linked with certain physiology, function of the cells.

Non coding RNA

One of such Non protein coding element is called non coding RNA, which is classified as small and long non coding RNA based on their transcript size. Small non coding RNAs are called as microRNAs which found to have diverse role in physiology and pathophysiology of the cells.

This class includes well documented miRNAs, which are of ~22 nucleotides (nt) in length and involved in the diverse functional regulation of both protein-coding, and putatively non-coding genes, by post-transcriptional silencing etc.

This small RNAS are found in 5' or 3' coding genes, intronic, exonic regions of a gene. Many of the miRNAs have found to have role in tumorigenesis and tumor suppressive function. Some of the examples of small RNAs are microRNA (miRNA), piRNA, siRNA, promoter associated RNA (PASR/PAR) and few enhancer RNAs.

In contrast there are long non codig RNAS are longer in length randing from 200bp- 17kb. Some of the investigators even reported that their sizes are ranging extensively >2 and some >100 kb , spliced and contain canonical polyadenylation signals [mercer et al 2009, ,Wilusz et al 2009].

Inaddition long noncoding RNAs can be classified according to their location and proximity to protein coding genes in to following cataegories. Sense, antisense, bidirectional, intronic and intergenic (Minna kaikonnen et al 2011 cardiovascular res). The intergenic is recently described based on its dictinctive chromatin signature of actively transcribed genes. These linc RNA have been suggested to guide chromatin modifying complexes to helps in establishing cell type specific epigenetic status (Khalil et al 2009).

The small non coding RNAs undergoes post transcriptional processing by certain proteins and finally found as 22 nt duplex transcript in the cytoplasm. The non coding RNA transcripts due its longer length, there is possibility of protein coding potency within its body of the gene or RNA. Very few reports have established such kind of short peptides could elicit phenotype inspite of presence of non coding RNA.

All RNAs including non coding RNAs they are bound with proteins to preserve the structure and function, stability etc. There are many RNA binding protein to date which may also assist the function of RNA or certain RNA binding protein need such RNAs to do its functions.RNA binding protein have certain RNA binding motifs RNA recognition motif (RRM), K homology (KH) ,cold shock domain (CSD), Zinc finger, PAZ domain etc. The present study interest is in KH domain protein which may be explained and other protein informations maynot be of interest for the present study.

RNA binding protein

RNA binding proteins are proteins that binds with RNAs and recognize single and double stranded RNAs and their 3 dimensional structures. Contacts with RNA binding proteins are mediated by ribose sugar, phosphate groups. Initial studies had identified certain proteins that are bound with pre-RNA or hetro nuclear RNAs are called as hnRNPs. These hnRNPs can stabilize the single strandedness of RNA and facilitate in splicing, pre-mRNA processing, pre RNA packaging etc. These hnRNP proteins that are described as chromatin associated RNA binding protein. Important feature of RNAs are whether they are coding or non coding they all get packaged by RNA binding proteins to protect from degradation and to facilitate the pre mRNA processing.

hnRNP proteins are further classified as major hnRNPs and minor hnRNP proteins. The major hnRNPs are most abundant RNA binding protein in the nucleus and include hnRNPA1(involved in alternative splicing) and hnRNP A2/B1 involved in splicing and trafficking, hnRNPC1/C2 involved in pre-mRNA packaging, splicing, nuclear retention, hnRNP F in splicing, and multifunctional hnRNPK translation regulation, splicing, mRNA stability). The minor class of hnRNP is HuR protein involved in RNA stability.

Among these, Hn RNPK is having KH domain of RNA binding motif that can recognize the both single stranded RNA, single stranded DNA and involves in wide range of processes including transcription, translation, splicing etc. This Kh domain consist of 70 amino acid which bears cetrin signature sequences (I/L/V)IGXXGXX(I/L/V) I the middle of domain. Regarding its function in tumorigenesis, it act as transcriptional activator in cMyc promoter, BRCA1(Thakur et al 2003) , eLF-4E (Lynch et al 2005) and p53 responsive genes (Moumen et al 2005).

Another Kh domain bearing protein is FXR2, a family of FMR which has role in fragile X syndrome. FMR1 gene family includes FXR1, FXR2. The protein coded by these genes harbors 2 KH domain. High level of this FMRP induces translational repression in *invitro* and *invivo.* The mechanism of such translational repression has not been elucidated.

Metastasis

Metastasis is dissemination of cells from primary site to distant organ or a site. Metastasis is process that result from succession of primary tumor evolution, which accumulates genetic mutations, deletion and other aberrations to surmount the physical boundaries and to colonize to new sites.

Many cellular events that accompany the metastasis events include the proliferation, invasion, intravasation, extravasation, dissemination at distant site and proliferation. These events are elicited in response to signaling pathways that are activated from accumulated genetic aberrations and many bio-molecules play role in metastasizing cancer cells. In addition to the established protein molecules, some of the RNA molecules also play a role in suppressing or enhancing the metastatic potential of a cancer cell. For example microRNAs 10b, miR200 family, miR429 have been shown to play role in dynamics of epithelial character of the cell. Likewise long non coding RNAs that assist gene expression as an enhancer or suppressor play role in metastasis process.

The present study aims at characterizing role of an undefined ncRNA in metastasis and role of RNA binding proteins on its function. Recently a ncNRA in the vicinity of Snail1 gene was found to enhance gene expression and the same nc RNA has taken for study for its involvement in metastasis.

Objective:

To study the role of possible short peptides encoded in the ncRNA gene

Identifying the role of RNA binding protein RNPK and FXR2 in metastasis.

Results

To identify the role of ncRNA which were already shown to have role in metastasis, we analyzed the ncRNA for possible peptides coding capabilities. Identification of ORF in the ncRNA, its possible peptide coding ORF were identified, The predicted ORF were mutated in a such way

that peptides/proteins won't be produced. The mutation along the length of ncRNA coding genes were made using concatenating PCR methodology/ site directed mutagenesis method and the mutated sequences were cloned in lentiviral vectors and sequenced.

The nc RNA sequence that are found to be at vicinity of snail1 gene was characterized to be regulating the adjacent genes. The sequence is as follows.

5'ctgtggcagagacggagaagatgaacagggattttataccaggcgtcagaagggaaccagtgctaaagaaaatgaaaacaccaggccggga
gaggcagctggcatggcgggccgtggtggttttacgtggccgatttgagagagtgagacccctggggtcttggagccaggcctgggaaaagctacttc
acgtcaggccaggggctgtagccctggcaacctccactccgcctggaaatcctccacctcggggcctctctttgcccagacctggcccaggaggagc
acatgggagccgggaccttcccaacaatccttgccgttggctccacaaacctcagccagtcctgcaacctgggatgccttttccaccaggatgcctgc
tactgtcactgttgtcattagataattaatgaactataattagaaatcatatcaataaaatttcacagtctaaggctgt-3'

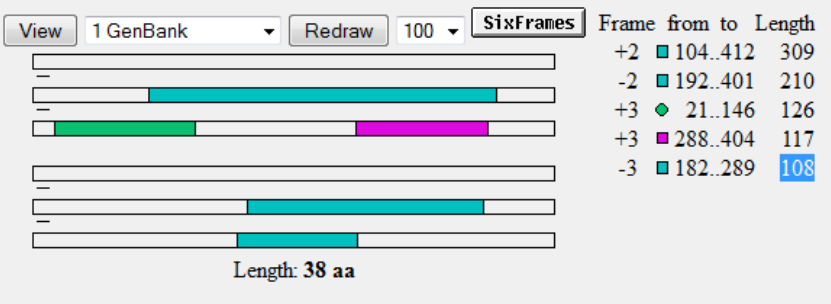

Fig1: Identification of ORF I ncRNA.The color in the length of long rectangles shows the ORFs.Mutations were created on the length of gene.

The PCR products and their respective mutated fragments were purified and cloned under CMV promoter vectors.

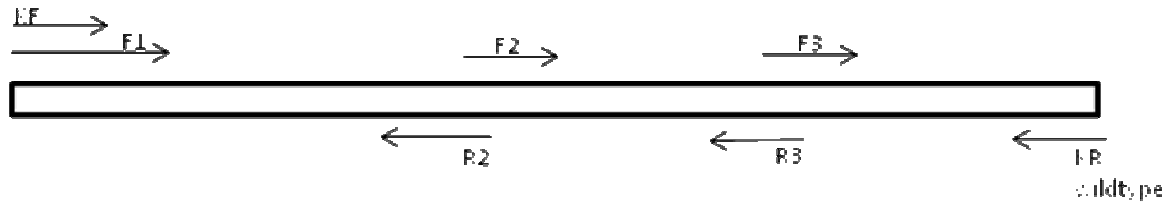

Fig2 : The scheme of the concatenating PCR that introduces mutation along the length of the possible ORF of the gene. KF, KR are wildtype forward and reverse full length primers. F2, F3 are forward mutant primers and R2, R3

are reverse primers.PCRs were performed with F1 and KR, KF with R2 and F2 with KR and F3 with KR, KF with R3.

The PCR products of Different primer combination were individually amplified and represented as below. The full length product of 476 bp was amplified with internal primers with mutations in the primers and mutant products were combined together to get final full length product.

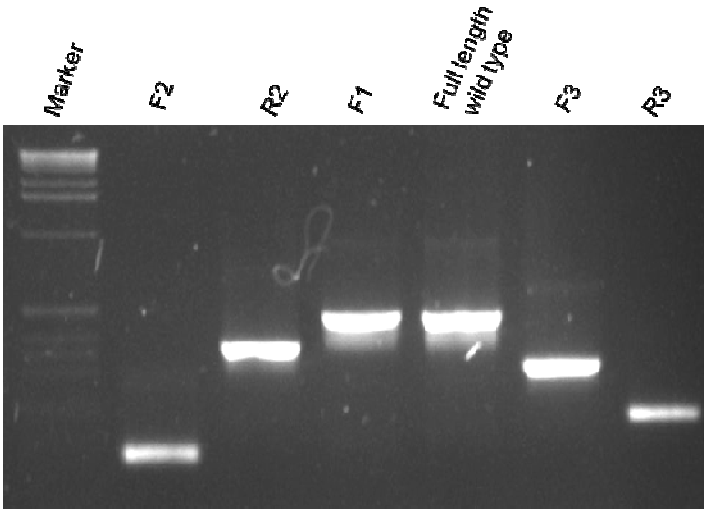

Fig 3: PCR products of mutaed ncRNA. The lane markings are respective to the PCR scheme depicted above.

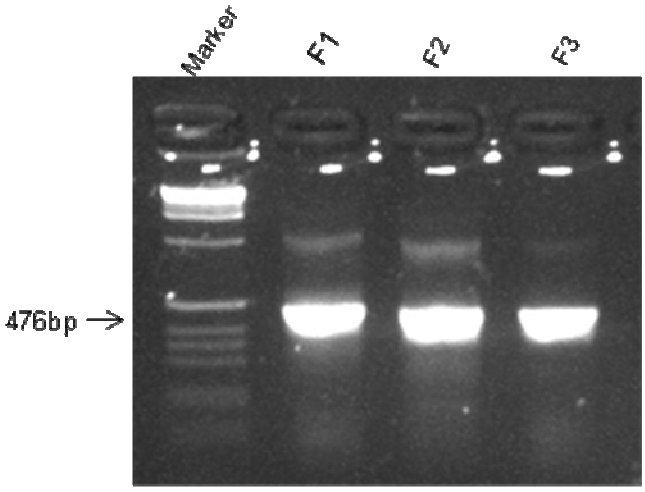

Fig 4: The PCR of the concatenated PCR products with mutated sequence to the full length of the ncRNA gene.

The full length PCR products were amplified from concatenation of PCR products and utilizing them as template for full length amplifications.

The amplified products were purified and cloned under the pFu –GFp vector bearing GFP gene under CMV promoter. The GFP gene was removed and utilized the site for cloning and expression purposes.

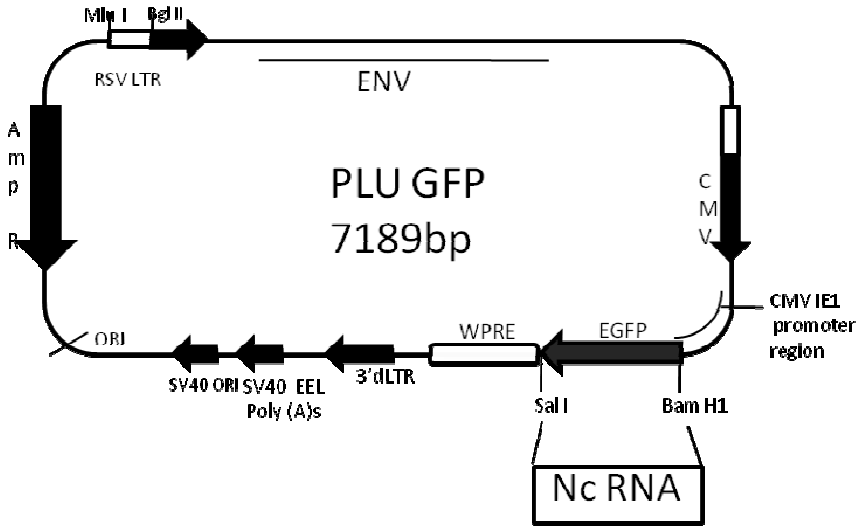

Fig 5: Cloning of mutant of F1, F2, F3 under CMR promoter of pLU-GFP lenti vector. The part of fragment was removed and mutant encoding nc RNA were introduced.

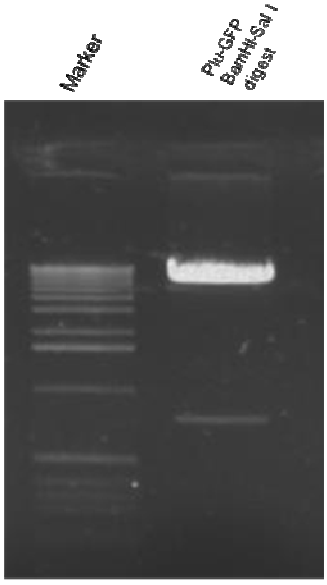

Fig 6: Removal of GFP fragment from the vector and linearizing the vector backbone for cloning the mutant products.

.

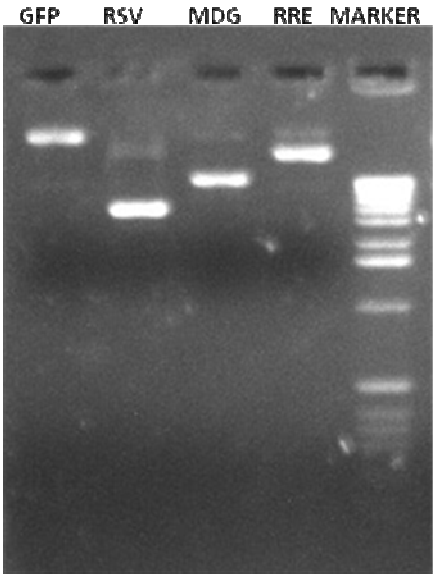

Fig 7: Plasmid isolation of packaging vectors RSV,MDG,RRE

The cloned plasmids were identified by restriction enzyme digestion and plasmids with insert were sent for sequencing. The agarose gel electrophoresis showing the plasmids of other packaging vectors which are used for further experiments. The sequences confirmed the mutations.

Once the plasmids bearing mutants were identified, the same were transducted using lentiviral expression system in MCF 7 cells. The cells were expanded in 10cm plate and were scraped for further analysis. Total RNA was isolated and checked for the expression of transducted sequences, using primers specific for the ncRNA. As a control un-transfected MCF cells were used in parallel. The agarose gel electrophoresis of RT PCR product shows the amplification of ncRNA with specific primers and confirmed the expression of the transducted sequences.

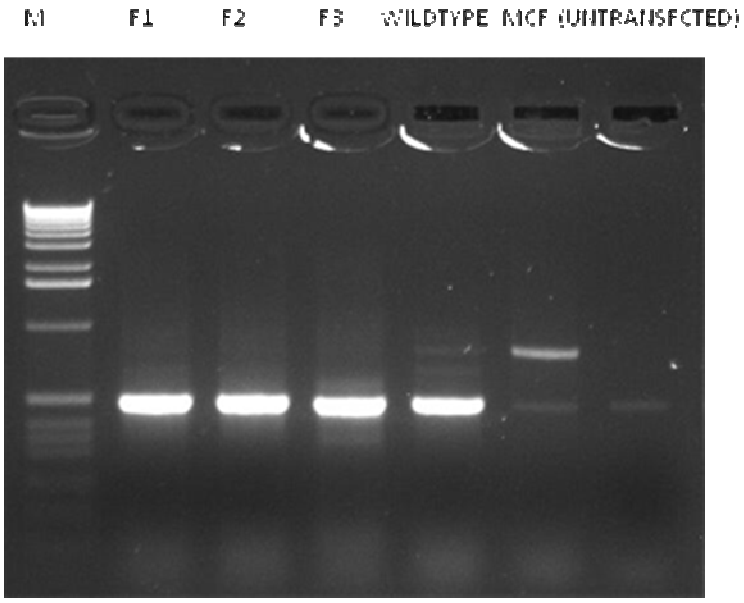

Fig 8: Expression of mutant and wildtype ncRNA in less invasive MCF cells by RT PCR.The F1 and F2, F3 are mutants of ncRNA and wildtype is wiltype ncRNA tranducted ,and untransfected MCF as control MCF

Next we try to access the effect of ncRNA and its mutants in MCF 7 cells. The cells transducted with the mutant and wild type Cells were analysed for E cadherin expression. This ncRNA earlier were identified that it has regulatory role in E cadherin expression and subsequently metastasis. The expression studies using western blotting on E-cadherin shows that mutant ncRNA doesn't affect the level of E-cadherin when compared with wild type ncRNA.

The westernblotting below shows the expression of E-cadherin. The same experiment was repeated twice with fresh transduction of the mutants. Both the times it had shown there was enhancement in E cadherin expression and there was no suppression or change in E cadherin expression in expressed cells. This experiment rule out the possibility of peptide coding capability of ncRNA and whatever the observed effect are elicited by ncRNA and not by predicted ORF or possible short peptides.

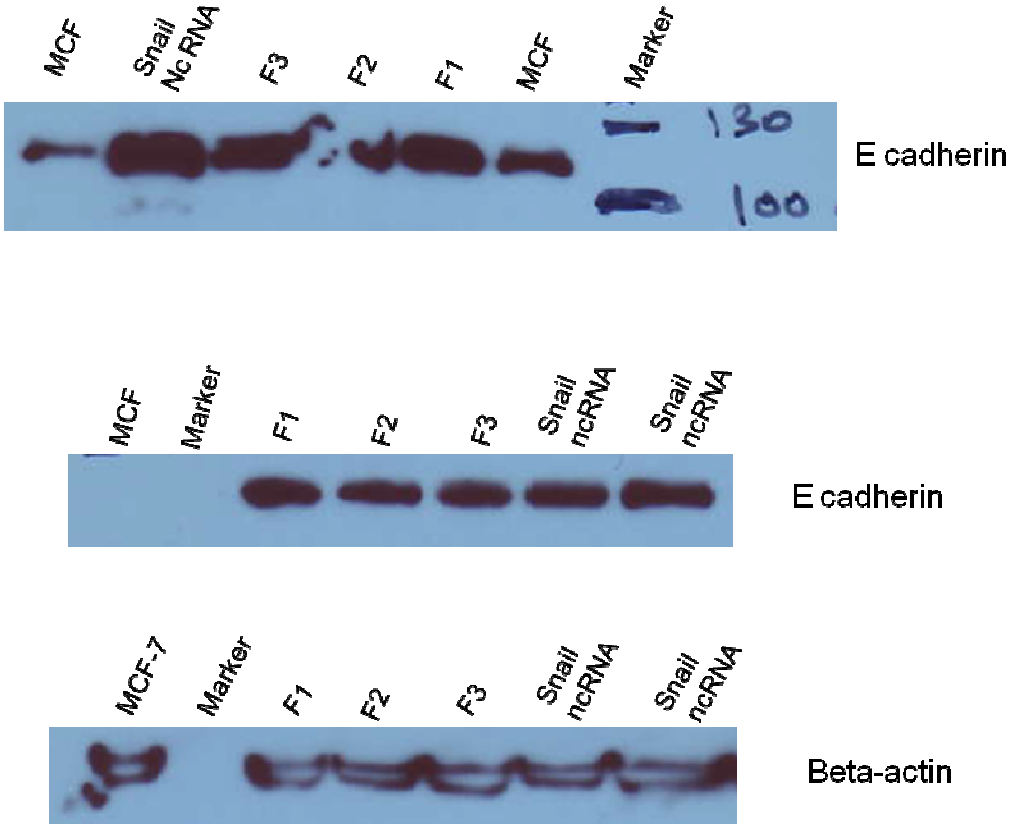

Fig 9: Effect of E cadherin expression upon ncRNA expression, The ncRNA controls the expression of E cadherin. Western blotting showing no differences in E cadherin expression in mutant expressing ncRNA compared to wild type and it enhances the expression of E-cadherin. The second wetsrn blot of e cadherin is repetition of transduction and expression studies.

Objective 2:
The Role of RNA Binding Proteins in Eliciting the Metastasis Process.

Next we try to study the role of RNA binding protein in eliciting the response in metastasis. This RNA binding protein hnRNPK and FXR2 were identified that they bind with snai 1 ncRNA and they may assist the function of ncRNA or protein itself elicit the response by binding with ncRNA. To study the role of RNA binding proteins the MDMBA cells were infected with viruses expressing the shRNA against the hnRNPK and FXR2. The knockdown effect of hnRNPK and FXR2 was assessed by western blotting.

Fig 10: Western blotting of hnRNPK and FXR2 proteins in knockdown cells.The blots confirms the knockdown of hnRNPK and FXR2.

The western blotting shows the knockdown of hnRNPK and FXR2 in MD MBA cells. The infected cells were further expanded in 10 or 15 cm plates and cells were washed twice with PBS having pencilin and streptomycin antibiotic and trypsinized. Again the cells were washed to remove t etrypsin and pellets were gently diluted with DMEM media with serum.

Due to time constraints and ethical issues the mouse model experiments couldn't be performed. However the cells infected with hnRNPK and FXR2 were given for intravenous/tail vein injection.

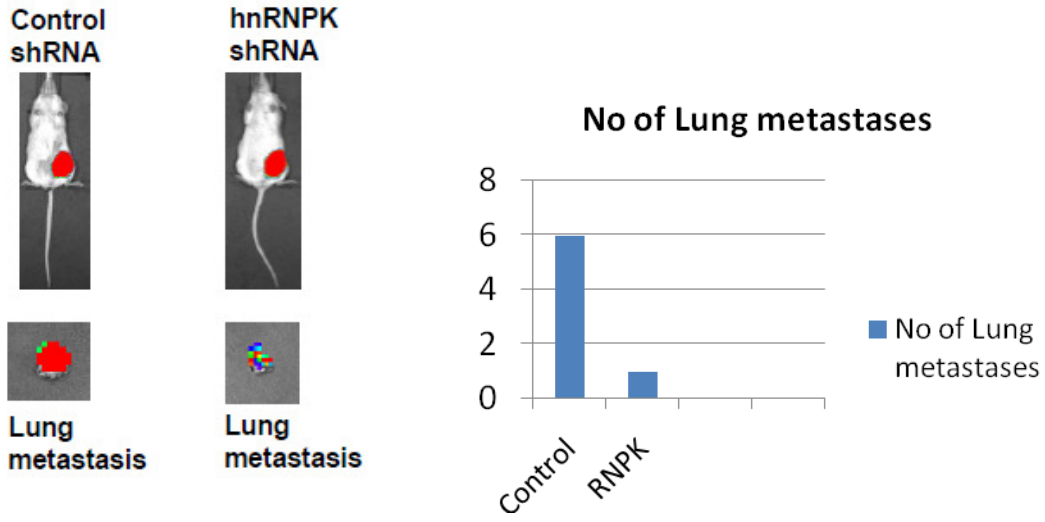

Fig 11: The metastasis formation of injected cells in lungs of mice, which shows there is reduction of no of metastatic nodules in lungs

To access the role of RNA binding protein hnRNPK, the hnRNPK knockdown cells were injected in tail vein and waited for its metastatic formation in lungs for 12-15 days. It was observed that there was reduction in number of metastatic nodules from the lung of the mouse injected with hnRNPK knockdown cells, compared with number of nodules in normal MCF 7 cells injected mouse.