January 2014

The Role Of Gene Regulation In Cancer: Studies Of Cancer-Related Phenotypes Mediated By Mex3d And By Microrna-618 Implicate Their Potential Oncogenic Role

Travis Whitfill
Yale University, twhitfill@gmail.com

THE ROLE OF GENE REGULATION IN CANCER: STUDIES OF CANCER-
RELATED PHENOTYPES MEDIATED BY MEX3D AND BY MICRORNA-618
IMPLICATE THEIR POTENTIAL ONCOGENIC ROLE

A Thesis
Presented to the Faculty of the School of Public Health
of
Yale University
in Candidacy for the Degree of
Masters of Public Health

by
Travis Michael Whitfill

Primary Thesis Advisor: Yong Zhu

May 2014

ABSTRACT

The role of gene regulation in cancer: Studies of cancer-related phenotypes mediated

by MEX3D and by microRNA-618 implicate their potential oncogenic role

Travis Michael Whitfill

2014

The control of gene expression is pivotal in the context of molecular

pathogenesis of a number of diseases, and thus is of critical relevance to public health.

An array of cellular tools exist in controlling gene expression, including epigenetic

effects, non-coding RNAs, and RNA-binding proteins. These tools are critical to the

modern study of public health, and are used in tandem with population-based studies.

This work focuses on specific examples of non-coding RNAs and RNA-binding

proteins, describing the effects of microRNA-618, a non-coding RNA, and MEX3D,

a post-transcriptional regulator, in cancer.

MicroRNAs (miRNAs) form a class of highly conserved endogenous RNAs

that inhibit gene expression and may act as oncogenes or as tumor suppressors,

regulating extensive cancer-related gene networks. Here, we show the association

between a single miRNA, miR-618, and cancer-related pathways in HeLa cells. MiR-

618 was identified as a potentially oncogenic microRNA, controlling a number of

cancer-related gene networks and pathways. Gain-of-function analysis reveals

differential expression of 110 transcripts following miRNA-618 transfection. Notably,

three upregulated genes are well-studied oncogenes—*KIT*, *JUN*, and *FOSB*—and three downregulated genes are well-known tumor suppressors—*PTPRO*, *STK11/LKB1*, and *IGFBP5*. Interestingly, investigation using the Ingenuity Pathway Analysis software tool reveals alterations in multiple cancer-related and cell cycle-related networks, including upregulated oncogenes in the top identified network "Post-translational modification, cellular development, cellular growth and proliferation" following miR-618 transfection. Further, miR-618 expression analysis shows overexpression in HeLa cells compared to normal cervical cells. Our findings present evidence for a novel oncogenic miRNA, miR-618, that is involved in cancer-related gene networks and is overexpressed in cancer.

This work also examined the role of a novel post-transcriptional regulator, MEX3D, in cancer. The Oncomine online database reveals that MEX3D is overexpressed in a number of solid tumors, notably in glioma. MEX3D is 3.01-fold overexpressed in glioma cells compared to non-cancerous, normal tissue. Kaplan-Meier survival analysis reveals that higher expression of *MEX3D* leads to poorer overall survival in overall glioma patients. Lastly, in a pilot case-control study of twelve glioma biopsies, we examined the effects of methylation in CpG sites in the *MEX3D* gene. The results were unclear, as we found a 3'UTR site that was 9.5% hypermethylated compared to normal tissue, a site in the body of the gene that was 24.8% hypermethylated, and second site in the body that was 15.4% hypomethylated compared to normal tissue.

Phenotypic studies reveal that MEX3D is responsible for two cancer phenotypes. Knockdown of MEX3D leads to increased cell proliferation and

decreased cell invasion, suggesting that overexpression of MEX3D is responsible for increased cell proliferation and decreased cell invasion.

This study is the first to describe the effects of miR-618 and of MEX3D in cancer. The findings presented in this work lay the foundation for further mechanistic studies of miR-618 and MEX3D. More work is needed to identify the mechanisms of oncogenesis controlled by these molecules. Our study indicates that miR-618 may be a biomarker for several types of cancer and warrants further investigation.

ACKNOWLEDGEMENTS

"Yale is at once a tradition, a company of scholars, a society of friends."

-- George Pierson, Yale College

Being at Yale has been a transformative experience in ways I did not foresee before starting my studies at Yale. In my two short years at this great University, I have made valuable connections, lasting friendships, and unforgettable memories. I owe everything in this work and the trajectory of my career path to this institution, and for those reasons, I am forever grateful for the opportunity to be here. Yale is full of incredible scholars from all across the world, and I am eternally appreciative to have been a part of the collective body of scholars at Yale—albeit a short period of two years.

I'd like to first thank my advisor, Yong Zhu, for his continuous support over the last two years. I am grateful to have had the opportunity to study in his lab, and I am glad to have found a lab in which the research encompasses so many aspects of medical research, from population-based research to intricate molecular studies. The members in his lab have been instrumental in this project. I'd especially like to thank Daniel Jacobs for his help in this work. He has spent many hours helping me, guiding me, and advising me, and I have thoroughly enjoyed working with him. He contributed to the design and execution of a number of the experiments represented in this thesis.

I would like to thank other members of the lab as well, especially Alan Fu. He has also been instrumental in the success of this thesis. His intellectual contributions fueled a number of aspects of this thesis. I'd also like to thank Allison Fritz for the expression analysis of miR-618 in HeLa cells.

Lastly, I would not be here without the support of my friends, both old and new, from home and here at Yale. They have gotten me through the drudgery of grad school and have provided moral support for which I am very thankful.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

3'UTR . . . . 3' untranslated region

ARE . . . . .AU-rich element

ATCC . . . . American Type Culture Collection

cDNA . . . . Complimentary DNA

DMEM . . . . Dulbecco's Modified Eagle's Medium

FBS . . . . .Fetal Bovine Serum

FOSB .FBJ murine osteosarcoma viral oncogene homolog B

GAPDH . . . . Glyceraldehyde 3-phosphate dehydrogenase

GWAS . . . . .Genome Wide Association Study

hsa-miR-618. . . . . Human micro-RNA-618

IGFBP . . . . Insulin-like growth factor-binding protein

IPA . . . . .Ingenuity Pathway Analysis

JUN . . . . Jun proto-oncogene

KH . . . . K-homology domain

LKB1 . . . . .Liver kinase B1

KIT . . . . v-kit HardyZuckerman 4 feline sarcoma viral oncogene homolog

MEX-3 . . . . Muscle excess-3 homolog (C. elegans)

miRNA . . . . .Micro ribonucleic acid

mRNP . . . . .Messenger ribonucleic protein

Pre-miRNA . . . . Preliminary micro-RNA

PTPRO . . . . .Protein tyrosine phosphatase, receptor type O

qPCR . . . . Quantitative polymerase chain reaction

RBP . . . . .RNA-binding protein

RING . . . . Really Interesting Element

CHAPTER 1: INTRODUCTION

The field of public health is changing. With the exponential decline of the cost of whole-genome sequencing in tandem with explosive growth in research and technology, we have seen a new paradigm emerge in public health research that integrates population-based observations with complex data describing the complexities of molecular mechanisms underlying complex diseases. No longer do molecular research and traditional epidemiology need to remain separated. Rather, an integrative approach is feasible—even warranted—and is gaining tractability and popularity. This approach is much more powerful than a dis-integrated approach, bridging gaps between medicine, bench-side research, and epidemiology, thus providing a more relevant and 21st-century approach to studying diseases.

For the past decade, the genome-wide association study (GWAS) has been at the forefront of genome-wide efforts to identify genetic variants that impact complex human traits, providing key insights into the underlying allelic architecture of complex human diseases. However, as we transition into what is commonly dubbed the "post-GWAS" era of cancer genomics, and genome-wide initiatives move from being novel to routine, it has become apparent that the main lesson learned from GWAS efforts is that most cancer-associated genetic variants have only minor effects on disease phenotype and, in total, explain only a small fraction of disease heritability. This is due in part to the limited scope of genetic analyses that fail to account for the complex interactions that occur between environmental stimuli and genetic and epigenetic phenomena and in part due to the stringent statistical requirements inherent to the GWAS.

1.1 Function of microRNAs

One such epigenetic mechanism involves microRNAs (miRNAs), a group of noncoding RNAs that have recently been found to silence genes post-transcriptionally. miRNAs are small RNA molecules of 20-22 nucleotides that silence target genes by either degrading the target mRNA or by preventing its translation.[1] Surprisingly, miRNAs have a large impact on genome-wide expression levels, as single miRNAs have been shown to regulate expression of several hundred genes.[2] It is estimated that miRNAs control the expression of at least a third of all mRNAs in the human genome.[3] Thus, the role of miRNAs is fundamental in cellular processes, and aberrant expression of these RNAs is tied to a number of diseases and many types of cancer, including breast, colorectal, prostate, and lung cancer.[3-7] The first identified miRNAs in *C. elegans*, *lin-4* and *let-7*[8]

Since their discovery in the early 1990s, microRNAs have been thoroughly investigated, and their role in gene expression has been well studied. MicroRNAs are small RNA molecules of 20-22 nucleotides.[1] The biogenesis of miRNAs is complex and includes a number of proteins, including members of the Argonaute family, polymerase II-dependent transcription, and RNAses Drosha and Dicer.[9] After transcription processing by the Drosha complex, the 60-110 nucleotide precursor miRNA, or pre-miRNA, is then exported from the nucleus and is further processed by Dicer and TAR RNA-binding protein 2 (TARBP2) to create the final 20-22 nucleotide mature, double-stranded miRNA in the cytoplasm.[10] Once the mature miRNA is formed, the antisense strand becomes incorporated into the miRNA-containing RNA-induced silencing complex (mi-RISC) and guides the complex to its

target mRNA. After hybridization between the miRNA and mRNA, the mRNA is degraded and thus prevents gene expression. Surprisingly, miRNAs have a large impact on genome-wide expression levels, as a single miRNA can downregulate expression of several hundred genes.[3]

1.2 Role of miRNAs in cancer

The role of aberrant miRNA expression in cancer and other diseases has been extensively investigated in the past 10 years. miRNAs are frequently located at fragile sites in the chromosome and in genomic regions that have been linked to cancer.[11] Importantly, miRNAs can have a role as either oncogenes or tumor suppressors (collectively termed as oncomirs).[3,9,12] Oncogenic miRNAs typically target tumor suppressor genes, while tumor suppressive miRNAs usually repress oncogene expression (See **Figure 1.2**). Aberrant miRNA expression is often due to hypermethlated CpG islands proximal to the encoded miRNA.[10] Notably, abnormal miRNA expression has been found in almost all solid tumors and hematologic malignancies.[9,13,14] We and other groups have also identified associations between single nucleotide variations in miRNAs and cancer.[15,16]

Due to their ability to modulate gene expression by post-translational repression, miRNAs that target oncogenes (tumor-suppressive miRNAs) are essential for maintaining stability and preventing tumorigenesis. However, it has been recently discovered that when these miRNAs are aberrantly downregulated, there is an increased risk of genomic instability, cell cycle aberrations, and possibly cancer. Thus, certain cancers are likely to display a profile of differentially expressed microRNAs. Following this discovery, several specific miRNAs have been identified that are

strongly downregulated in certain types of cancers. Conceivably, miRNAs could be introduced into the cell as miRNA "replacement therapy," thereby artificially reversing the aberrant lack of presence of these regulatory RNAs (see Figure 1 for a conceptual summary). Indeed, a number of recent studies have successfully demonstrated efficacious treatment of cancers *in vivo* that support this "replacement therapy" hypothesis.

1.3 Effects of DNA methylation in cancer

Gene expression is not determined solely by the DNA base sequence; it also depends on epigenetic phenomena, defined as gene-regulating activities that do not involve changes in DNA sequence whose effects may be inherited. [17] Epigenetic control of gene expression occurs through two primary mechanisms: DNA methylation (i.e., 5-methylcytosine content of DNA) and chromatin modification (i.e., the preservation of DNA protein complexes). Although epigenetic regulation of gene expression has been known and studied for decades, little disease-related work was done in this area until the close of the last century. Epigenetic regulation of gene expression is a dynamic process, and the epigenetic status of a gene can change over the lifetime of an organism. This ability to change and adapt in response to developmental cues, environmental cues, or both rests primarily on post-synthetic modifications either of the DNA itself (e.g., via methylation) [18] or of proteins that intimately associate with DNA (e.g., the acetylation, methylation, or phosphorylation of histones). [19] These modifications can then be interpreted by proteins (e.g., DNA methyltransferases and histone deacetylases) that recognize a particular modification and facilitate the appropriate downstream biological events.[20]

A major focus of epigenetic research has been the identification of CpG-rich sequences, termed CpG islands, and the determination of methylation patterns within these regions in relation to disease status. The most common post-synthetic modification of the mammalian genome is the methylation of the 5' position of cytosines in dinucleotide CpG sequences. Approximately 60–90% of all dinucleotide CpG sequences in the genome are methylated, while unmethylated dinucleotides are mainly clustered in CpG islands in gene promoters.[21] Although the addition of a methyl group does not alter base pairing, it can impinge on protein-DNA interactions via the attenuation of binding affinity of sequence-specific DNA binding proteins for their cognate *cis* elements and result in changes in chromatin structure and histone acetylation status.[22] Normally, both the core promoter and transcription start site are located within the CpG island, and gene expression is completely repressed when this region becomes hypermethylated. As such, the predominant role of DNA methylation encompasses a wide range of functions, including developmental and tissue-specific gene transcription,[23,24] imprinted gene expression,[25] expression of transposable elements,[26] chromatin organization [27] and suppression of integrated foreign DNA.[28,29] The relative plasticity of epigenetic mechanisms makes the study of epigenetics a powerful paradigm in the investigation of disease etiologies. In particular, epigenetics may offer an opportunity to further our understanding of a number of features of complex diseases that traditional DNA sequence–based genetics has been unable to explain.[30]

1.4 RNA binding proteins and disease relevance

RNA-binding proteins (RBPs) form a class of highly evolutionarily conserved post-transcriptional regulators found in a variety of organisms. Nascent mRNAs associate with RBPs during transcription to form messenger ribonucleoprotein (mRNP) complexes. RBPs direct and control the localization, stability, or translation of their target RNAs.[31] Recent studies underscore the importance of such proteins in a variety of physiological functions, including stem cell differentiation, embryogenesis, and cancer.[32] Most RBPs recognize the secondary structures formed by their target mRNA, but are also able to recognize certain intrinsic features of RNA such as certain sequence motifs, the 5' cap or the 3' poly (A) tail. Furthermore, RBPs are involved in miRNA stabilization and processing, and disruption between the interplay between miRNA and RBPs have clear implications in cancer.[33,34]

Recent high throughput techniques such as photoactivatable-ribonucleoside-enhanced UV crosslinkning and immunoprecipitation (PAR-CLIP), iCLIP, and CLIP-seq have identified the mRNA-bound proteome.[35-37] From these advanced techniques, more than 800 mammalian RBPs with over 40 RNA-binding domains (RBDs) for recognition of secondary structures of specific sequence motifs have been identified.[31,32] This includes RNA recognition motifs (RRMs), the K-homology (KH) domain, double-stranded RNA binding motif (dsRBM), the Arg-Gly-Gly (RGG) box, DEAD/DEAH box, and Piwi/Argonaute/Ziwi (PAZ) domains.[38] KH domains are important motifs in RNA-binding activity, and a number of studies have elucidated the importance of the motif's function in RNA binding and gene regulation.[39,40]

Many RBPs are relevant to a number of diseases, particularly in cancer. For example, the RBP insulin-like growth factor 2 mRNA-binding protein 1 (IGF2BP1) is

upregulated in several types of cancer, and has been shown to display oncogenic activity through binding and stabilizing c-MYC and MK167 mRNA.[41] Other members of the IGF2BP family such as IGF2BP2 and IGF2BP3 have been shown to be post-transcriptional drivers of cancer progression and can regulate important drivers in cancer including β-catenin, let-7 microRNA, KRAS, and MDR1.[41-44]

1.5 MEX-3

MEX-3 is a translational repressor and RNA-binding proteins originally discovered in *C. elegans.* Vertebrates have four MEX-3 analogues, and are highly evolutionarily conserved across species. The MEX-3 family plays an important role in a number of cellular functions, but most recently they have been found to play a central role in cell differentiation and cell renewal, and are thus of major concern to stem biology and cell biology.[45]

Studies in *C. elegans* underscore the importance of MEX-3 in embryonic development. Mutations in *Mex-3* are embryonic lethal and cause defects in the AB anterior blastomere descendants, particularly in the body wall muscle, thus the name *muscle excess 3* for *Mex-3*.[46] In oocytes, *Mex-3* is uniformly distributed, but becomes asymmetrically overexpressed in the two anterior blastomeres at the four-cell stage.[47] MEX-3 targets PAL-1 in the *PAL-1* 3'UTR and prevents its translation in the anterior blastomeres.[48,49] Thus, MEX-3 spatially represses *PAL-1*. PAL-1 is necessary for establishing the posterior blastomeres. In addition to its role in establishing the anterior blastomeres MEX-3 is also involved in ensuring segregation of P granules at the eight-cell stage.[46] P granules contain maternal mRNAs and are thought to be

involved in determining the germline.[47] Thus, this function suggests the importance of MEX-3 in establishing the specification of this cell lineage.

Importantly, *MEX-3* family members are highly evolutionarily conserved across species. Phylogenetic analysis reveals that ancestral *mex-3* underwent two round of duplication after vertebrate lineage divergence, which accounts for the presence of four homologues in both mice and humans.[50] MEX-3 contains two K homology (KH) domains that contain RNA-binding activity. The structure of the KH domain bound to its target RNA is shown in **Figure 1.3.** However, vertebrates contain an additional domain, perhaps a neo-functionalization that may have arisen during species divergence, the carboxy terminus really interesting new gene (RING) finger domain, which is absent in *C. elegans*. RING domains have been shown to mediate E3 ubiquitin ligase activity. Such activity has been shown in MEX-3C, which is responsible for degradation of human leukocyte antigen serotype A2 (*HLA-A2*) mRNA.[51]

1.6 Disease relevance of MEX-3 family

While the primary function of MEX-3 proteins in humans is unknown, recent evidence suggests they carry a role in cancer and other diseases, including hypertension, metabolic diseases, and cancer.[52-54] Interestingly, MEX3A overexpression in an intestinal cell line lead to impaired differentiation and altered polarity.[55] Further, MEX-3 in *C. elegans* represses cell-cycle inhibitor cyclin-dependent kinase inhibitor-2 (CKI-2)[56], whose human orthologues p21, p27, and p57 contain tumor suppressive activity.[45] Furthermore, silencing of *MEX-3C* causes

chromosome mis-segregation, structural abnormalities, and DNA replication stress.[45] Data from the Oncomine database (http://www.oncomine.com) strengthens the possible role of MEX-3 family members in cancer, as there are large amounts of data that suggest its overexpression in cancer compared to normal tissue.

MEX-3D, a member of the hMEX-3 family, is expressed in somatic tissue in humans, contrasting with the expression profile of its analogues, which are restricted to the embryo and the germ line. Interestingly, the Oncomine database reveals that MEX-3D is highly overexpressed in a number of cancer samples compared to normal tissue. Yet, there is a gross lack of understanding of the role of this protein in both normal cell physiology and in cancer. As such, we aimed to elucidate these roles by with an emphasis on glioma. The success of this project would have profound implications both in cancer biology and in the clinical relevance of these proteins, as they could serve as potential biomarkers or candidates for gene therapy in breast cancer.

Despite gaining recent momentum in research, there is a gross lack of understanding of the role of these proteins. Namely, there are four major deficits in MEX3 family research (adapted by Pereria *et al.* 2013[45]): (1) Identification of RNA targets and mechanism or recognition and regulation; (2) Distinction of functional redundancy between MEX-3 family members; (3) Determination of interactions between signaling pathways and molecular processes; and (4) Analysis of physiological and pathological roles, especially in cancer.

Interestingly, the Oncomine database reveals that MEX-3D is highly overexpressed in a number of cancer samples compared to normal tissue. Yet, there is a gross lack of understanding of the role of this protein in both normal cell physiology and in cancer. The goal of this study was to determine the effects of MEX3D in glioma.

1.7 Epidemiology and clinical characterization of glioma

Malignant brain tumors comprise a frightening and deadly diagnosis. Little advances in this area have improved outcomes. An estimated 138,000 patients were newly diagnosed with a primary malignant brain tumor in the United States in 2010.[57] Dismally, the 5- 10- and 20- year overall survival rates for all gliomas are 54%, 45%, and 40%, respectively.[57] The World Health Organization (WHO) classification system categorizes glioma into four distinct histological grades differentiated by increasing degrees of dedifferentiation, anaplasia, and aggressiveness.[58] Grade III tumors include astrocytomas, oligodendroglioma, and oligoastrocytoma.[58] Grade IV tumors, or glioblastoma, account for 82% of cases of malignant glioma and are defined histologically by considerable cellularity, mitotic activity, vascular proliferation, and necrosis.[59] Sadly, the 5-year survival rate for glioblastoma is 4.7%.[60]

Malignant gliomas arise from sequential genetic alterations from intrinsic and environmental influences. Contrary to public perception, risk factors for glioma do not include exposure to cell phones.[59,61,62] These studies are limited by a number of factors including recall bias, inconsistent measurement of cell phone usage, and varying cell phone technologies over time. Established risk factors for glioma are

limited, but include ioninizing radiation.[63,64] Gliomas are most common later in life, especially in the sixth to eighth decade of life.[59] Family history has a moderate influence on risk for glioma and is associated with about a 2-fold increase in risk of developing glioma.[59] Genome-wide association studies have identified a few genes that are weakly associated with glioma, including 0q13.33 (*RTEL*), 5p15.33 (*TERT*), 9p21.3 (*CDKN2BAS*), 7p11.2 (*EGFR*), 8q24.21 (*CCDC26*), and 11q23.3 (*PHLDB1*).[65,66]

Figure 1.1 Role of miRNAs in gene expression

Gene

Transcription

mRNA

miRNA

Translation

Protein
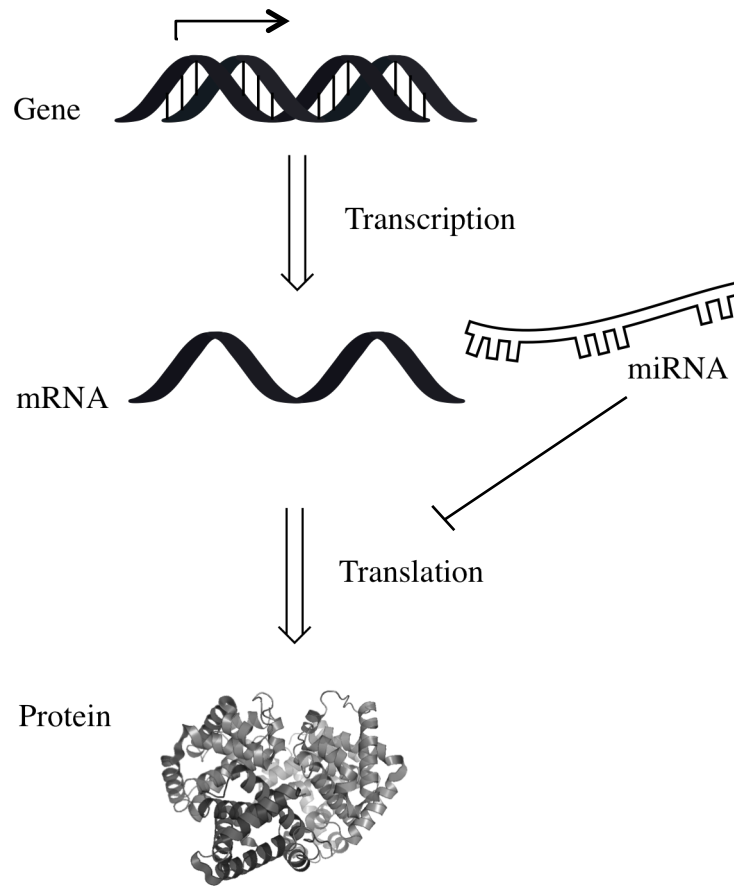
Figure 1.1 Role of miRNAs in gene expression

miRNA silences translation of a gene through post-transcriptional targeting. First, the target gene is transcribed in the nucleus. Once exported into the cytoplasm, the miRNA binds to the target mRNA, usually in the 5'UTR region. This blocks the translation of the mRNA by preventing binding of translational machinery and thus preventing the encoded protein from being translated. It is important to note that a single miRNA can have up to hundreds of targets through non-specific binding.

Figure 1.2 Role of miRNAs in molecular carcinogenesis

a. Overexpression of miRNAs

b. Underexpression of miRNAs

mRNA

mRNA

Decreased tumor suppressor gene expresion

Increased oncogene expression
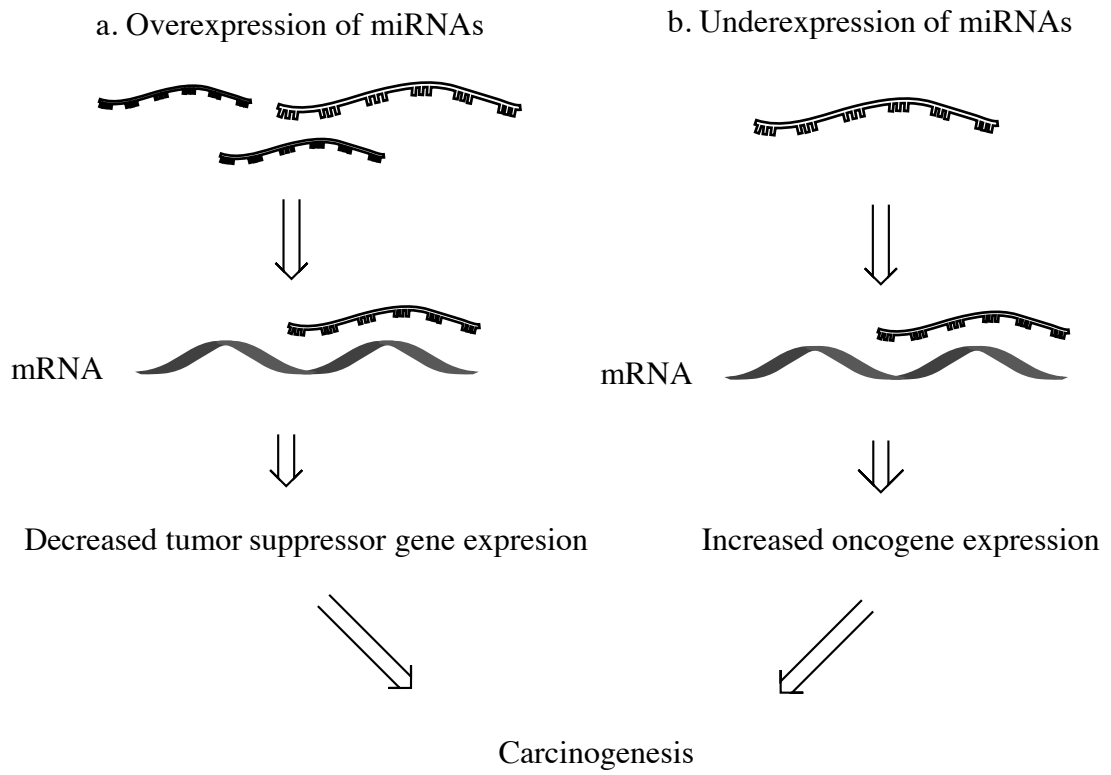
Carcinogenesis

Figure 1.2 Role of miRNAs in molecular carcinogenesis

miRNAs can act as either oncogenes or tumor suppressor genes. Oncogenic miRNAs typically silence translation of tumor suppressors and the expression of these miRNAs remains at a relatively low level of expression in normal cells. However, if overexpressed, they can inhibit translation of tumor suppressors, leading to decreased tumor suppressor expression, and increased likelihood of carcinogenesis (a). Conversely, tumor suppressive miRNAs typically inhibit expression of oncogenes. When these miRNAs are underexpressed, they lead to overexpression of tumor suppressors, thus potentially leading to carcinogenesis (b).

Figure 1.3 Schematic of mRNA in humans and regulation of post-transcriptional activity
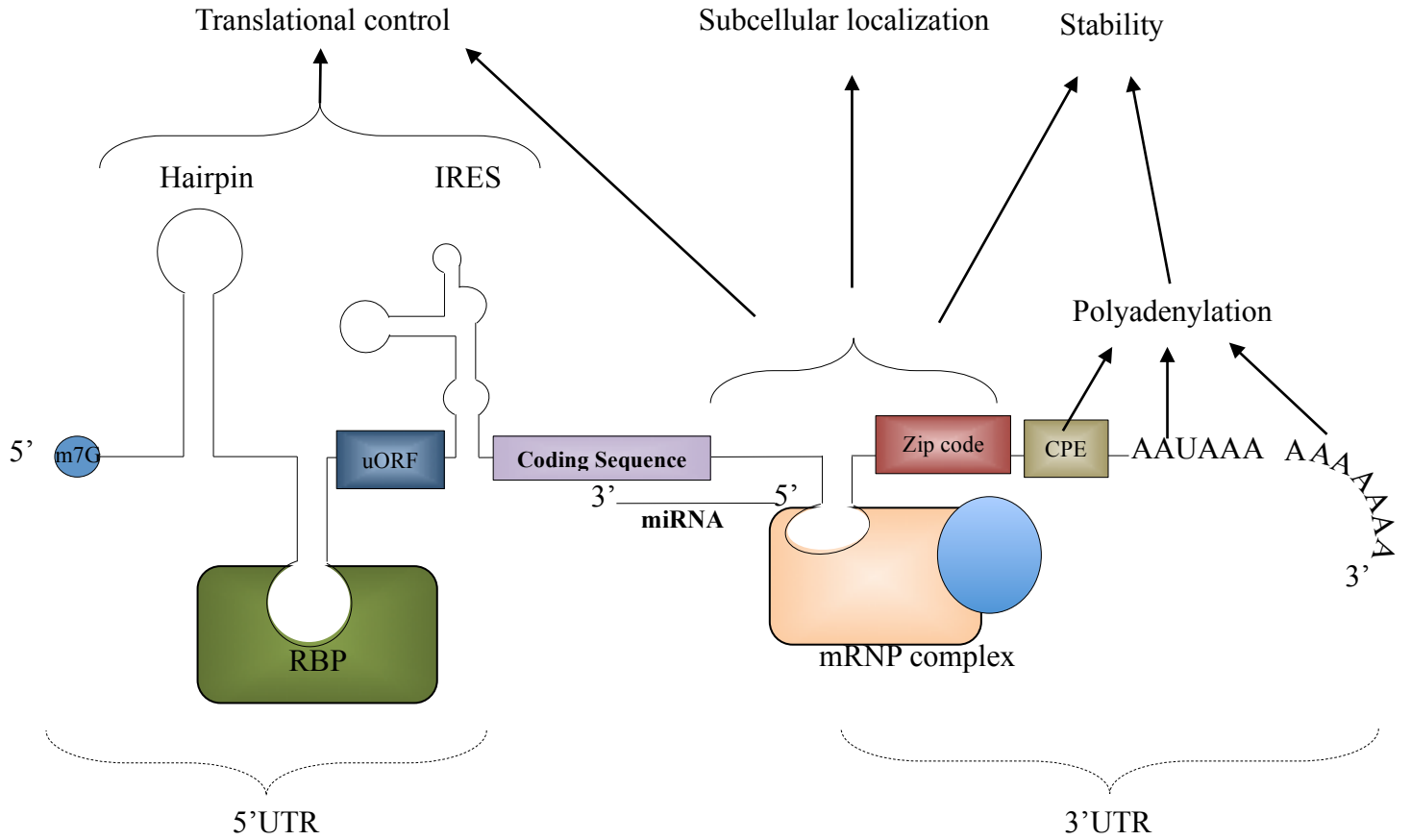
Figure 1.3 Schematic of mRNA in humans and regulation of post-transcriptional activity

The generic structure of a eukaryotic mRNA, illustrating some post-transcriptional regulatory elements that affect gene expression. Abbreviations (from 5' to 3'): UTR, untranslated region; m7G, 7-methyl-guanosine cap; hairpin, hairpin-like secondary structures; RBP, RNA-binding protein; uORF, upstream open reading frame; IRES, internal ribosome entry site; mRNP, messenger ribonucleoprotein; CPE, cytoplasmic polyadenylation element; AAUAAA, polyadenylation signal. Based on Mignone *et al.* 2012.[67]

Figure 1.4 3D structure of the KH domain
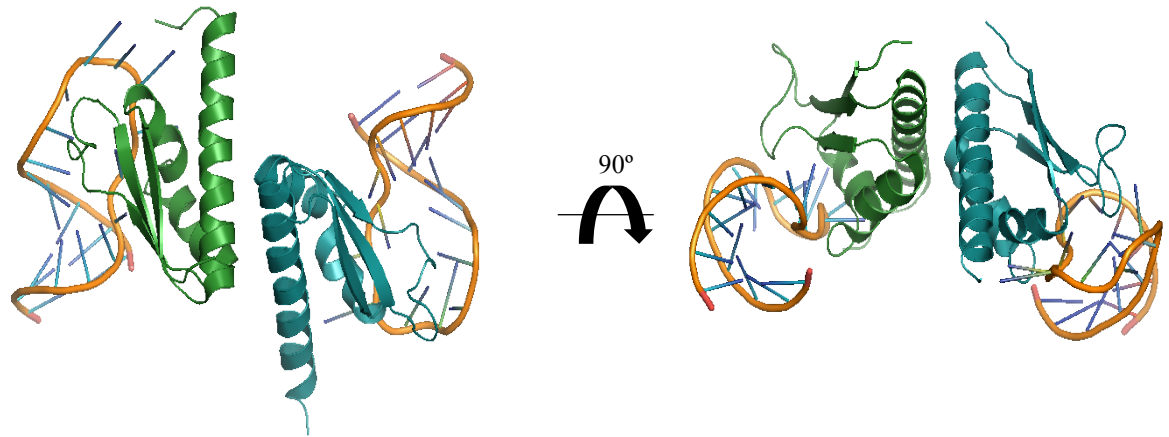
Figure 1.4 3D structure of the KH domain

The structure was adapted by Lewis *et al.* 2000.[68] It shows the Nova-2 KH3-RNA

complex. The structures highlight the dimeric assembly of the KH domains.

CHAPTER 2: MIR-618 DISPLAYS POSSIBLE ONCOGENIC ACTIVITY

2.1 INTRODUCTION

One newly identified miRNA, hsa-miR-618, is located in the chromosomal region 12q21.31.[22] Interestingly, an increased risk for prostate cancer has been associated with low copy number variation with this region.[69] Also, two recent studies have suggested a possible role of miR-618 in cancer. One study identified a single nucleotide polymorphism in miR-618 and found a weak, non-significant positive association with risk of breast cancer.[15] A second recent study identified miR-618 as a possible biomarker for hepatocellular carcinoma among patients infected with Hepatitis C.[16] Yet, these findings have not been supported, and the potential role of miR-618 in cancer remains unclear.

In order to elucidate the role miR-618 plays in cancer, we investigated the differential gene expression associated with miR-618 transfection into cultured cells using genome-wide expression microarray analysis. After the gene expression levels were obtained, significantly altered transcripts were investigated for network and functional interrelatedness using the Ingenuity Pathway Analysis software tool. Our study indicates a strong correlation between miR-618 and cancer-related networks and genes. Three common oncogenes were upregulated—i.e. *JUN, FOSB,* and *KIT*— while three tumor suppressor genes were downregulated as a result of the transfection, *PTPRO, IGFBP5*, and *STK11* (this will henceforth referred to as *LKB1*). Furthermore, we observed a number of genes involved in hepatotoxicity and hepatocellular carcinoma, which supports the recent study that found an association between aberrant miR-618 expression and hepatocellular carcinoma. Finally, we found an

overexpression of miR-618 in HeLa cells compared to non-cancerous End1 cervical cells. Overall, our results seem to implicate an oncogenic role for miR-618, and as such, warrant further investigation.

2.2 RESULTS

Effect of genome-wide expression levels by miR-618 transfection

The microarray assay determined genome-wide expression in miR-618 transfected cells compared with cells transfected with the scrambled control. 110 identified transcripts fit our criteria for significance (Q < 0.05 and fold change ≥ |2|). Of these 109 transcripts, 79 showed increased expression (fold change ≥ 2), and 30 transcripts showed decreased expression (fold change ≤ -2). The predicted secondary structure of pre-miR-618 is shown in **Supplementary Figure S1**, and an exhaustive list of all significantly differentially expressed genes can be found in **Supplementary Table 1**.

Cancer-related networks formed by miR-618-mediated genes

To determine which biological functions and networks may be influenced by miR-618, we examined the set of genes with altered expression using the Ingenuity Pathway Analysis software tool. We identified eight biological groups of molecules that contained at least ten differentially expressed transcripts. Cancer is the top biological-related function associated with miR-618 transfection, involving 29 differentially expressed genes (p-value <0.05). Other identified functions include immunological disease (22 genes), organismal injury and abnormalities (10 genes),

cellular movement (16 genes), cellular development (16 genes), cellular growth and proliferation (20 genes), cell cycle (11 genes), and cell death (19 genes) (p-values <0.05). These biologically relevant functions are summarized in **Table 1**, and the cancer-related transcripts are summarized in **Table 2 and detailed in Supplementary Table S2**. Furthermore, differentially expressed transcripts were associated with a number of cancer-associated functions. In addition to the genes identified related to cancer, cell cycle, cell death, and cellular growth and proliferation, we also identified significant associations with tumor morphology (4 molecules), and DNA repair (7 molecules). These cancer-related functions are summarized in **Figure 1**.

The top network identified, "Post-translational modification, cellular development, cellular growth and proliferation", consists of 11 overexpressed transcripts (*ANG, DCN, FGFBP1, FOSB, HMGB1, JUN, KIT, KRT18, MIF, PCDH1, and PDGFRB*) and 5 underexpressed transcripts (*CBX5, IGFBP5, PDGFRA, PTPRO, and LKB1*). Importantly, several of these upregulated genes, namely FOSB, JUN, and KIT, are well known oncogenes. Known functions and respective fold changes of genes in this network are summarized in **Table 2**. **Figure 2** presents the peptides encoded by these genes in their respective locations of function in the cytoplasm, nucleus, cell membrane, or extracellular space. The second top network identified was "Cellular movement, cell death and survival, and cell cycle." The software identified eight overexpressed transcripts (*SGK3, EDN2, HUWE1, IFI27, MAPK4, MIF, NUDCD2*, and *TM4SF1*) and two underexpressed transcripts (*C1QTNF5* and *CLDN5*). **Supplementary Figure S2** presents a graphical representation related to

this network. All microarray expression data have been uploaded to the Gene Expression Omnibus (GEO) database, and can be accessed via their website (http://www.ncbi.nlm.nih.gov/geo/; accession # pending).

Overexpression of miR-618 in Hela cells

In order to determine if miR-618 is overexpressed in cancer compared to normal tissue, we performed an expression analysis in miR-618 non-transfected HeLa cervical cancer cells relative to End1 cervical cells. The expression analysis in HeLa cells revealed an overexpression of miR-618 by 2.68-fold relative to End1 cells (**Figure 3**).

2.3 DISCUSSION

The importance of miRNAs in regulating and maintaining a variety of cellular functions is widely known and studied. Certain miRNAs carry the potential for either oncogenic or tumor suppressive activity, and can regulate cancer-related gene networks. After analyzing the effects of miR-618 transfection on genome-wide expression levels in HeLa cells, we report here the oncogenic role of miR-618 and its regulatory impact on cancer-related gene networks. miR-618 is located in the 12q21.31 region, and interestingly, an increased risk for prostate cancer has been associated with low copy number variation in this region.[69] The gene *MGAT4C* also maps to 12q21.31, and its copy number variation has been associated with prostate

Recently, our group has described the effects of miR-618 in non-Hodgkin lymphoma.[70] A total of 128 miR-618 targets were identified that were enriched for genes that have functional roles in lymphoma-relevant pathways. The study also revealed a significant association between rs2682818 G>T in the miR-618 and follicular lymphoma (FL) (OR: 1.65, 95% CI: 1.05-2.60). In vitro analysis of rs2682818 functional impact revealed that the variant T allele resulted in reduced levels of mature miR-618, which in turn may lead to deregulation of miR-618-controlled pathways relevant to follicular lymphoma.

We identified three overexpressed oncogenes from the miR-618 transfection. *JUN, FOSB,* and *KIT* are commonly studied oncogenes and were upregulated by fold changes of 2.79, 3.62, and 2.20, respectively (Q-values <0.05). These genes all belong to the top identified network "Post-translational modification, cellular development, cellular growth, and proliferation." Notably, FosB has been demonstrated to lead to increased cell invasion in MCF7 cells.[71] Together, FosB and Jun form a heterodimer and are both highly overexpressed in a variety of cancers.[72] The Jun/FosB complex is essential for regulating a number of cancer-related functions, including invasion and metastasis, cell proliferation, cell differentiation, angiogenesis, and cell survival.[73] Further, the c-Kit receptor, a transmembrane protein, is a notable member of the type III receptor tyrosine kinase family and also regulates similar biological functions as Jun/FosB, including cell proliferation, cell differentiation, cell development, and cell migration.[74] Overexpression or abnormal regulation of c-Kit has been strongly linked to cancer.[75] Thus, miR-618 carries

important oncogenic activity by upregulating these oncogenes involved in mechanisms central to cancer and tumorigenesis.

PTPRO, IGFBP5, and *LKB1* were significantly underexpressed tumor suppressors following miR-618 transfection with fold changes of -2.07, -2.02, and -2.38, respectively (Q-values <0.05). PTPRO encodes a protein tyrosine phostphatase receptor that is involved in proper entry into the G_1 phase and regulates apoptotic pathways.[76] Loss-of-function of this receptor is highly associated with cancer.[76] IGFBP5, which encodes an insulin-like growth factor binding protein, is another important protein in cancer that is essential in inhibiting mitogenesis, cell differentiation, and in preventing metastasis.[77,78] Aberrant *IGFBP5* expression has been documented in a number of cancers. *LKB1* encodes an important serine/threonine kinase-activating kinase in which loss-of-function leads to Peutz-Jehers syndrome and sporadic cancers such as lung and cervical cancer.[79] LKB1 plays a role in phosphorylation cascades that regulate essential cellular processes such as cytoskeleton organization, cell proliferation, and metabolism.[79] This strengthens support of the putative role of miR-618 in cancer, as the loss of tumor suppressor genes and overexpression of oncogenes is a common phenotype in nearly all cancers.

Importantly, a recent study found a significant association between upregulated miR-618 expression and hepatocellular carcinoma (HCC) among Egyptian patients infected with Hepatitis C virus (HCV),[80] identifying miR-618 as a potential risk biomarker for HCC. Interestingly, our IPA analysis found an upregulated gene that is associated with hepatocellular carcinoma, *IL22RA,* due to miR-618 transfection. *IL22RA* is directly associated with carcinoma and carcinoma *in*

situ. More importantly, *IL22RA* is found within the same interaction network as *IL22,* which is associated with hepatocellular carcinoma and liver tumorigenesis, proliferation, hyperplasia, and hypertrophy (all p-values <0.05). Furthermore, we identified seven molecules associated with hepatotoxicity and hepatocellular carcinoma that were differentially expressed: *TM4SF1, KIT, PDGFRB, HUWE1, HMGB1, PTPRO, PDGFRA,* and *MATN3* (p-value = 5.72 x 10[-3]; see **Supplementary Table S1** for more information). *PTPRO* has been associated with hepatocellular carcinoma progression when underexpressed, and exhibited a -2.07-fold change in expression (Q<0.05) following miR-618 transfection.[81] Thus, taken in the context of miR-618 as a biomarker for hepatocellular carcinoma, our analysis implicates miR-618 as a potential upstream effector of hepatocellular carcinogenesis via its interactions with genes relevant to HCC proliferative and hepatotoxicity pathways.

Interestingly, an A>C single nucleotide polymorphism (SNP), rs2682818, has been identified in miR-618 that may be involved in cancer. According to the NCBI SNP database, the hsa-miR-618 SNP variant (C to A) has been most commonly found among Japanese and African American individuals, and least common among Caucasians (A genotype frequency 0.320, 0.320, and 0.170, respectively).[82] A study by Zhang *et al.* recently investigated the role of this SNP in miR-618 in breast cancer among Chinese females, but found no significant association.[15] However, the study was limited due to its small sample size, thus further investigation is necessary to explore the potential role of the SNP in miR-618 in tumorigenesis in breast cancer and in other cancers. A study in 2010 by Cicatiello *et al.* identified miR-618 as a target of 17β-estradiol in estrogen-responsive breast cancer MCF-7 breast cancer

cells.[83] The findings both associate miR-618 in breast cancer and demonstrate that miR-618 is responsive to estrogen control of target gene activity. These studies support our findings by providing a possible role of aberrant miR-618 expression in breast cancer specifically.

As with all *in vitro* studies, there are several limitations in this study. As the miR-618 transfection was conducted in HeLa cells, the findings are limited to genes with detectable expression levels in this cell line. HeLa cells have a gene expression profile that differs from those of other cell lines, and, as such, it is likely that some observed expression changes may be HeLa-specific. Further, since we determined following the transfection and network analysis that miR-618 is upregulated in HeLa cells compared to End1 cells, the results may have been more meaningful in a cell line in which miR-618 was not overexpressed. Nevertheless, our findings indicate a strong potential role for miR-618 in cancer. In combination with the previous suggestion of miR-618 as a potential biomarker for hepatocellular carcinoma, evidence strongly supports a possible role of this miRNA in hepatocellular carcinoma, liver cancer, and other diseases. A number of future studies are warranted based on our results, including an expansion of this analysis to other cell lines, and an investigation of the role of this miR-618 in clinical samples.

As previous evidence regarding the role of the novel miRNA miR-618 in cancer remains unclear, we present evidence that miR-618 may serve an important role in cancer, finding that it affects the expression of key oncogenes and tumor suppressors and that it operates in cancer-related pathways. Our observations of miR-618 overexpression in cancer cells support an oncogenic role of miR-618. Taken

together with our findings of overexpressed oncogenes and underexpressed tumor suppressor genes following miR-618 transfection, miR-618 is a likely oncomir. Our data indicate that when overexpressed, miR-618 may carry an oncogenic role by modulating pathways leading to increased cell growth, cell migration, cell invasion, and cell proliferation. Notably, our findings suggest a role for miR-618 in hepatocellular carcinoma, as a number of genes associated with hepatocellular carcinoma and hepatotoxicity were dysregulated as a result of the miR-618 transfection. Further *in vitro* and *in vivo* studies are warranted to validate and elucidate the association between miR-618 and cancer and to determine the clinical relevance of miR-618 as a possible biomarker in cancer.

Table 2.1 Top biological functions of differentially expressed target genes of miR-618

Name	No. of molecules	P-value[a]
Cancer	29	8.61×10^{-7}
Immunological disease	22	8.61×10^{-7}
Organismal injury and abnormalities	10	8.61×10^{-7}
Cellular movement	16	2.95×10^{-4}
Cellular development	16	3.89×10^{-2}
Cellular growth/proliferation	20	3.89×10^{-4}
Cell cycle	11	1.55×10^{-3}
Cell death	19	2.04×10^{-3}

`[a] P-values represent the pooled value of the target genes and were calculated with the Fisher t-test.

Table 2.2 Identified cancer-related differentially expressed miR-618 targets

Gene name	Expression fold changes[a]	Gene name	Expression fold changes
ANG	2.25	*IL22RA1*	2.52
ANXA8L2	2.19	*JUN*	2.79
BIK	2.66	*KCNE4*	2.75
C8ORF8/SGK1	2.17	*KIT*	2.20
CDH7	2.26	*KRT18*	2.99
CLDN5	-2.47	*MAFA*	-2.02
DCN	2.35	*MATN3*	-2.70
DESI2	2.22	*MIF*	2.01
FGFBP1	2.03	*MX1*	2.04
FOSB	3.62	*PDGFRA*	-2.10
HMGB1	2.04	*PDGFRB*	2.04
HUWE1	2.04	*RNASE4*	2.01
IFI27	2.14	*STK11/LKB1*	-2.38
IGFBP5	-2.02	*TM4SF1*	3.16

[a] All fold changes have a Q value of <0.05.

Figure 2.1 Cancer-related functions of selected differentially expressed genes
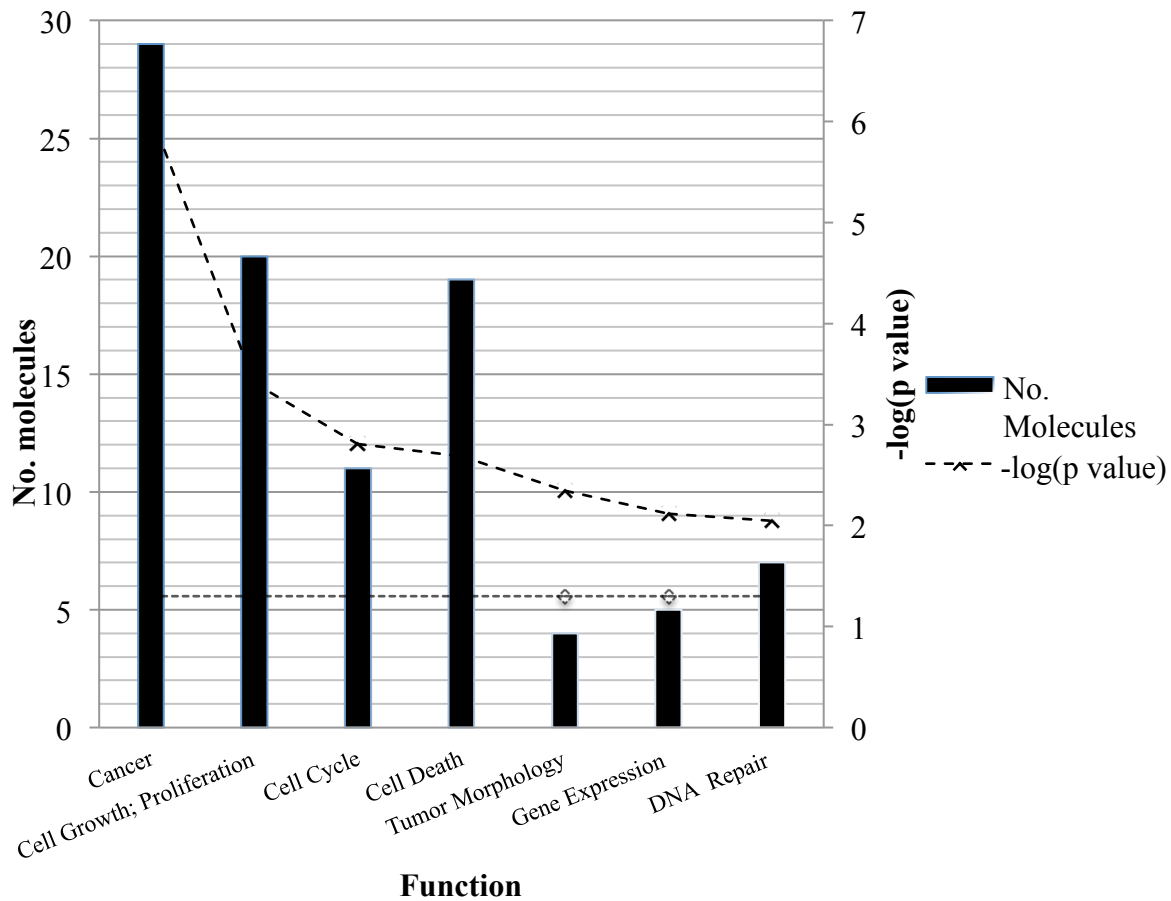
Figure 2.1: Cancer-related functions of differentially expressed transcripts.

Seven biological functions associated with cancer were identified using the IPA software: cancer, cell growth and proliferation, cell cycle, cell death, tumor morphology, gene expression, and DNA repair. The left y-axis indicates the number of identified molecules for each function represented by the bars. The right y-axis indicates the –log of the pooled p-values for each function, represented by the marked line. The statistical significance threshold of –log(p=0.05) is given by the straight line.

Figure 2.2 Differentially expressed genes in the "Post-translation modification, cellular development, and cellular growth and proliferation" network
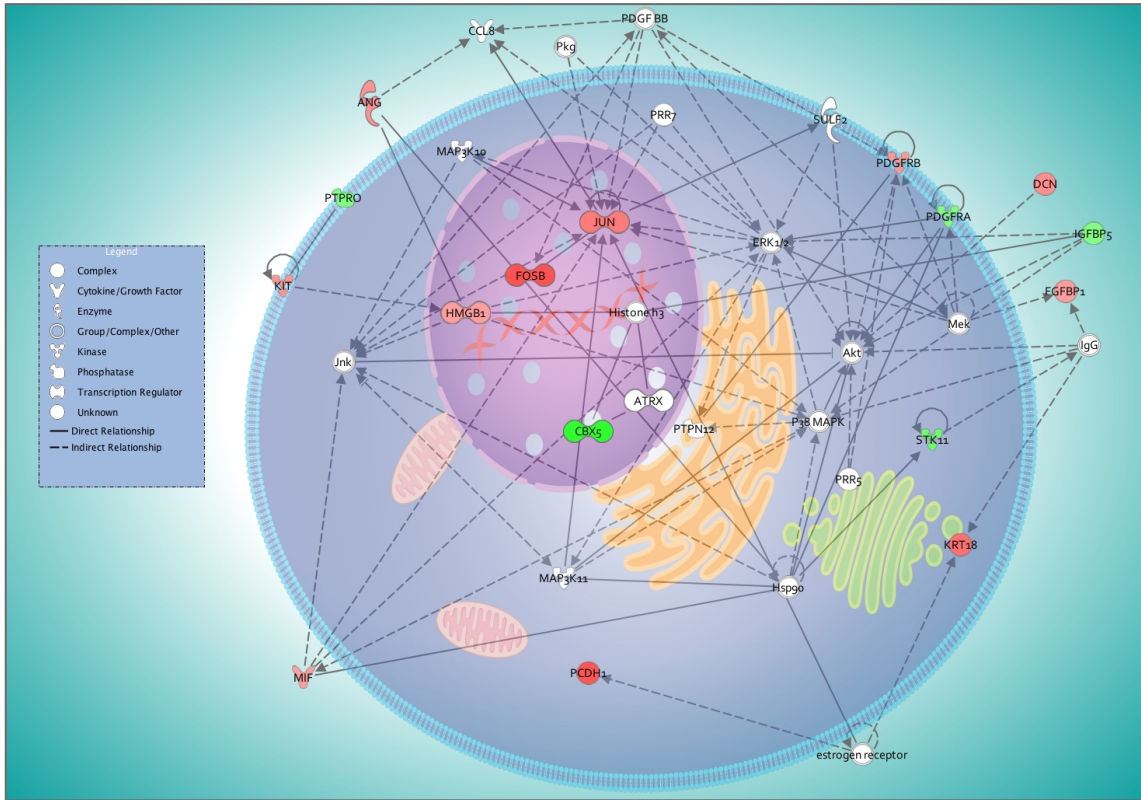
Figure 2.2 Differentially expressed genes in the "Post-translational modification, cellular development, and cellular growth and proliferation" network

Network of differentially expressed transcripts following miR-618 transfection under the "post-translational modification, cellular development, and cellular growth and proliferation" network. This network was identified by the Ingenuity Pathway Analysis software as significantly associated with the set of transcripts with altered expression following the introduction of the miR-618 mimic (pooled p-value = 8.61 x 10^{-7} and FDR $\geq$ |2|). Overexpressed transcripts are indicated in red, and underexpressed transcripts are indicated in green. Transcripts indicate the relative location by either in the nucleus, cytoplasm, cell membrane, or extracellular matrix.

Figure 2.3 miR-618 is overexpressed in HeLa cells relative to End1 cells.

miR-618 Relative Expression
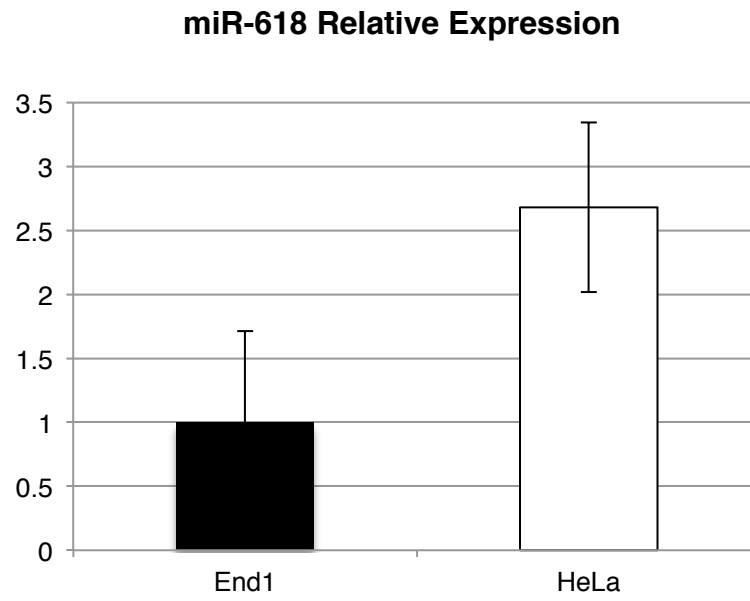
Figure 2.3 miR-618 is overexpressed in HeLa cells relative to End1 cells

qPCR results of quantifying miR-618 levels in HeLa cells relative to End1 cells reveals a 2.68 fold change in cancer vs. non-cancer cells. Expression levels were quantified using the $2^{-\Delta\Delta Ct}$ method with normalization to the stably expressed ncRNA U6B, where $\Delta\Delta Ct = (C_{t:miR\text{-}618\_HeLa} - C_{t:U6B\_HeLa}) - (C_{t:LmiR\text{-}618\_End1} - C_{t:U6B\_End1})$. Error bars represent the standard deviation of the C_T values.

cancer among Caucasian males (OR = 4.8, p = 2.6 x 10^{-3}), thus highlighting the importance of this genetic region in cancer.[69]

CHAPTER 3: MEX3D IS OVEREXPRESSED ACROSS CANCER TYPES

3.1: INTRODUCTION

RNA-binding proteins (RBPs) form a class of highly evolutionarily conserved post-transcriptional regulators found in a variety of organisms. Recent studies underscore the importance of such proteins in a variety of physiological functions, including stem cell differentiation, embryogenesis, and cancer.[32] MEX-3 is a translational repressor and RNA-binding proteins originally discovered in *C. elegans*. Vertebrates have four MEX-3 analogues, and are highly evolutionarily conserved across species. The MEX-3 family plays an important role in a number of cellular functions, but most recently they have been found to play a central role in cell differentiation and cell renewal, and are thus of major concern to stem biology and cell biology.[45]

Importantly, *MEX-3* family members are highly evolutionarily conserved throughout species. Phylogenetic analysis reveals that ancestral *mex-3* underwent two round of duplication after vertebrate lineage divergence, which accounts for the presence of four homologues in both mice and humans.[50] MEX-3 contains two K homology (KH) domains that contain RNA-binding activity. However, vertebrates contain an additional domain, perhaps a neo-functionalization that may have arisen during species divergence, the carboxy terminus really interesting new gene (RING) finger domain, which is absent in *C. elegans*. RING domains have been shown to mediate E3 ubiquitin ligase activity. Such activity has been shown in MEX-3C, which

is responsible for degradation of human leukocyte antigen serotype A2 (*HLA-A2*) mRNA.[51]

While the primary function of MEX-3 proteins in humans is unknown, recent evidence suggests they carry a role in cancer and other diseases, including hypertension, metabolic diseases, and cancer.[52-54] Interestingly, MEX3A overexpression in an intestinal cell line lead to impaired differentiation and altered polarity.[55] Further, MEX-3 in *C. elegans* represses cell-cycle inhibitor cyclin-dependent kinase inhibitor-2 (CKI-2)[56], whose human orthologues p21, p27, and p57 contain tumor suppressive activity.[45] Furthermore, silencing of *MEX-3C* causes chromosome mis-segregation, structural abnormalities, and DNA replication stress.[45] However, the overall function of MEX3 proteins in humans is largely unknown, especially in the context of cancer development. As such, this study aimed to profile the expression of MEX3D as well as the other members of the MEX3 family, MEX3A-C, in a number of tumors compared to normal tissue in order to guide molecular studies presented in **Chapter 4**.

Case-control studies deposited in the Oncomine database (http://www.oncomine.com) reveal that MEX3 family members are highly overexpressed in many types of cancer. In particular, we show here that MEX3D is overexpressed in melanoma, sarcoma, glioma, pancreatic cancer, cervical cancer, colorectal cancer, lymphoma, gastric cancer, and oral cancer. These findings are crucial, as they point to clear evidence that MEX3 family members play a role in tumor development. Future studies are needed to understand the mechanistic role of the proteins.

3.2 RESULTS

MEX3D is overexpressed in many types of cancer

According to the Oncomine database, *MEX3D* is overexpressed in five types of gastrointestinal cancers. In gastric cancer, there was a fold change of 2.52 (p = 2.03×10^{-8}) in *MEX3D* measured in 31 paired gastric carcinoma and adjacent normal gastric mucosa. The results are from a human genome U133 Plus 2.0 Array.[84] 70 colorectal carcinoma and 12 normal colon samples were analyzed with the U133 Plus 2.0 Array. There was a 2.59-fold change in colorectal cancer vs. normal samples (p = 2.04×10^{-8}).[85] In pancreatic cancer, 36 pancreatic carcinoma and 16 paired normal samples were analyzed with the Human Genome Plus 2.0 Array. The analysis revealed that *MEX3D* is 2.88-fold overexpressed in pancreatic cancer vs. normal tissue (p = 1.89×10^{-9}).[86] Another study looked at 7 rectal adenoma samples and 25 colon adenoma samples matched with 32 normal samples using the U133 Plus 2.0 Array. The study found that *MEX3D* is 2.76-fold overexpressed in rectal adenoma vs. normal tissue (p = 9.33×10^{-6}) and 2.09-fold overexpressed in colon adenoma vs. normal tissue (p = 5.76×10^{-11}).[87]

MEX3D was also overexpressed in skin cancer. The gene expression analysis in melanoma included 45 cutaneous melanoma samples, 18 benign melanocytic skin samples, and 7 normal skin samples analyzed with the U133 Plus 2.0 Array. The analysis revealed that *MEX3D* is overexpressed 2.42-fold in benign melanoma vs. normal skin (p = 2.58×10^{-6}), and 4.41 in cutaneous melanoma vs. normal skin (p = 6.96×10^{-9}).[88]

The Oncomine database also reveals that *MEX3D* is overexpressed in at least two reproductive cancers. In prostate adenocarcinoma, using 69 prostate tumor samples, 18 adjacent normal prostate samples, and 2 pooled normal prostate samples with the U133 Plus 2.0 Array, *MEX3D* was found to be overexpressed in prostate tumor vs. normal tissue by 2.10-fold ($p = 9.70 \times 10^{-4}$).[89] In cervical cancer, 20 cervical cancers were used with 8 cervical normal samples using the U133 Plus 2.0 Array. The study found *MEX3D* was overexpressed in cervical cancer vs. normal tissue by 2.80-fold ($p = 2.43 \times 10^{-7}$).

Five types of sarcomas showed overexpression of *MEX3D* compared to normal tissue. In a study of 149 soft tissue sarcomas compared to 9 normal adipose tissue specimens, *MEX3D* was overexpressed 3.11-fold in pleomorphic liposarcoma vs. normal was ($p = 5.51 \times 10^{-8}$), 2.99-fold in myxofibrosarcoma vs. normal ($p = 2.82 \times 10^{-10}$), 2.33-fold overexpressed in dedifferentiated liposarcoma vs. normal ($p = 4.28 \times 10^{-8}$), and 2.73 in myxoid/round cell liposarcoma vs. normal tissue ($p = 9.35 \times 10^{-8}$).[90] In a separate study of 39 sarcoma samples and 15 normal tissue samples using the Human Genome U133A Array, *MEX3D* was overexpressed by 3.02-fold in fibrosarcoma vs. normal tissue ($p = 5.48 \times 10^{-6}$).[91]

A few other types of cancer also exhibited overexpression of *MEX3D* compared to normal tissue. Tongue cancer, a form of oral cancer, exhibits a 2.12-fold overexpression compared to normal tissues ($p = 2.11 \times 10^{-4}$).[92] The study used 42 head and neck cancers compared to 14 head and neck normal samples on the Human Genome U133 Plus 2.0 Array. *MEX3D* was also 2.84-fold overexpressed in germinal center B-cell-like diffuse large B-cell lymphoma vs. normal ($p = 4.85 \times 10^{-5}$), 2.37-

fold overexpressed in activated B-cell-like diffuse large B-cell lymphoma vs. normal ($p = 1.45 \times 10^{-8}$), and 2.35-fold overexpressed in diffuse large B-cell lymphoma vs. normal ($p = 1.57 \times 10^{-12}$).[93] A total of 73 diffuse large B-cell lymphoma samples and 20 normal B-lymphocytes were analyzed using the Human Genome U133 Plus 2.0 Array. Lastly, in a study of 547 glioblastoma samples compared to 10 normal brain samples, *MEX3D* was found to be 3.00-fold overexpressed in glioblastoma vs. normal tissue ($p = 8.65 \times 10^{-9}$).[94] These data are summarized in **Figure 3.1** and **Figure 3.2.**

MEX3A-C are also overexpressed in many types of cancer

We also examined the expression levels of the other members of the MEX3 family, MEX3A-C. We found that these three genes displayed an expression profile similar to that as MEX3D. We found that MEX3A was overexpressed in breast cancer, melanoma, colorectal cancer, lung cancer, glioma, ovarian cancer, uterine cancer, and prostate cancer. Notably, MEX3A is 19.6-fold overexpressed in basal cell carcinoma compared to normal skin tissue ($p = 1.47 \times 10^{-7}$).

We found that MEX3B is overexpressed in breast cancer, uterine cancer, lung cancer, and leukemia compared to normal tissue. Lastly, we examined the expression profile of MEX3C in cancer compared to normal tissue, and found overexpression in cervical cancer, glioblastoma, esophageal cancer, oral cancer, leukemia, sarcoma, and gastric cancer. The cutoff values of *p-value* < 1.0 x 10^{-4} and fold-change > |2.0| were used to search the Oncomine database.

3.3 DISCUSSION

Human Mex-3A proteins, including hMex-3A, -3B, -3C and -3D, are located on distinct chromosomes at *1q22*, *15q25.2*, *18q21.1* and *19p13.3*, respectively (9). All the four proteins contain two highly conserved KH domains specifically for RNA binding and a RING finger domain in their carboxy- terminal region probably involved in mediating protein-protein interactions (15-17). The similarity in structure of the hMex-3 proteins implies that they may share certain similar biological functions. In this study, we investigated the expression profiles of the MEX3 family proteins in cancer compared to normal tissue. The data demonstrated that MEX3 family members are overexpressed in a broad range of types of cancer. Notably, MEX3B is nearly 20-rfold overexpressed in basal cell carcinoma compared to normal skin tissue. Furthermore, we found that MEX3D is overexpressed across a broad range of cancers, including melanoma, sarcoma, glioma, pancreatic cancer, cervical cancer, colorectal cancer, lymphoma, gastric cancer, and oral cancer.

Interestingly, one study associated the closely related gene *MEX3A* with increased cell proliferation and migration in human gastric cancer cells.[95] However, the literature sheds little light otherwise on the role of the proteins in human and especially their relation to disease development. This study is the first to describe *MEX3A-D* expression in a variety of tumors *as* displaying possible oncogenic activity. Few studies have looked at this gene family, and none have yet associated *MEX3D* with cancer. The finding that *MEX3D* is overexpressed in many types of cancer is alarming and warrants further investigation. These findings are crucial, as they point

to clear evidence that MEX3 family members play a role in tumor development.

Future studies are needed to understand the mechanistic role of the proteins.

Figure 3.1 Overexpression of MEX3D in tumors
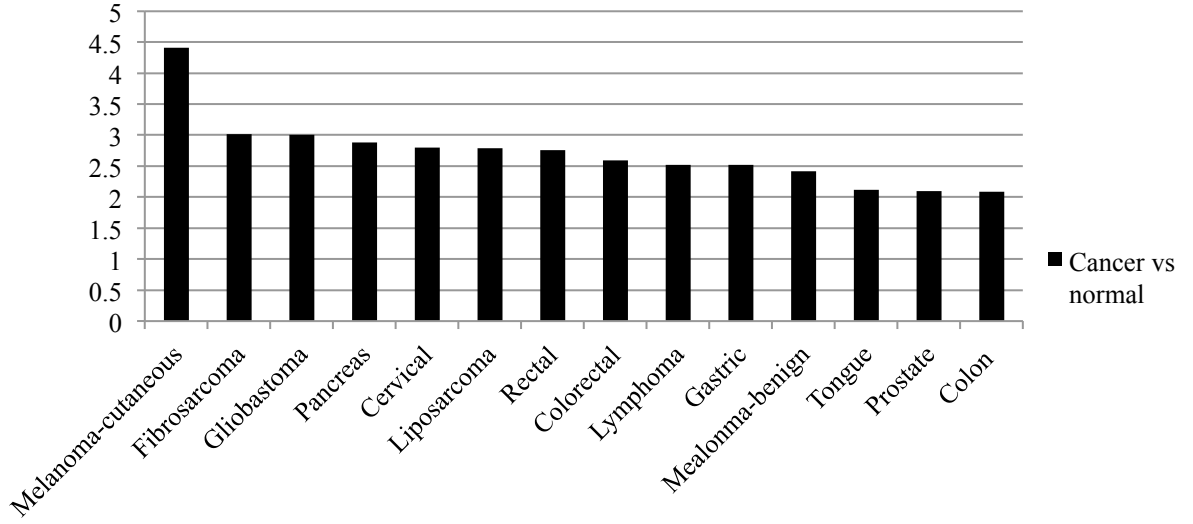
Figure 3.1 Overexpression of MEX3D in tumors

Using fold-change cutoff of $> | 2.0 |$ and *p*-value cutoff of 1×10^{-4}, the Oncomine

database was searched for cancer vs. normal expression data for *MEX3D*. The search

yielded studies in which *MEX3D* was only overexpressed.

Figure 3.2 MEX3A-C are also overexpressed across tumor types

A.
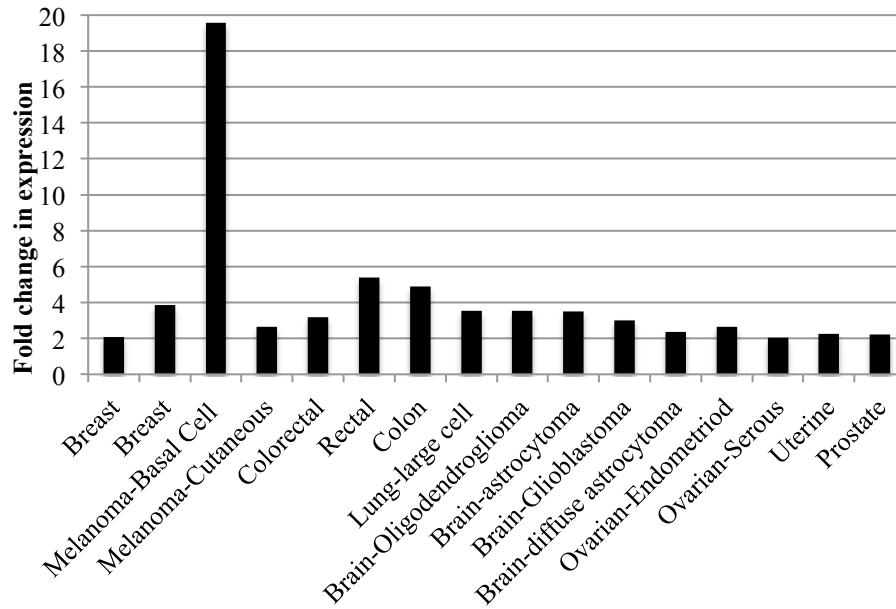
B.
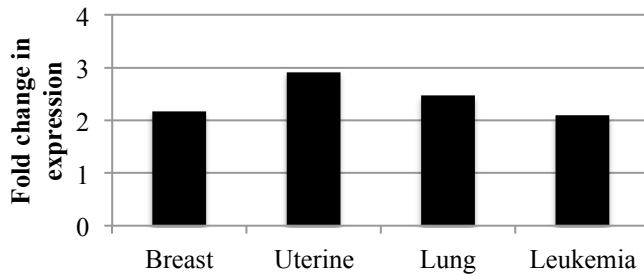
C.
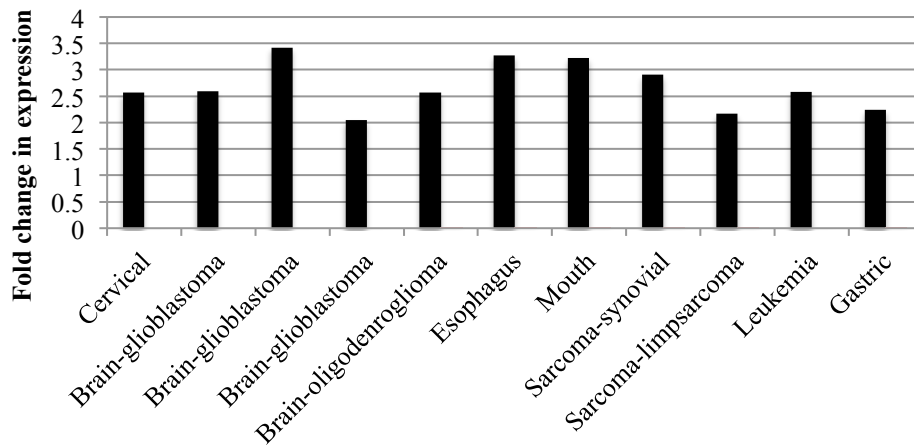
Figure 3.2 MEX3A-C are also overexpressed across tumor types

Using fold-change cutoff of $> | 2.0 |$ and *p*-value cutoff of 1×10^{-4}, the Oncomine

database was searched for cancer vs. normal expression data for *MEX3A* (**A**), *MEX3B*

(**B**), and *MEX3C* (**C**). The search yielded studies in which *MEX3A-C* were only

overexpressed.

Table 3.1 Expression profiles of MEX3A-D in cancer vs. normal tissue

Cancer	Expression vs. normal	Cancer	Expression vs. normal
MEX3A			
Breast	2.09	Brain-Oligodendroglioma	3.53
Breast	3.85	Brain-astrocytoma	3.50
Melanoma-Basal Cell	19.57	Brain-Glioblastoma	3.01
Melanoma-Cutaneous	2.64	Brain-diffuse astrocytoma	2.35
Colorectal	3.18	Ovarian-Endometriod	2.65
Rectal	5.38	Ovarian-Serous	2.03
Colon	4.89	Uterine	2.26
Lung-large cell	3.53	Prostate	2.20
MEX3B			
Breast	2.17	Lung	2.48
Uterine	2.91	Leukemia	2.41
MEX3C			
Cervical	2.57	Mouth	3.23
Brain-glioblastoma	2.59	Sarcoma-synovial	2.90
Brain-glioblastoma	3.42	Sarcoma-limpsarcoma	2.17
Brain-glioblastoma	2.05	Leukemia	2.58
Brain-oligodenroglioma	2.56	Gastric	2.25
Esophagus	3.27		
MEX3D			

Lymphoma	2.84	Sarcoma-fibrosarcoma	3.02
Cervical	2.80	Sarcoma-synovial	2.73
Melanoma-Cutaneous	4.41	Brain-glioblastoma	3.05
Melaonma-Benign	2.42	Pancreas	2.27
Renal	4.31	Tongue	2.12
Sarcoma-liposarcoma	3.11	Colorectal	2.59
Sarcoma-myxofibrosarcoma	2.99	Rectal Adenoma	2.76
Sarcoma-myxoid	2.73	Colon	2.09
Sarcoma-fibrosarcoma	3.02	Gastric	2.52
Sarcoma-synovial	2.73	Prostate	2.10

CHAPTER 4: *MEX3D* IS AN ONCOGENE IN GLIOMA

4.1 INTRODUCTION

RNA-binding proteins (RBPs) form a class of highly evolutionarily conserved post-transcriptional regulators found in a variety of organisms. Recent studies underscore the importance of such proteins in a variety of physiological functions, including stem cell differentiation, embryogenesis, and cancer.[32] MEX-3 is a translational repressor and RNA-binding proteins originally discovered in *C. elegans*. Vertebrates have four MEX-3 analogues, and are highly evolutionarily conserved across species. The MEX-3 family plays an important role in a number of cellular functions, but most recently they have been found to play a central role in cell differentiation and cell renewal, and are thus of major concern to stem biology and cell biology.[45]

Importantly, *MEX-3* family members are highly evolutionarily conserved throughout species. Phylogenetic analysis reveals that ancestral *mex-3* underwent two round of duplication after vertebrate lineage divergence, which accounts for the presence of four homologues in both mice and humans.[50] MEX-3 contains two K homology (KH) domains that contain RNA-binding activity. However, vertebrates contain an additional domain, perhaps a neo-functionalization that may have arisen during species divergence, the carboxy terminus really interesting new gene (RING) finger domain, which is absent in *C. elegans*. RING domains have been shown to mediate E3 ubiquitin ligase activity. Such activity has been shown in MEX-3C, which

is responsible for degradation of human leukocyte antigen serotype A2 (*HLA-A2*) mRNA.[51]

While the primary function of MEX-3 proteins in humans is unknown, recent evidence suggests they carry a role in cancer and other diseases, including hypertension, metabolic diseases, and cancer.[52-54] Interestingly, MEX3A overexpression in an intestinal cell line lead to impaired differentiation and altered polarity.[55] Further, MEX-3 in *C. elegans* represses cell-cycle inhibitor cyclin-dependent kinase inhibitor-2 (CKI-2)[56], whose human orthologues p21, p27, and p57 contain tumor suppressive activity.[45] Furthermore, silencing of *MEX-3C* causes chromosome mis-segregation, structural abnormalities, and DNA replication stress.[45] Data from the Oncomine database (http://www.oncomine.com) strengthens the possible role of MEX-3 family members in cancer, as there are large amounts of data that suggest its overexpression in cancer compared to normal tissue.

MEX-3D, a member of the hMEX-3 family, is expressed in somatic tissue in humans, contrasting with the expression profile of its analogues, which are restricted to the embryo and the germ line. Interestingly, the Oncomine database reveals that MEX-3D is highly overexpressed in a number of cancer samples compared to normal tissue. Yet, there is a gross lack of understanding of the role of this protein in both normal cell physiology and in cancer. The goal of this study was to determine the effects of MEX3D in glioma. We found that MEX3D is 3.00-fold overexpressed in glioma tissue compared to normal brain tissue. Kaplan-Meier analysis reveals that overexpression of MEX3D also leads to poor survival in glioma patients. Furthermore, we found that MEX3D leads to increased cell proliferation in A172 glioma cells.

Concurrently, MEX3D also leads to decreased cell invasion in both A172 and U87 glioma cells. This study is the first to report the effect of MEX3D in cancer, and describes MEX3D as a potential oncogene in glioma. Further studies are warranted to examine the mechanistic role of MEX3D in gene regulation and in tumorigenesis.

4.2 RESULTS

MEX3D **is overexpressed in glioma and leads to poor survivals.**

In order to determine the clinical relevance of *MEX3D* in glioma, we searched the Oncomine database for expression data in glioma compared to normal tissue. *MEX3D* is highly overexpressed in glioblastoma. In a study of 547 glioblastoma samples compared to 10 normal brain samples, *MEX3D* was found to be 3.00-fold overexpressed in glioblastoma vs. normal tissue ($p = 8.65 \times 10^{-9}$) (**Figure 4.1**).[94] We also examined the effects of *MEX3D* overexpression on overall survival in glioma patients. *MEX3D* expression data from n=273 glioma patients were analyzed using the Human Genome U133 Plus 2.0 Array. The Kaplan-Meier curve was calculated using a median-cutoff expression of 64.1 and two curves were calculated: a high group (n=143) with *MEX3D* expression greater than the median 64.1 *MEX3D* expression value, and a low group (n=130) with *MEX3D* expression lower than the median 64.1 expression value. After correcting for multiple comparisons using the Bonferoni method, there was a statistically significant difference between the high- and low-*MEX3D* expression profile groups ($p = 6.2 \times 10^{-3}$).

Differential methylation in *MEX3D*

In order to understand the mechanism underlying the change in gene expression in cancer, we searched for evidence of differential methylation in *MEX3D* in glioma samples. We tested the methylation status of twelve matched glioma and normal samples. The normal tissues came from the same patient and were adjacent normal tissue to the tumor site. The analysis revealed that three CpG sites showed significant differential methylation ($p<0.05$) between tumor and normal tissue. One site in the body of *MEX3D* showed hypermethylation of 24% ($\Delta\beta=0.2477$, $p=3.4\text{x}10^{-4}$) in tumor vs. normal tissue. Concurrently, a second site in the body of *MEX3D* was hypomethlyated in tumor vs. normal tissue by 15.4% ($\Delta\beta= -0.154$, $p=0.001$). A third location in *MEX3D* in the 3'UTR was found to be hypomethylated by 9.5% ($\Delta\beta= -0.095$, $p=0.041$).

MEX3D leads to increased cell proliferation in A172 glioma cells

As unregulated cell proliferation is a clear hallmark of cancer, we tested whether knockdown of *MEX3D* has an effect on cell proliferation *in vitro* using RNA interference against MEX3D in A172 glioma cells. As shown in Figure **4**, the chemically synthesized siRNA exhibited an efficient knockdown of MEX3D relative to the siRNA control. There was approximately a 10-fold reduction in expression of the knockdown compared to the control measured by qPCR. Remarkably, cell proliferation assays showed that si-RNA knockdown of *MEX3D* leads to significant decrease in cell proliferation in A172 glioma cells after five days compared to the negative control si-RNA transfected cells ($p <0.001$) (**Figure 4.2**). The negative

control exhibited normal cell growth, whereas the knockdown showed linear growth at best from days 1-4.

MEX3D leads to decreased cell invasion *in vitro*

In order to test the effects of MEX3D on another hallmark of cancer, cell invasion, we performed a cell invasion assay. *MEX3D* was knocked down in A172 glioma cells using si-RNA and U87 cells and compared to negative control si-RNA. Unexpectedly, knockdown of *MEX3D* led to significant increases in cell invasion, suggesting that MEX3D normally suppresses cell invasion. The relative fold increase in *MEX3D* knockdown compared to the negative control in A172 cells was 4.80 ($p = 0.003$), and in U87 cells was 2.58 ($p = 1.67 \times 10^{-5}$) (**Figure 4.5**).

4.3 DISCUSSION

Human Mex3 proteins, including hMex3A, -3B, -3C and -3D, are located on distinct chromosomes at *1q22*, *15q25.2*, *18q21.1* and *19p13.3*, respectively (9). All the four proteins contain two highly conserved KH domains specifically for RNA binding and a RING finger domain in their carboxy- terminal region probably involved in mediating protein-protein interactions (15-17). The similarity in structure of the MEX3 proteins implies that they may share certain similar biological functions. In the present study, one of the four proteins, MEX3D, was selected to investigate its function in glioma cancer development and progression for the first time. The final data demonstrated that a reduction in the MEX3D expression level led to a suppression of cell growth. Furthermore, the silencing of MEX3D also resulted in a increase in the invasion ability of gastric cancer cells. These findings shed light on the role of MEX3 RNA-binding proteins, particularly for MEX3D in tumorigenesis.

In recent years, RNA interference has been employed as a mature and powerful strategy for exploring gene functions by downregulating the expression of targeted genes. We applied this technique to specifically knockdown MEX3D in A172 cells and in U87 cells. We found that MEX3D appears to promote cell proliferation while limiting tumor invasion. Futhermore, we found that MEX3D is overexpressed in glioma, and that overexpression of MEX3D in glioma patients leads to poorer rates of survival.

While it is well known that methylation is an important epigenetic effect and alters gene expression, the results of this study are inconclusive and cannot be interpreted. Several reasons explain this. First, the quality of the samples was compromised. The samples were stored in formalin fixed paraffin for a long period of time, thus contaminating the samples and degrading them. This was evidenced by heavy background noise in the data set. Further, a small sample size (n=12) limited the power of the study. The methylation of *MEX3D* is an important effect that warrants closer investigation and a bigger sample size.

The fact that MEX3D appears to promote cell proliferation while limiting tumor invasion is somewhat paradoxical. However, the evidence we present of the high level of over-expression of MEX3D as well as the poorer tumor survival in glioma patients with high expression of MEX3D indicates that this protein plays a major role in tumor development, especially in the context of glioma. Thus, we conclude that MEX3D is a potential oncogene.

In conclusion, our present study is the first to show that the knockdown of hMex-3A using RNA interference can effec- tively inhibit cell proliferation and migration in gastric cancer cells. These results further indicate that hMex-3A may serve as a potential target for the therapy of gastric cancer, although additional studies *in vivo* are necessary.

Figure 4.1 *MEX3D* **is overexpressed in glioma and leads to poor survival in glioma patients**

A

**p=8.65x10^{-9}

B
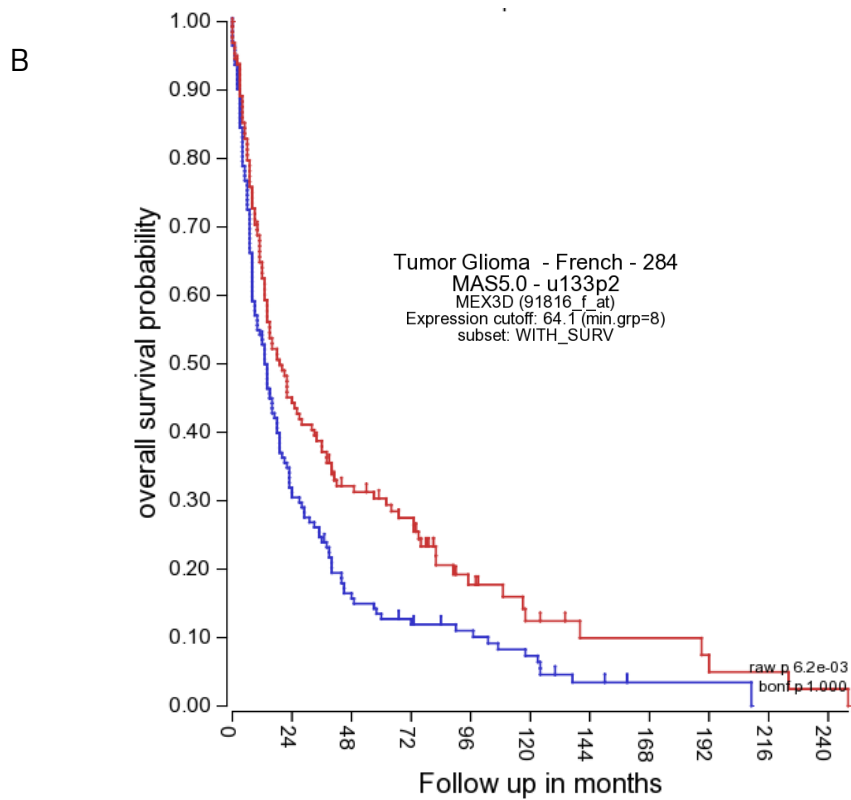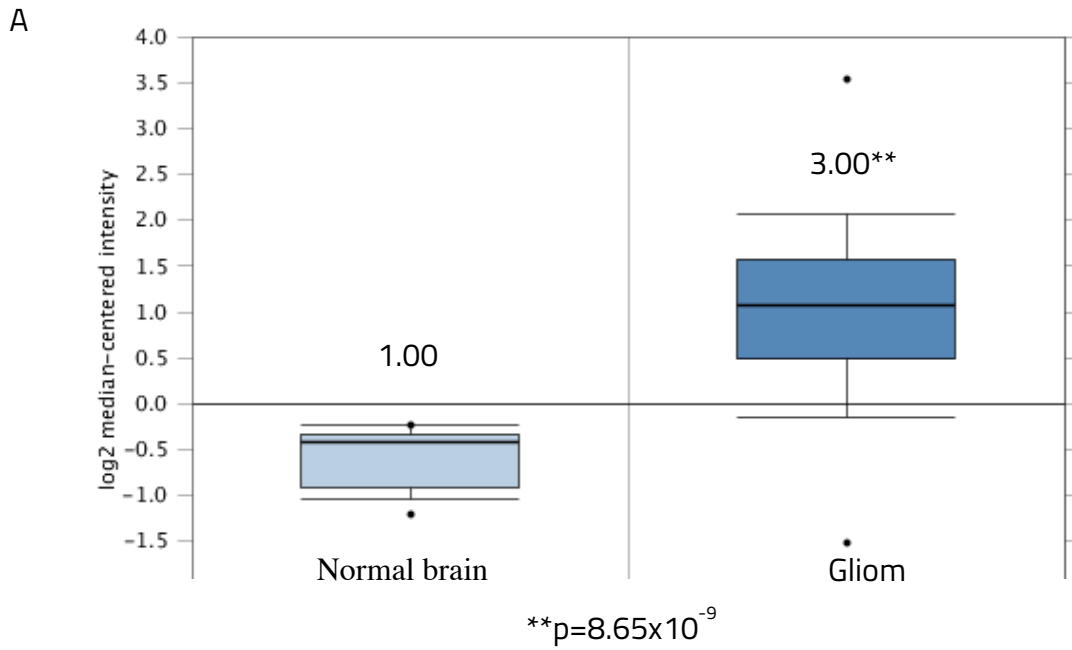
Figure 4.1 *MEX3D* **is overexpressed in glioma and leads to poor survival in glioma patients**

Oncomine database reveals that *MEX3D* is 3.00-fold overexpressed in glioma tumors vs. normal brain tissue ($p = 8.65 \times 10^{-9}$) (A). Using median-expression cutoffs in the Kaplan Meier survival analysis, overexpression of *MEX3D* is associated with poorer survival in all glioma patients (B). The red, top line represents lower expression of *MEX3D*, whereas the bottom blue line represents higher expression of *MEX3D*.

FIGURE 4.2 Methylation of CpG sites in *MEX3D* in glioma vs. normal samples
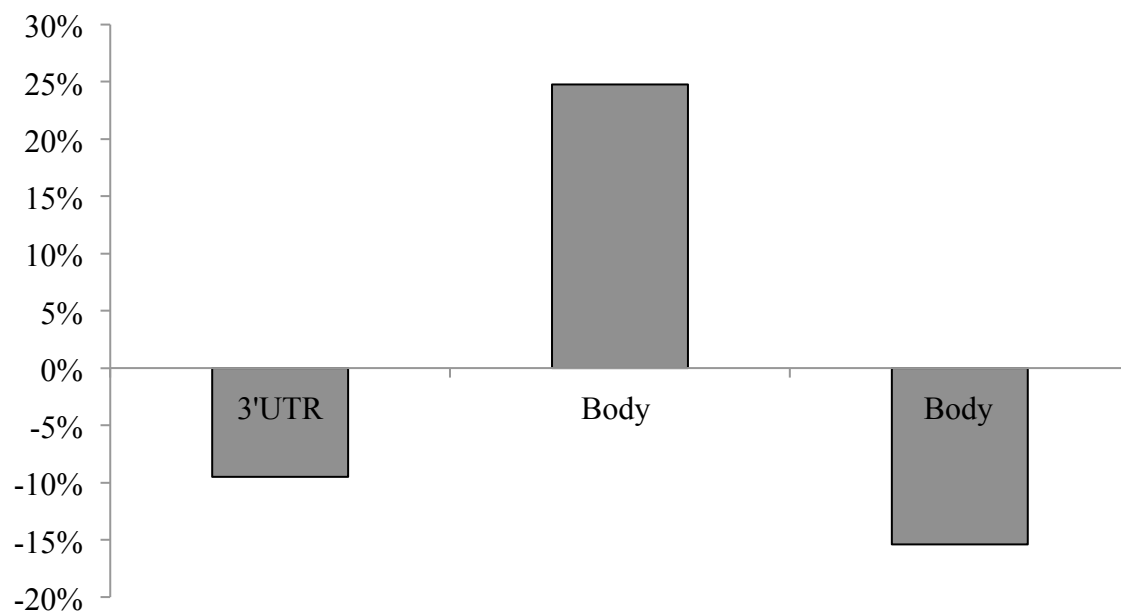
Figure 4.2 Methylation of CpG sites in *MEX3D* in glioma vs. normal samples

Analysis of methylation status of *MEX3D* reveals that one site in the body of *MEX3D* showed hypermethylation of 24% ($\Delta\beta$=0.2477, *p*=3.4x10^{-4}) in tumor vs. normal tissue. Concurrently, a second site in the body of *MEX3D* was hypomethlyated in tumor vs. normal tissue by 15.4% ($\Delta\beta$= -0.154, *p*=0.001). A third location in *MEX3D* in the 3'UTR was found to be hypomethylated by 9.5% ($\Delta\beta$= -0.095, *p*=0.041).

Figure 4.3 Knockdown of *MEX3D* in A172 cells
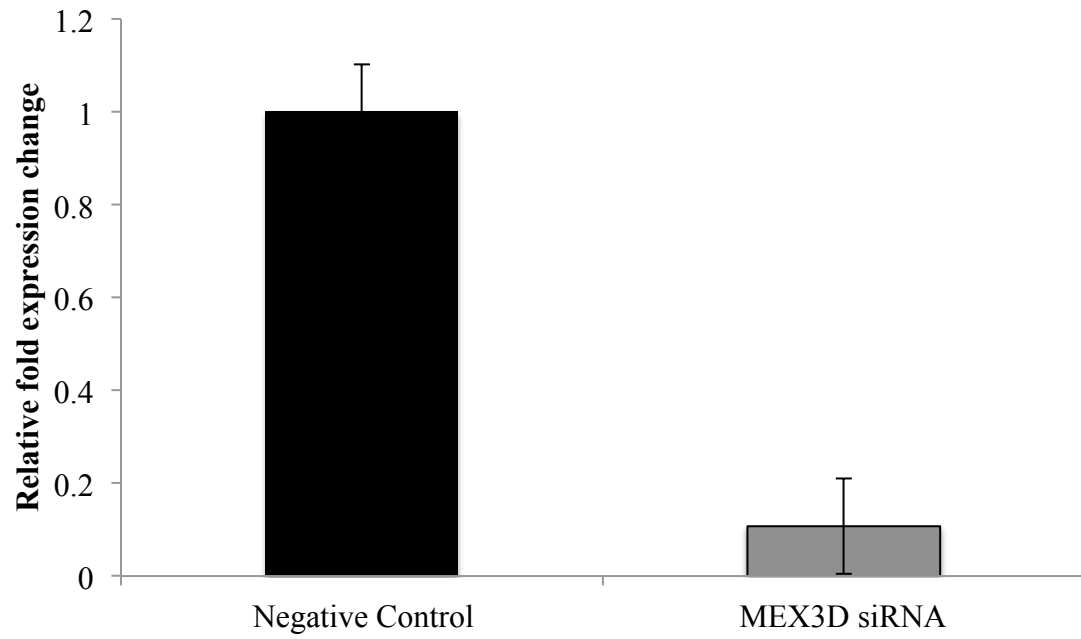
FIGURE 4.3 Knockdown of *MEX3D* in A172 cells

si-RNA knockdown of *MEX3D*. Expression levels were quantified using the $2^{-\Delta\Delta Ct}$ method with normalization to the stably expressed gene GAPDH, where $\Delta\Delta Ct = (C_{t:siRNA\text{-}NC\_A172} - C_{t:GAPDH\_A172}) - (C_{t:siRNA\text{-}MEX3D\_A172} - C_{t:APDH\_A172})$. Error bars represent the standard deviation of the C_T values.

Figure 4.4 *MEX3D* knockdown decreases cell proliferation in A172 glioma cells
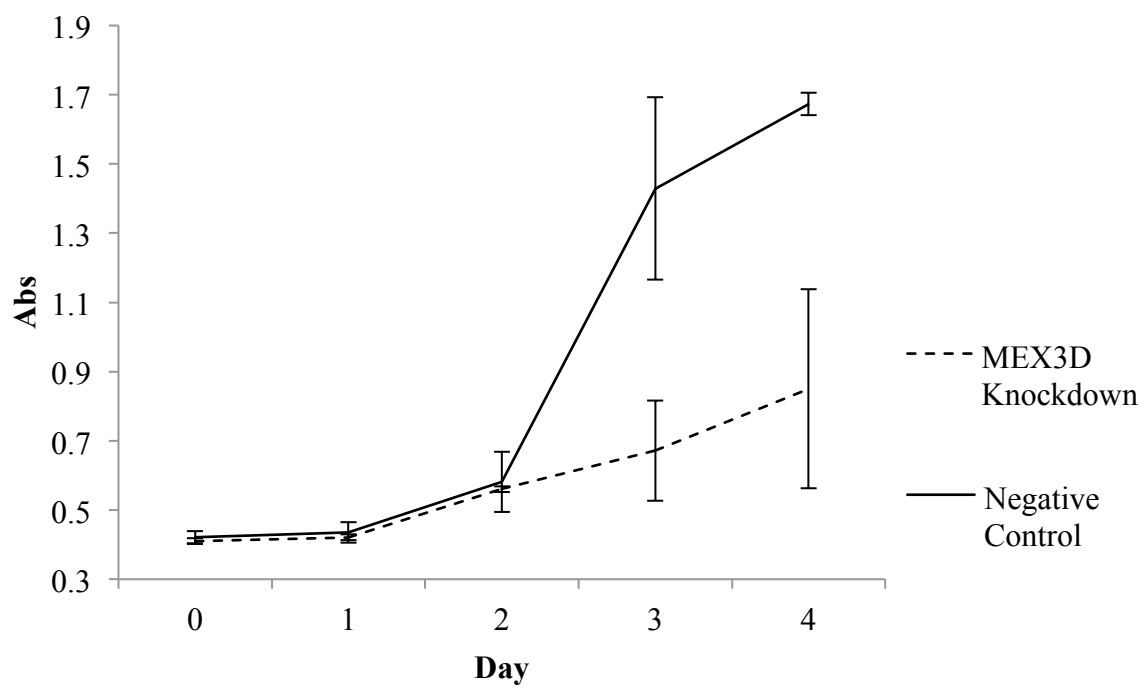
Figure 4.4 *MEX3D* knockdown decreases cell proliferation in A172 glioma cells

MEX3D was knocked down in A172 glioma cells using siRNA against *MEX3D*.

Knockdown of *MEX3D* (bottom dashed line) leads to significant decrease in cell

proliferation in A172 glioma cells after five days compared to the negative control si-

RNA transfected cells (top solid line) (p <0.001). The experiment was performed in

triplicates, and the error bars represent the standard deviation between the triplicate

experiments.

Figure 4.5 MEX3D knockdown increases cell invasion in A172 cells and U87 cells

A.
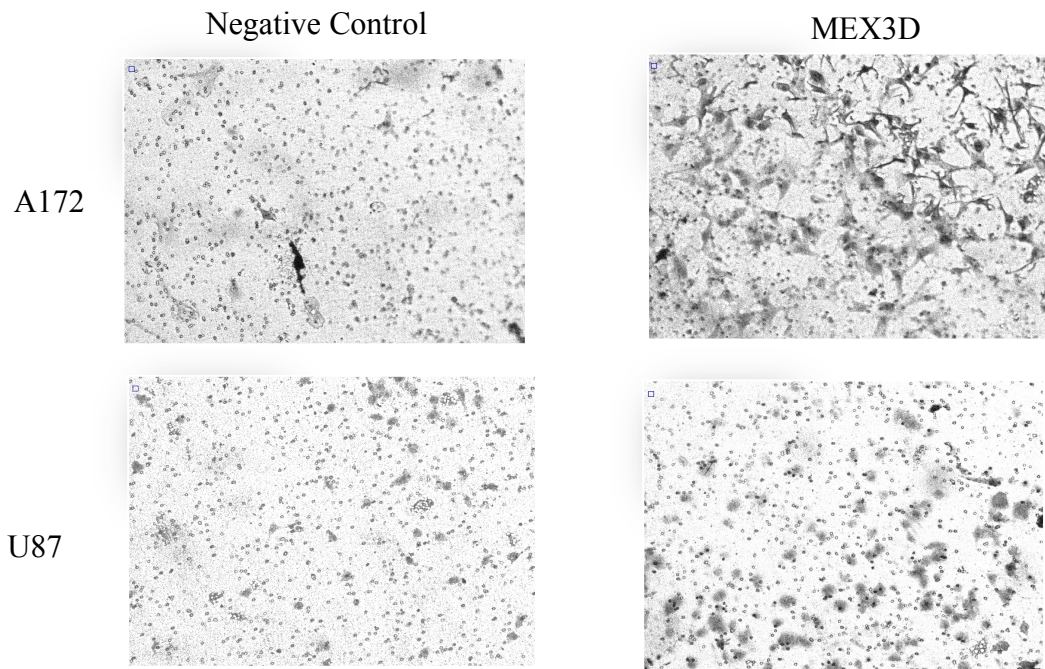
B
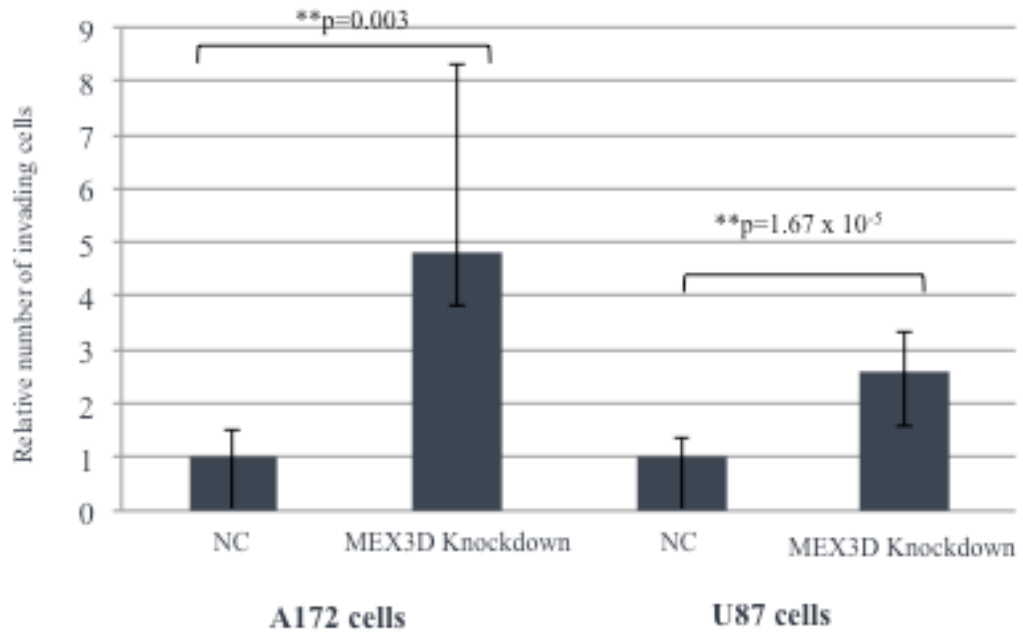
Figure 4.5 MEX3D knockdown increases cell invasion in A172 cells and U87

cells

MEX3D was knocked down using si-RNA against MEX3D in A172 cells and in U87

cells. Negative control si-RNA (NC) served as the control. The cells were transfected

and incubated for 24 hours. **A.** Ten random fields from each of the invasion assays

were selected for cell counting using phase-contrast microscopy. **B.** Error bars

represent the standard deviation between the ten fields. Statistical significance of thde

difference between the NC and the knockdown was determined using the Students t-

test.

CHAPTER 5: CONCLUSIONS

The control of gene expression is pivotal in the context of molecular pathogenesis of a number of diseases, and thus is of critical relevance to public health. An array of cellular tools exist in controlling gene expression, including epigenetic effects, non-coding RNAs, and RNA-binding proteins. These tools are critical tools to the modern of public health, and are used in tandem with population-based studies. This work focuses on specific examples of non-coding RNAs and RNA-binding proteins, describing the effects of microRNA-618 and MEX3D, a post-transcriptional regulator, in cancer.

MicroRNAs (miRNAs) form a class of highly conserved endogenous RNAs that inhibit gene expression and may act as oncogenes or as tumor suppressors, regulating extensive cancer-related gene networks. Here, we showed the association between a single miRNA, miR-618, and cancer-related pathways in HeLa cells. MiR-618 was identified as a potentially oncogenic microRNA, controlling a number of cancer-related gene networks and pathways. Gain-of-function analysis reveals [96]differential expression of 110 transcripts following miRNA-618 transfection. Notably, three upregulated genes are well-studied oncogenes—*KIT*, *JUN*, and *FOSB*—and three downregulated genes are well-known tumor suppressors—*PTPRO*, *STK11/LKB1*, and *IGFBP5*. Interestingly, investigation using the Ingenuity Pathway Analysis software tool reveals alterations in multiple cancer-related and cell cycle-related networks, including upregulated oncogenes in the top identified network "Post-translational modification, cellular development, cellular growth and proliferation" following miR-618 transfection. Further, miR-618 expression analysis

shows overexpression in HeLa cells compared to normal cervical cells. Our findings present evidence for a novel oncogenic miRNA, miR-618, that is involved in cancer-related gene networks and is overexpressed in cancer.

RNA binding proteins also play a critical role in regulating gene expression, and are often critical in tumor development. This work also examined the role of a novel post-transcriptional regulator, MEX3D, in cancer. [97] The Oncomine online database reveals that MEX3D is overexpressed in a number of solid tumors, notably in glioma. MEX3D is 3.01-fold overexpressed in glioma cells compared to non-cancerous, normal tissue. Kaplan-Meier survival analysis reveals that higher expression of *MEX3D* leads to poorer overall survival in overall glioma patients. Lastly, in a pilot case-control study of twelve glioma biopsies, we examined the effects of methylation in CpG sites in the *MEX3D* gene. The results were unclear, as we found a 3'UTR site that was 9.5% hypermethylated compared to normal tissue, a site in the body of the gene that was 24.8% hpyermethylated, and second site in the body that was 15.4% hypomethylated compared to normal tissue..

Phenotypic studies reveal that MEX3D is responsible for two cancer phenotypes. Knockdown of MEX3D leads to increased cell proliferation and decreased cell invasion, suggesting that overexpression of MEX3D is responsible for increased cell proliferation and decreased cell invasion.

This study is the first to describe the effects of miR-618 and of MEX3D in cancer. The findings presented in this work lay the foundation for further mechanistic studies of miR-618 and MEX3D. More work is needed to identify the mechanisms of

oncogenesis controlled by these molecules. Our study indicates that miR-618 may be

a biomarker for several types of cancer and warrants further investigation.

CHAPTER 6: Materials and methods

6.1 Cell culture and miRNA transfection (Chapters 2 and 4)

 HeLa cells (American Type Culture Collection) were cultured in Dulbecco's modified Eagle medium (Invitrogen) with 10% fetal bovine serum (Invitrogen) and 1% penicillin/streptomycin (Sigma-Aldrich). A172 cells will be purchased from ATCC (Manassas, VA) and maintained as monolayer cultures in 25 cm^2 polystyrene flasks in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% fetal bovine serum (FBS). Cells will be incubated at 37ºC in a humid atmosphere containing 5% CO_2 and subcultured every 2-3 days.

6.2 miRNA transfection (Chapter 2)

 A synthetic miRNA mimic of hsa-miR-618 (5'-AAACUCUACUUGUCCUUCUGAGU-3') was synthesized by Qiagen for transfection. Cells were transfected with either the miR-618 mimic or a scrambled negative control (Qiagen) using the Lipofectamine RNAiMAX transfection reagent (Invitrogen), according to the manufacturer's protocol. 120 pmol of miR-618 or negative control was mixed with 20 µl of RNAiMAX reagent and 2 ml OPTI-MEM (Invitrogen). The mixtures were transfected with ~100,00 cells in 12-well plates, and cells were harvested 48 hours post-transfection. RNA was subsequently extracted for analysis using the RNeasy Mini Kit (Qiagen), according to the manufacturer's protocol.

6.3 Genome-wide expression microarray (Chapter 2)

Gene expression differences in cells transfected with the miR-618 mimic were interrogated by whole genome microarray (Agilent, Inc 41k chip, performed by MoGene). The miRNA transfection and microarray were conducted in two independent biological replicates. Due to the multiple comparisons inherent in the microarray analysis, the false discovery rate (FDR) method as described by Benjamini and Hochberg was applied in order to obtain an adjusted p-value (Q).[98] Transcripts with inadequate signal intensity (Intensity < 50) were discarded. Expression differences were considered to be statistically and biologically significant at Q < 0.05 and fold change $\geq |2|$.

6.4 Network Analysis (Chapter 2)

Genes that were differentially expressed from the miR-618 transfection were investigated for functional interrelatedness using the IPA software tool (Ingenuity Systems). This tool elucidates networks and functionality by using information in the Ingenuity Pathways Knowledge Base, which is a database of known functional interactions that have been previously identified in peer-reviewed publications.[99] Statistical significance was obtained by performing a Fisher's exact test based on the hypergeometric distribution for each identified network to calculate the likelihood of obtaining at least the same number of related expressed genes by chance for the input gene set.

6.5 Expression analysis of miR-618 (Chapter 2)

Total RNA from HeLa cells and End1 cells (American Type Culture

Collection) was isolated using the miRNeasy Mini Kit (QIAGEN), with on-column

DNA digestion. To determine levels of mature miRNA, polyadenylated mature

miRNA sequences were first generated and converted to cDNA using the NCode

miRNA First-Strand cDNA Synthesis Kit (Invitrogen). The cDNA was then

amplified using a custom miR-618-specific forward primer (5'-

AAACTCTACTTGTCCTTCT -3') and a universal reverse primer targeting on the

polyadenylated region of the miRNA. Mature miRNA levels in HeLa cells relative to

End1 cervical cells were assessed using the $2^{-\Delta\Delta Ct}$ method with normalization to the

stably expressed ncRNA U6B, where $\Delta\Delta Ct = (C_{t:miR\text{-}618\_HeLa} - C_{t:U6B\_HeLa}) - (C_{t:LmiR\text{-}618\_End1} - C_{t:U6B\_End1})$. All qPCR reactions were performed in triplicate on an ABI

7500 Fast Real-Time PCR instrument (Applied Biosystems) using the Kapa

Biosystems SYBR Fast qPCR kits (Kapa Biosystems) according to the

manufacturer's protocols.

6.6 Oncomine Database Search (Chapters 3 and 4)

In order to determine differential expression of *MEX3D* in cancer vs. normal

tissue, the Oncomine Database was searched on Sunday, March 30, 2014 using the

search terms "MEX3D" and filtering for cancer vs. normal analysis

(http://www.oncomine.org/). The following filters were added: threshold fold change

of |2.0| and a threshold *p*-value of 1×10^{-4}.

6.7 Kaplan-Meier curve (Chapter 4)

MEX3D expression data from 273 glioma patients were analyzed using the Human Genome U133 Plus 2.0 Array. The Kaplan-Meier curve was calculated using a median-cutoff expression of 64.1 and two curves were calculated: a high group (n=143) with *MEX3D* expression greater than the median 64.1 value, and a low group (n=130) with *MEX3D* expression lower than the median 64.1 expression value. The Bonferoni method was to correct for multiple comparisons.

6.8 *MEX3D* methylation analysis (Chapter 4)

The Illumina Custom Model, as implemented in the Illumina GenomeStudio software, was used to assess the statistical significance of methylation differences between glioma cases and controls at each CpG site. This model operates under the assumption that the methylation value β is normally distributed at each locus. The variance function is estimated for all values of β by repeatedly measuring loci with known methylation fractions ranging from 0 to 1, and then fitting a parabola to the standard deviation as a function of β. The standard deviation estimate is then given by $s = A\beta^2 + B\beta + C$, where A = -0.1511, B = 0.1444, and C = 0.01646.

P-values were calculated using the following formula:

$$p = z\left(\frac{|\beta_{cond} - \beta_{ref}|}{\frac{s^2_{ref}}{N_{ref}} + \frac{s^2_{cond}}{N_{cond}}}\right)$$

where z denotes the standard score of a two-sided normal probability distribution, β_{ref} and β_{cond} denote the average β values for the reference (normal tissue controls) and condition (glioma tumor samples) groups, s^2_{ref} and s^2_{cond} denote the variances of β_{ref}

and β_{cond}, and N_{ref} and N_{cond} denote the sample sizes of the reference and condition groups. Due to the multiple comparative nature of our epigenome-wide analysis, *P*-values generated by the array were corrected for multiple comparisons using the Benjamini-Hochberg method [100] based on the number of CpG sites measured (485,577) to minimize the probability of false positives. Only CpG sites exhibiting methylation differences with corrected *P*-values of <0.05 were considered as significantly altered.

6.9 RNA interference and cell transfection (Chapter 4)

In order to suppress *MEX3D* expression in glioma cells, MEX3D-specific small interference RNA (siRNA) was chemically synthesized from Invitrogen (Catalog no. 1299001). A negative control siRNA served as the control for the cancer phenotype assays (Invitrogen). A172 cells cells will be reverse transfected in 12-well plates with either siRNA or the siRNA control using LipofectAMINE RNAiMAX transfection reagent (Invitrogen, Carlsbad, CA). Approximately 25,000 cells suspended in 10% FBS-supplemented DMEM will be added to a pre-incubated mixture of RNAiMAX (1 µl) and RNA oligo in each well to a final oligo concentration of 25 nM. Wells will be incubated at 37ºC and 5% CO_2 for 72 hours prior to RNA isolation (RNeasy Mini Kit, QIAGEN, Valencia, CA). All transfections were conducted in biological triplicate. Total RNA from A172 cells (American Type Culture Collection) was isolated using the RNeasy Mini Kit (QIAGEN), with on-column DNA digestion. To determine levels of mature mRNA, polyadenylated mature mRNA sequences were first generated and converted to cDNA using the NCode mRNA First-Strand cDNA Synthesis Kit (Invitrogen). The cDNA was then

amplified using a custom MEX3D specific forward primer (5'-CCAGGGCTGCAAGATCAAG-3'). Mature mRNA levels in A172 cells were assessed using the $2^{-\Delta\Delta Ct}$ method with normalization to the stably expressed gene GAPDH, where $\Delta\Delta Ct = (C_{t:siRNA-NC\_A172} - C_{t:GAPDH\_A172}) - (C_{t:siRNA-MEX3D\_A172} - C_{t:APDH\_A172})$. All qPCR reactions were performed in triplicate on an ABI 7500 Fast Real-Time PCR instrument (Applied Biosystems) using the Kapa Biosystems SYBR Fast qPCR kits (Kapa Biosystems) according to the manufacturer's protocols.

6.10 Cell proliferation assay (Chapter 4)

Effects of MEX-3D on cell proliferation will be measured using a cell proliferation assay kit from Invitrogen (Cat. No C35006). This compares *MEX-3D* knockdown in A172 cells to A172 cells treated with the negative control siRNA. This assay measures cell proliferation by quantifying DNA to indicate a relative cell number, and will be performed according to the manufacturer's protocols. Each time point was measure in triplicate and were measured in 24-hour intervals.

6.11 Cell invasion assay (Chapter 4)

A172 cells and U87 cells were seeded into the upper wells of 24-well Transwell plates on Matrigel-coated (BD Biosciences Inc., San Jose, CA, USA) 8-μm poly(ethylene terephthalate) membranes (Millipore Corpo., Billerica, MA , USA) in DMEM without fetal bovine serum (FBS). The lower wells contained complete medium with 10% FBS. After 48 h of incubation in a humidified atmosphere with 5% CO_2, the contents of the upper well content were removed using Q-tips. The

invasion chambers were processed following the manufacturer's protocols, and the migrated cells were stained with haematoxylin (??). Ten random fields from each of the invasion assays were selected for cell counting using phase-contrast microscopy.

6.12 Statistical analysis

With the exception of differential gene expression analysis, network analysis, and methylation analysis using the statistical methods described above, all other statistical analyses were evaluated using the Student's t-test. $P<0.05$ was considered to indicate a statistically significant difference.

CHAPTER 7: SUPPLEMENTARY INFORMATION

TABLES

Supplementary Table S1. All identified differentially expressed genes by miR-618 transfection with expression fold changes ≥ |2|

Gene Symbol	Description[a]	Fold change[b]
AU145408	Unknown function	5.24
A_24_P942036	Unknown function	4.66
A_32_P6841	Unknown function	4.46
AK026750	Unknown function	4.38
CCDC66	Coiled-coil domain containing 66; encoded protein linked to progressive retinal atrophy	4.30
A_32_P233713	Unknown function	4.06
TEX13B	Testis expressed 13b	4.04
A_32_P4882	Unknown function	3.79
FOSB	FBJ murine osteosarcoma viral oncogene homolog B. Encodes protein that can dimerize with proteins of the FUN family. Implicated in cell proliferation, differentiation, and transformation.	3.62
A_24_P611965	Unknown function	3.59
PCDH1	Protocadherin 1; encodes membrane protein found at cell-cell boundaries.	3.52
HBBP1	Hemoglobin, beta pseudogene 1.	3.49
NAIF1	Nuclear apoptosis inducing factor 1; encoded protein inhibits tumor cell proliferation by inducing apoptosis.	3.41
NM_052958	Unknown function	3.36
CA441361	Unknown function	3.16
FA2H	Fatty acid 2-hyrdoxylase. Encoded protein catalyzes synthesis of 2-hydroxysphingolipids.	3.13

KRT18	Keratin 18; encodes type I intermediate filament chain keratin 18. Expressed in epithelial tissues and mutations in *KRT18* have been linked to cryptogenic cirrhosis.	2.99
BF899060	Unknown function	2.93
COL8A2	Collagen, type VIII alpha 2. This encoded protein is a major component of the basement membrane of corneal epithelium.	2.90
SAMD9L	Sterile alpha motif domain containing 9-like	2.85
JUN	Jun proto-oncogene; Putative transforming gene of avian sarcoma virus 17. Encoded gene interacts with specific DNA to regulate gene expression.	2.79
KCNE4	Potassium voltage-gated channel, Isk-related family, member 4. Encoded protein is a type I membrane protein with a potassium channel.	2.75
BC063641	Unknown function	2.67
W95609	Unknown function	2.67
BIK	BCL2-interacting killer (apoptosis-inducing). Encoded protein has pro-apoptotic activity and its expression is suppressed in the presence of survival-promoting proteins.	2.66
AI263083	Unknown function	2.66
ZBED2	Zinc finger, BED-type containing 2	2.59
THC2340539	Unknown function	2.57
IL22RA1	Interleukin 22 receptor, alpha 1. Encoded protein belongs to class II cytokine receptor family and has been shown to be a receptor for IL22.	2.52
ABLIM3	Actin binding LIM protein family, member 3; protein contains LIM domain, double zinc finger. It plays roles in embryonic development, cell lineage determination, and cancer.	2.50
ENST00000379557	Unknown function	2.50
SRRM3	Serine/arginine repetitive matrix 3	2.50
THC2442829	Unknown function	2.44
THC2409451	Unknown function	2.40

ENST00000320662	Unknown function	2.39
THC2316768	Unknown function	2.38
CPM	Carboxypeptidase M; encoded protein is a membrane-bound arginine/lysine carboxypeptidase and is associated monocyte differentiation	2.38
EDN2	Endothelin 2; encodes member of endothellin protein family of secretory vasoconstrictive peptides.	2.36
DCN	Decorin; encoded protein is a small cellular or pericellular matrix proteoglycan that is a component of connective tissue, binding to type I collage fibrils and playing a role in matrix assembly.	2.35
BC038972	Unknown function	2.27
KCNJ10	Potassium inwardly-rectifying channel, subfamily J, member 10. Encodes potassium channel that allows potassium to flow into, rather than out of, a cell.	2.27
AA586832	Unknown function	2.26
CDH7	Cadherin 7, type 2; encoded protein is membrane protein that is a calcium dependent cell-cell adhesion glycoprotein	2.26
ANG	Angiogenin, ribonuclease, RNase A family, 5; encodes protein that strongly mediates blood vessel formation	2.25
BE671182	Unknown function	2.25
THC2398758	Unknown function	2.25
KREMEN2	Kringle containing transmembrane protein 2. Encodes a high-affinity dickkopf homolog 1 (DKK1) transmembrane receptor and complexes with DKK1 And lipoprotein receptor-related protein 6 (LRP6) to induce rapid endocytosis and removal of LRP6 from plasma membrane.	2.25
DESI2	Desumoylating isopeptidase 2	2.22
THC2436337	Unknown function	2.21
TM4SF1	Transmembrane 4L six family member 1; encodes protein of transmembrane 4 superfamily that mediates signal transduction to play a role in regulating cell development, activation, growth, and motility.	2.21
KIT	v-KIT Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog; encodes human homolog of proto-oncogene c-kit. Aberrant expression associated with	2.20

	cancer.	
ANXA8	Annexin A8; encodes protein Ca^{2+} and phospholipid binding proteins. Encoded protein may function as an anticoagulant that indirectly inhibits thromboplastin-specific complex. Overexpression of ANXA8 has been associated with acute myelocytic leukemia.	2.19
MAPK4	Mitogen-activated protein kinase 4; activated MAPK4 translocates to nucleus to phosphorylate nuclear targets.	2.17
SGKL	Serum/glucocorticoid regulated kinase family, member 3; encoded kinase phosphorylates several target proteins and plays role in neutral amino acid transport and activation of potassium and chloride channels.	2.17
A_24_P919283	Unknown function	2.15
NGEF	Neuronal guanine nucleotide exchange factor	2.15
FBXO2	F-box protein 2; encodes protein involved in phosphorylation-dependent ubiquitination.	2.15
LGSN	Lengsin, lens protein with glutamine synthetase domain; encoded protein is localized to the lens and by be associated with cataract disease.	2.14
AI915259	Unknown function	2.14
IFI27	Interferon alpha-inducible protein 27	2.14
THC2282944	Unknown function	2.13
ENST00000375256	Unknown function	2.12
N34499	Unknown function	2.11
ENST00000302001	Unknown function	2.10
SMIM15	Small integral membrane protein 15	2.10
BI963219	Unknown function	2.08
AF086329	Unknown function	2.08
PIGZ	Phosphatidylinosital glycan anchor biosynthesis, class Z; encodes protein that localizes to the endoplasmic reticulum and is involved in glycosylphosphatidylinositol anchor biosynthesis.	2.05
AW440350	Unknown function	2.05

THC2265157	Unknown function	2.05
H43551	Unknown function	2.04
BM504117	Unknown function	2.04
AF159295	Unknown function	2.04
MX1	Myxovirus resistance 1, interforn-inducible protein p78; encoded protein is responsible for specific antiviral state against influenza virus infection.	2.04
FGFBP1	Fibroblast growth factor binding protein 1; Encodes secreted fibroblast growth factor carrier protein that plays major role in cell differentiation, proliferation, and migration. FGFBP1 has been linked to cancer, including pancreatic and colorectal adnenocarcinoma.	2.03
MIF	Macrophage migration inhibitory factor (glycosylation-inhibiting factor; encodes lymphokine involved in cell-mediated immunity, immunoregulation, and inflammation.	2.01
ASB4	Ankyrin repeat and SOCS box containing 4; encoded protein is member of ankyrin repeat and SOCS box-containing (ASB) family of proteins.	2.01
RNASE4	Ribonuclease, RNase A family, 4; encoded protein is member of pancreatic ribonuclease family. Plays important role in mRNA cleavage and has marked specificity towards the 3' side of uridine nucleotides.	2.01
ATP8B3	ATPase, aminophospholipd transporter, class I, type 8B, member 3; encoded protein belongs to family of P-type cation transport ATPases, and to the subfamily of aminophospholipid-transporting ATPases.	2.00
MAFA	v-maf musculoaponeurotic oncogene (avaian); MAFA is a transcription factor that binds RIPE3b, a conserved enhancer element that regulates pancreatic beta cel-specific expression of the insulin gene.	-2.02
PRLH	Prolactin releasing hormone	-2.02
A_24_P375751	Unknown function	-2.05
PTPRO	Protein tyrosine phosphatase, receptor type, O; encodes member of R3 subtype family of receptor-type protein tyrosine phosphatases. This gene is believed to be a tumor suppressor, and its decreased expression has been associated with several types of cancer.	-2.07
PDGFRA	Platelet-derived growth factor receptor, alpha polypeptide; encoded protein is a cell surface tyrosine	-2.10

	kinase receptor for members of the platelet-derived growth factor family. These growth factors are mitogens for mysenchymal clles. Mutations in this gene have been associated with tumor progression.	
TNRC18	Trinucleotide repeat containing 18	-2.11
BC013423	Unknown function	-2.17
BQ015140	Unknown function	-2.17
DHH	Desert hedgehog; this gene encodes member of hedgehog family that encodes signaling molecules. The protein may be involved in male gonadal differentiation and perineurial development.	-2.17
A_24_P862251	Unknown function	-2.18
HAUS1	HAUS augmin-like complex, subunit 1; encoded protein is a microtubule-binding complex involved in microtubule generation within the mitotic spindle.	-2.18
HOX-AS3	HOXB cluster antisense RNA3	-2.24
GPHB5	Glycoprotein hormone beta 5; encodes cysteine knot-forming polypeptide and a subunit of the dimeric glycoprotein hormone family.	-2.29
SLC35F5	Solute carrier family 35, member F5	-2.31
ENST00000074056	Unknown function	-2.32
NM_173573	Unknown function	-2.34
ARTN	Artemin; this protein is a member of the glial cell line-derived neurotrophic (GDNF) family of ligands and has neurotrophic properties.	-2.37
STK11	Serine/threonine kinase 11; regulates cell polarity and functions as a tumor suppressor. Mutations in the gene associated Peutz-Jeghers syndrome and other neoplasms.	-2.38
LRRC15	Leucine rich repeat containing 15	-2.40
ARHGAP27	Rho GTPase activating protein 27; encodes a GTPases that is involved in many cellular processes.	-2.44
CLDN5	Claudin 5; encodes integral membrane protein and forms tight junctions. Mutations in this gene have been associated with velocardiofacial syndrome.	-2.47
C1QTNF5	C1q and tumor necrosis factor related protein 5; encodes protein that may be a component of basement	-2.47

	membranes and may play role in cell adhesion. Mutations in this gene have been associated with late-onset retinal degeneration	
BX105253	Unknown function	-2.65
MATN3	Matrilin 3; encodes member of von Willebrand factor A domain containing family of proteins. Thought to be involved in forming filamentous networks in the extracellular matrices in several tissues. Mutations in this gene result in multiple epiphyseal dysplasia.	-2.69
GLTP	Glycolipid transfer protein; encoded protein accelerate transfer of various glycosphingolipids and glyceroglycolipids between membranes.	-2.79
A_32_P185089	Unknown function	-2.80
EFCAB4A	EF-hand calcium binding domain 4a	-3.00
CBX5	Chromobox homolog; encodes member of heterochromatin protein family. Enriched in the heterochromatin and associated with centromeres. Can bind to histones via methylated lysine residues and is involved in the formation of functional kinetochore.	-3.36
AVP	Arginine vasopressin. This encoded proteins is a posterior pituitary hormone that acts as a growth factor and is involved in antidiuretic activity in the kidney. It can also cause vasoconstriction of the peripheral vessels.	-3.62

[a] Gene function description was summarized from information obtained from database Entrez Gene Summary.

[b] All fold changes are statistically significant (Q<0.05; False Discovery Rate (FDR) <0.05).

Supplementary Table S2. Identified cancer-related differentially expressed miR-618

targets

Gene name	Expression fold changes[a]	Cancer-related functions[b]
ANG	2.25	Tumorigenesis of cervical cancer cell lines, mammary tumor
ANXA8L2	2.19	Carcinoma, adenocarcinoma, ovarian cancer; gonadal tumor
BIK	2.66	Carcinoma, adenocarcinoma, uterine cancer
C8ORF8/SGK1	2.17	Mammary tumor skin cancer
CDH7	2.26	Skin cancer
CLDN5	-2.47	Carcinoma, adenocarcinoma, uterine cancer
DCN	2.35	Mammary tumor, arrest in G1 phase of colon carcinoma cells
DESI2	2.22	Mammary tumor, carcinoma
FGFBP1	2.03	Epidermal hyperplasia
FOSB	3.62	Mammary tumor, carcinoma, ovarian cancer, gonadal cancer
HMGB1	2.04	Carcinoma, liver cancer, large cell lymphoma, Non-Hodgkin's lymphoma, Hodgkin's lymphoma
HUWE1	2.04	Mammary tumor, carcinoma, head and neck cancer, liver, cancer, benign neoplasia, pancreatic tumor, lung cancer, pancreatic cancer, carcinoma in the lung, thyroid carcinoma
IFI27	2.14	Carcinoma, ovarian cancer, gonadal cancer, head and neck cancer, skin squamous cell carcinoma
IGFBP5	-2.02	Carcinoma, adenocarcinoma, head and neck cancer, benign neoplasia, pancreatic tumor, binding of mammary tumor cells, leiomyomatosis, mucoepidermoid carcinoma, metastatic colorectal cancer
IL22RA1	2.52	Carcinoma
JUN	2.79	Mammary tumor, carcinoma, ovarian cancer, gonadal tumor, uterine cancer, skin cancer, squamous cell tumor, large-cell lymphoma, Non-Hodgkin's lymphoma, non-Hodgkin's lymphoma, head and neck cancer, benign neoplasia, lung cancer, skin squamous cell carcinoma, leiomyomatosis,

KCNE4	2.75	Leukemia, bone marrow cancer
KIT	2.20	Mammary tumor, carcinoma, adenocarcinoma, ovarian cancer, gonadal cancer, skin cancer, liver cancer, large-cell lymphoma, Non-Hodgkin's lymphoma, head and neck cancer, benign neoplasia, pancreatic tumor, lung cancer, carcinoma in lung, thyroid carcinoma, leiomyomatosis, leukemia, bone barrow cancer, dysgerminoma, uticaria pigmentosa, chordoma, aggressive fibromatosis, mixed germ cell tumor, seminoma, melanoma, sporadic mastocytosis, sarcoma, systemic mastocytosis, islet-cell carcinoma, neurofibrosarcoma, primary peritoneal cancer, gastrointestinal stromal tumor, Fallopian tube cancer, testicular cancer
KRT18	2.99	Carcinoma, adenocarcinoma, uterine cancer, head and neck cancer, mucoepidermoid carcinoma, metastatic colorectal cancer
MAFA	-2.02	Carcinoma, squamous cell tumor
MATN3	-2.70	Carcinoma, liver cancer
MIF	2.01	Carcinoma, ovarian cancer, gonadal cancer, sarcoma
MX1	2.04	Epidermal hyperplasia
PDGFRA/B	-2.10 and +2.04, respectively	Mammary tumor, carcinoma, adenocarcinoma, ovarian cancer, gonadal cancer, liver cancer, head and neck cancer, benign neoplasia, pancreatic tumor, lung cancer, pancreatic cancer, carcinoma in lung, leiomyomatosis, leukemia, bone marrow cancer, sarcoma, systemic mastocytosis, islet-cell carcinoma, neurofibrosarcoma, primary peritoneal cancer, gastrointestinal stromal tumor, Fallopian tube cancer
RNASE4	2.01	Carcinoma, adenocarcinoma, uterine cacner
STK11/LKB1	-2.38	Carcinoma, adenocarcinoma, gonadal tumor, uterine cancer, skin cancer, squamous cell tumor, pancreatic tumor, lung cancer, pancreatic cancer, carcinoma in lung, testicular cancer, adenoma malignum, Peutz Jagher syndrome
TM4SF1	3.16	Carcinoma, adenocarcinoma, skin cancer, liver cancer, metastatic colorectal cancer

[a] All fold changes have a Q value of <0.05.

[b] All associations between the transcript and disease have a p-value of <0.05.

Supplementary Table S2: Affected genes related to hepatotoxicity and

hepatocellular carcinoma

Gene	Effect	Fold change[a]
TM4SF1	Upregulated	3.16
KIT	Upregulated	2.20
PDGFRB	Upregulated	2.14
HUWE1	Upregulated	2.04
HMGB1	Upregulated	2.04
PDGFRA	Downregulated	-2.10
MATN3	Downregulated	-2.70

[a] Overlap p-value: 5.72×10^{-3}

Supplementary Figure S1: Secondary structure prediction of pre-miR-618.
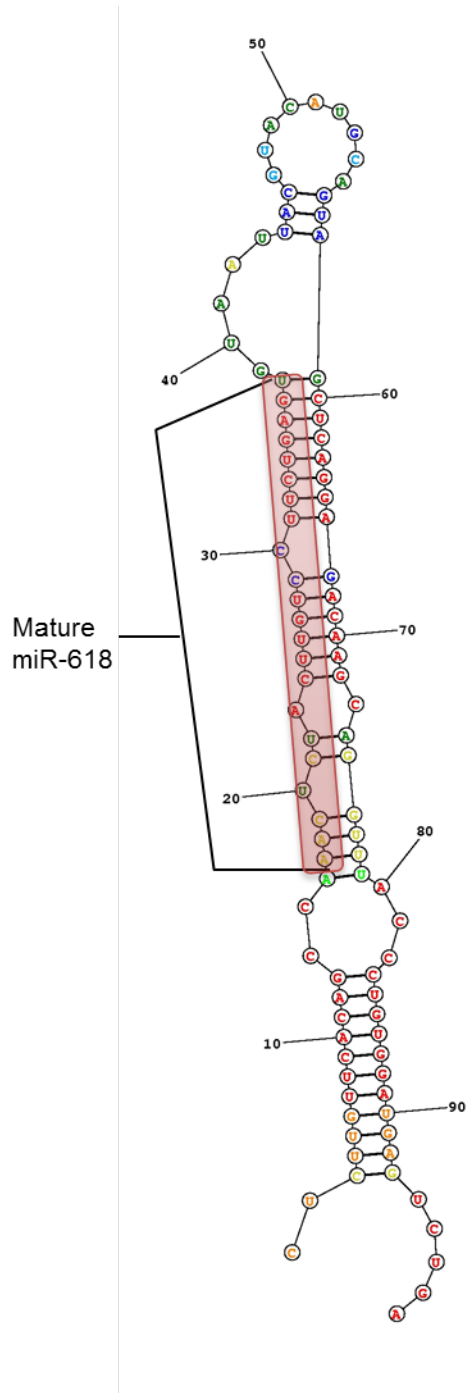
Supplementary Figure S1: Predicted secondary structure of pre-miR-618. Mature miR-618 is highlighted in red. Base pairs are colored according to the probability of predicted accuracy. Red represents >99% probability, whereas blue represents 50-60% accuracty. Secondary structure generated using RNAstructure.[101]

Supplementary Figure S2: Differentially expressed transcripts following miR-618 transfection in the "Cellular movement, cell death and survival, and cell cycle" network.
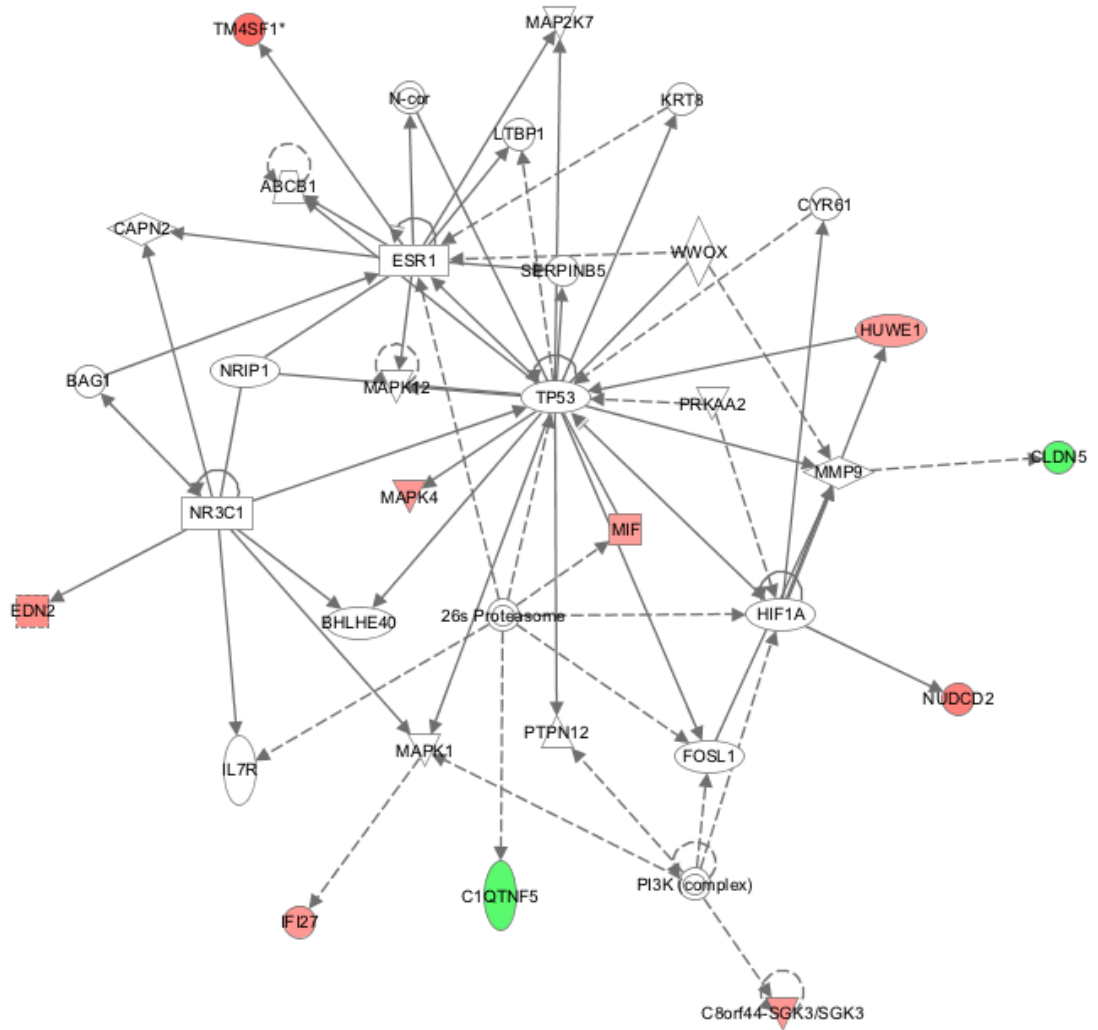
Supplementary Figure S2: Differentially expressed miR-618 targets in the "Cellular movement, cell death and survival, and cell cycle" network. This network was identified by the Ingenuity Pathway Analysis software as significantly associated with the set of transcripts with altered expression following the introduction of pre-miR-618 (pooled p-value = 8.61 x 10^{-7} and FC ≥ |2|). Overexpressed transcripts are indicated in red, and underexpressed transcripts are indicated in green.

CHAPTER 8: REFERENCES

1   Kala, R., Peek, G. W., Hardy, T. M. & Tollefsbol, T. O. MicroRNAs: an emerging science in cancer epigenetics. *Journal of clinical bioinformatics* **3**, 6, doi:10.1186/2043-9113-3-6 (2013).

2   Lim, L. P. *et al.* Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* **433**, 769-773, doi:10.1038/nature03315 (2005).

3   Esquela-Kerscher, A. & Slack, F. J. Oncomirs - microRNAs with a role in cancer. *Nature reviews. Cancer* **6**, 259-269, doi:10.1038/nrc1840 (2006).

4   Calin, G. A. *et al.* Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proc Natl Acad Sci U S A* **99**, 15524-15529, doi:10.1073/pnas.242606799 (2002).

5   Iorio, M. V. *et al.* MicroRNA gene expression deregulation in human breast cancer. *Cancer Res* **65**, 7065-7070, doi:10.1158/0008-5472.can-05-1783 (2005).

6   O'Day, E. & Lal, A. MicroRNAs and their target gene networks in breast cancer. *Breast cancer research : BCR* **12**, 201, doi:10.1186/bcr2484 (2010).

7   Takamizawa, J. *et al.* Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival. *Cancer Res* **64**, 3753-3756, doi:10.1158/0008-5472.can-04-0637 (2004).

8   Lee, R. C., Feinbaum, R. L. & Ambros, V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **75**, 843-854 (1993).

9   Calin, G. A. & Croce, C. M. MicroRNA signatures in human cancers. *Nature reviews. Cancer* **6**, 857-866, doi:10.1038/nrc1997 (2006).

10   Esteller, M. Non-coding RNAs in human disease. *Nature reviews. Genetics* **12**, 861-874, doi:10.1038/nrg3074 (2011).

11   Calin, G. A. *et al.* Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proc Natl Acad Sci U S A* **101**, 2999-3004, doi:10.1073/pnas.0307323101 (2004).

12   He, L. *et al.* A microRNA polycistron as a potential human oncogene. *Nature* **435**, 828-833, doi:10.1038/nature03552 (2005).

13   Lu, J. *et al.* MicroRNA expression profiles classify human cancers. *Nature* **435**, 834-838, doi:10.1038/nature03702 (2005).

14   Moore, A. E., Chenette, D. M., Larkin, L. C. & Schneider, R. J. Physiological networks and disease functions of RNA-binding protein AUF1. *Wiley interdisciplinary reviews. RNA*, doi:10.1002/wrna.1230 (2014).

15   Zhang, M. *et al.* Associations of miRNA polymorphisms and female physiological characteristics with breast cancer risk in Chinese population. *European Journal of Cancer Care* **21**, 274-280, doi:10.1111/j.1365-2354.2011.01308.x (2012).

16   Hoffman, A. E. *et al.* microRNA miR-196a-2 and breast cancer: a genetic and epigenetic association study and functional analysis. *Cancer Res* **69**, 5970-5977, doi:10.1158/0008-5472.CAN-09-0236

0008-5472.CAN-09-0236 [pii] (2009).

17   Pennisi, E. Behind the scenes of gene expression. *Science* **293**, 1064-1067 (2001).

18   Jones, P. A. & Takai, D. The role of DNA methylation in mammalian epigenetics. *Science* **293**, 1068-1070 (2001).

19   Jenuwein, T. & Allis, C. D. Translating the histone code. *Science* **293**, 1074-1080 (2001).

20   Jaenisch, R. & Bird, A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet* **33 Suppl**, 245-254 (2003).

21   Ng, H. H. & Bird, A. DNA methylation and chromatin modification. *Curr Opin Genet Dev* **9**, 158-163 (1999).

22   Prendergast, G. C. & Ziff, E. B. Methylation-sensitive sequence-specific DNA binding by the c-Myc basic region. *Science* **251**, 186-189 (1991).

23   Krude, T. & Keller, C. Chromatin assembly during S phase: contributions from histone deposition, DNA replication and the cell division cycle. *Cell Mol Life Sci* **58**, 665-672 (2001).

24   Holliday, R. & Pugh, J. E. DNA modification mechanisms and gene activity during development. *Science* **187**, 226-232 (1975).

25   Naveh-Many, T. & Cedar, H. Active gene sequences are undermethylated. *Proc Natl Acad Sci U S A* **78**, 4246-4250 (1981).

26   Rakyan, V. K., Preis, J., Morgan, H. D. & Whitelaw, E. The marks, mechanisms and memory of epigenetic states in mammals. *Biochem J* **356**, 1-10 (2001).

27   Li, E., Beard, C. & Jaenisch, R. Role for DNA methylation in genomic imprinting. *Nature* **366**, 362-365 (1993).

28   Kanduri, C. & Raman, R. Characterisation of developmentally regulated chromatin structure in the coding region of the proto-oncogene, c-fos, in the male laboratory mouse. *Int J Dev Biol* **43**, 279-282 (1999).

29   Simon, D. *et al.* Retrovirus genomes methylated by mammalian but not bacterial methylase are non-infectious. *Nature* **304**, 275-277 (1983).

30   Petronis, A. Human morbid genetics revisited: relevance of epigenetics. *Trends Genet* **17**, 142-146 (2001).

31   Muller-McNicoll, M. & Neugebauer, K. M. How cells get the message: dynamic assembly and function of mRNA-protein complexes. *Nature reviews. Genetics* **14**, 275-287, doi:10.1038/nrg3434 (2013).

32   Kim, M. Y., Hur, J. & Jeong, S. Emerging roles of RNA and RNA-binding protein network in cancer cells. *BMB reports* **42**, 125-130 (2009).

33   van Kouwenhove, M., Kedde, M. & Agami, R. MicroRNA regulation by RNA-binding proteins and its implications for cancer. *Nature reviews. Cancer* **11**, 644-656, doi:10.1038/nrc3107 (2011).

34   Re, A., Joshi, T., Kulberkyte, E., Morris, Q. & Workman, C. T. RNA-Protein Interactions: An Overview. *Methods in molecular biology (Clifton, N.J.)* **1097**, 491-521, doi:10.1007/978-1-62703-709-9_23 (2014).

35   Baltz, A. G. *et al.* The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. *Molecular cell* **46**, 674-690, doi:10.1016/j.molcel.2012.05.021 (2012).

36   Konig, J. *et al.* iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nature structural & molecular biology* **17**, 909-915, doi:10.1038/nsmb.1838 (2010).

37   Zhang, C., Lee, K. Y., Swanson, M. S. & Darnell, R. B. Prediction of clustered RNA-binding protein motif sites in the mammalian genome. *Nucleic Acids Res* **41**, 6793-6807, doi:10.1093/nar/gkt421 (2013).

38   Lunde, B. M., Moore, C. & Varani, G. RNA-binding proteins: modular design for efficient function. *Nature reviews. Molecular cell biology* **8**, 479-490, doi:10.1038/nrm2178 (2007).

39   Gibson, T. J., Thompson, J. D. & Heringa, J. The KH domain occurs in a diverse set of RNA-binding proteins that include the antiterminator NusA and is probably involved in binding to nucleic acid. *FEBS letters* **324**, 361-366 (1993).

40   Wen, W. L. *et al.* Vgl1, a multi-KH domain protein, is a novel component of the fission yeast stress granules required for cell survival under thermal stress. *Nucleic Acids Res* **38**, 6555-6566, doi:10.1093/nar/gkq555 (2010).

41   Gutschner, T. *et al.* The Insulin-like growth factor 2 mRNA-binding protein 1 (IGF2BP1) is an important pro-tumorigenic factor in hepatocellular carcinoma. *Hepatology (Baltimore, Md.)*, doi:10.1002/hep.26997 (2014).

42   Yisraeli, J. K. VICKZ proteins: a multi-talented family of regulatory RNA-binding proteins. *Biology of the cell / under the auspices of the European Cell Biology Organization* **97**, 87-96, doi:10.1042/bc20040151 (2005).

43   Bell, J. L. *et al.* Insulin-like growth factor 2 mRNA-binding proteins (IGF2BPs): post-transcriptional drivers of cancer progression? *Cell Mol Life Sci* **70**, 2657-2675, doi:10.1007/s00018-012-1186-z (2013).

44   Wachter, K., Kohn, M., Stohr, N. & Huttelmaier, S. Subcellular localization and RNP formation of IGF2BPs (IGF2 mRNA-binding proteins) is modulated by distinct RNA-binding domains. *Biological chemistry* **394**, 1077-1090, doi:10.1515/hsz-2013-0111 (2013).

45   Pereira, B., Le Borgne, M., Chartier, N. T., Billaud, M. & Almeida, R. MEX-3 proteins: recent insights on novel post-transcriptional regulators. *Trends in biochemical sciences* **38**, 477-479, doi:10.1016/j.tibs.2013.08.004 (2013).

46   Draper, B. W., Mello, C. C., Bowerman, B., Hardin, J. & Priess, J. R. MEX-3 is a KH domain protein that regulates blastomere identity in early C. elegans embryos. *Cell* **87**, 205-216 (1996).

47   Buchet-Poyau, K. *et al.* Identification and characterization of human Mex-3 proteins, a novel family of evolutionarily conserved RNA-binding proteins differentially localized to processing bodies. *Nucleic Acids Research* **35**, 1289-1300, doi:10.1093/nar/gkm016 (2007).

48   Huang, N. N., Mootz, D. E., Walhout, A. J., Vidal, M. & Hunter, C. P. MEX-3 interacting proteins link cell polarity to asymmetric gene expression in Caenorhabditis elegans. *Development (Cambridge, England)* **129**, 747-759 (2002).

49   Hunter, C. P. & Kenyon, C. Spatial and temporal controls target pal-1 blastomere-specification activity to a single blastomere lineage in C. elegans embryos. *Cell* **87**, 217-226 (1996).

50   Buchet-Poyau, K. *et al.* Identification and characterization of human Mex-3 proteins, a novel family of evolutionarily conserved RNA-binding proteins differentially localized to processing bodies. *Nucleic Acids Res* **35**, 1289-1300, doi:10.1093/nar/gkm016 (2007).

51   Cano, F. *et al.* The RNA-binding E3 ubiquitin ligase MEX-3C links ubiquitination with MHC-I mRNA degradation. *The EMBO journal* **31**, 3596-3606, doi:10.1038/emboj.2012.218 (2012).

52   Guzmán, B. *et al.* Implication of Chromosome 18 in Hypertension by Sibling Pair and Association Analyses: Putative Involvement of the RKHD2 Gene. *Hypertension* **48**, 883-891, doi:10.1161/01.HYP.0000244085.52918.a0 (2006).

53   Jiao, Y. *et al.* Mex3c mutation reduces adiposity and increases energy expenditure. *Molecular and cellular biology* **32**, 4350-4362, doi:10.1128/mcb.00452-12 (2012).

54   Jiang, H. *et al.* Knockdown of hMex-3A by small RNA interference suppresses cell proliferation and migration in human gastric cancer cells. *Molecular medicine reports* **6**, 575-580, doi:10.3892/mmr.2012.943 (2012).

55   Pereira, B. *et al.* CDX2 regulation by the RNA-binding protein MEX3A: impact on intestinal differentiation and stemness. *Nucleic Acids Research* **41**, 3986-3999, doi:10.1093/nar/gkt087 (2013).

56   Kalchhauser, I., Farley, B. M., Pauli, S., Ryder, S. P. & Ciosk, R. FBF represses the Cip/Kip cell-cycle inhibitor CKI-2 to promote self-renewal of germline stem cells in C. elegans. *The EMBO journal* **30**, 3823-3829, doi:10.1038/emboj.2011.263 (2011).

57   Porter, K. R., McCarthy, B. J., Freels, S., Kim, Y. & Davis, F. G. Prevalence estimates for primary brain tumors in the United States by age, gender, behavior, and histology. *Neuro-oncology* **12**, 520-527, doi:10.1093/neuonc/nop066 (2010).

58   Louis DN, O. H., Wiestler OD, Cavenee WK. WHO Classification of Tumours of the Central Nervous System. *Lyon, France: International Agency for Research on Cancer* **4th ed.** (2007).

59   Omuro, A. & DeAngelis, L. M. Glioblastoma and other malignant gliomas: a clinical review. *JAMA : the journal of the American Medical Association* **310**, 1842-1850, doi:10.1001/jama.2013.280319 (2013).

60   Dolecek, T. A., Propp, J. M., Stroup, N. E. & Kruchko, C. CBTRUS statistical report: primary brain and central nervous system tumors diagnosed in the United States in 2005-2009. *Neuro-oncology* **14 Suppl 5**, v1-49, doi:10.1093/neuonc/nos218 (2012).

61   Corle, C., Makale, M. & Kesari, S. Cell phones and glioma risk: a review of the evidence. *Journal of neuro-oncology* **106**, 1-13, doi:10.1007/s11060-011-0663-9 (2012).

62   Hardell, L., Carlberg, M. & Hansson Mild, K. Use of mobile phones and cordless phones is associated with increased risk for glioma and acoustic

neuroma. *Pathophysiology : the official journal of the International Society for Pathophysiology / ISP* **20**, 85-110, doi:10.1016/j.pathophys.2012.11.001 (2013).

63   Shete, S. *et al.* Genome-wide association study identifies five susceptibility loci for glioma. *Nat Genet* **41**, 899-904, doi:10.1038/ng.407 (2009).

64   Rajaraman, P. *et al.* Genome-wide association study of glioma and meta-analysis. *Human genetics* **131**, 1877-1888, doi:10.1007/s00439-012-1212-0 (2012).

65   Hottinger, A. F. & Khakoo, Y. Update on the management of familial central nervous system tumor syndromes. *Current neurology and neuroscience reports* **7**, 200-207 (2007).

66   Gu, J., Liu, Y., Kyritsis, A. P. & Bondy, M. L. Molecular epidemiology of primary brain tumors. *Neurotherapeutics : the journal of the American Society for Experimental NeuroTherapeutics* **6**, 427-435, doi:10.1016/j.nurt.2009.05.001 (2009).

67   Mignone, F., Gissi, C., Liuni, S. & Pesole, G. Untranslated regions of mRNAs. *Genome biology* **3**, Reviews0004 (2002).

68   Lewis, H. A. *et al.* Sequence-specific RNA binding by a Nova KH domain: implications for paraneoplastic disease and the fragile X syndrome. *Cell* **100**, 323-332 (2000).

69   Demichelis, F. *et al.* Identification of functionally active, low frequency copy number variants at 15q21.3 and 12q21.31 associated with prostate cancer risk.

Proc Natl Acad Sci U S A **109**, 6686-6691, doi:10.1073/pnas.1117405109 (2012).

70   Fu, A. *et al.* Targetome profiling and functional genetics implicate miR-618 in lymphomagenesis. *Epigenetics : official journal of the DNA Methylation Society* **9** (2014).

71   Milde-Langosch, K. *et al.* The role of the AP-1 transcription factors c-Fos, FosB, Fra-1 and Fra-2 in the invasion process of mammary carcinomas. *Breast Cancer Res Treat* **86**, 139-152, doi:10.1023/b:brea.0000032982.49024.71 (2004).

72   van Dam, H. & Castellazzi, M. Distinct roles of Jun : Fos and Jun : ATF dimers in oncogenesis. *Oncogene* **20**, 2453-2464, doi:10.1038/sj.onc.1204239 (2001).

73   Milde-Langosch, K. The Fos family of transcription factors and their role in tumourigenesis. *Eur J Cancer* **41**, 2449-2461, doi:10.1016/j.ejca.2005.08.008 (2005).

74   Liang, J. *et al.* The C-kit receptor-mediated signal transduction and tumor-related diseases. *International journal of biological sciences* **9**, 435-443, doi:10.7150/ijbs.6087 (2013).

75   Antonescu, C. R. The GIST paradigm: lessons for other kinase-driven cancers. *The Journal of pathology* **223**, 251-261, doi:10.1002/path.2798 (2011).

76   Motiwala, T. *et al.* Protein tyrosine phosphatase receptor-type O (PTPRO) exhibits characteristics of a candidate tumor suppressor in human lung cancer.

Proc Natl Acad Sci U S A **101**, 13844-13849, doi:10.1073/pnas.0405451101 (2004).

77   Jones, J. I. & Clemmons, D. R. Insulin-like growth factors and their binding proteins: biological actions. *Endocrine reviews* **16**, 3-34 (1995).

78   Gullu, G., Karabulut, S. & Akkiprik, M. Functional roles and clinical values of insulin-like growth factor-binding protein-5 in different types of cancers. *Chinese journal of cancer* **31**, 266-280, doi:10.5732/cjc.011.10405 (2012).

79   Ollila, S. & Makela, T. P. The tumor suppressor kinase LKB1: lessons from mouse models. *Journal of molecular cell biology* **3**, 330-340, doi:10.1093/jmcb/mjr016 (2011).

80   Abdalla, M. A. & Haj-Ahmad, Y. Promising Candidate Urinary MicroRNA Biomarkers for the Early Detection of Hepatocellular Carcinoma among High-Risk Hepatitis C Virus Egyptian Patients. *Journal of Cancer* **3**, 19-31 (2012).

81   Hou, J. *et al.* Estrogen-sensitive PTPRO expression represses hepatocellular carcinoma progression by control of STAT3. *Hepatology (Baltimore, Md.)* **57**, 678-688, doi:10.1002/hep.25980 (2013).

82   Nagata, C., Kawakami, N. & Shimizu, H. Trends in the incidence rate and risk factors for breast cancer in Japan. *Breast Cancer Res Treat.* **44**, 75-82. (1997).

83   Cicatiello, L. *et al.* Estrogen receptor alpha controls a gene network in luminal-like breast cancer cells comprising multiple transcription factors and microRNAs. *The American journal of pathology* **176**, 2113-2130, doi:10.2353/ajpath.2010.090837 (2010).

84   D'Errico, M. *et al.* Genome-wide expression profile of sporadic gastric cancers with microsatellite instability. *Eur J Cancer* **45**, 461-469, doi:10.1016/j.ejca.2008.10.032 (2009).

85   Hong, Y., Downey, T., Eu, K. W., Koh, P. K. & Cheah, P. Y. A 'metastasis-prone' signature for early-stage mismatch-repair proficient sporadic colorectal cancer patients and its implications for possible therapeutics. *Clinical & experimental metastasis* **27**, 83-90, doi:10.1007/s10585-010-9305-4 (2010).

86   Pei, H. *et al.* FKBP51 affects cancer cell response to chemotherapy by negatively regulating Akt. *Cancer cell* **16**, 259-266, doi:10.1016/j.ccr.2009.07.016 (2009).

87   Sabates-Bellver, J. *et al.* Transcriptome profile of human colorectal adenomas. *Molecular cancer research : MCR* **5**, 1263-1275, doi:10.1158/1541-7786.mcr-07-0267 (2007).

88   Talantov, D. *et al.* Novel genes associated with malignant melanoma but not benign melanocytic lesions. *Clinical cancer research : an official journal of the American Association for Cancer Research* **11**, 7234-7242, doi:10.1158/1078-0432.ccr-05-0683 (2005).

89   Wallace, T. A. *et al.* Tumor immunobiological differences in prostate cancer between African-American and European-American men. *Cancer Res* **68**, 927-936, doi:10.1158/0008-5472.can-07-2608 (2008).

90   Barretina, J. *et al.* Subtype-specific genomic alterations define new targets for soft-tissue sarcoma therapy. *Nat Genet* **42**, 715-721, doi:10.1038/ng.619 (2010).

91   Detwiller, K. Y. *et al.* Analysis of hypoxia-related gene expression in sarcomas and effect of hypoxia on RNA interference of vascular endothelial cell growth factor A. *Cancer Res* **65**, 5881-5889, doi:10.1158/0008-5472.can-04-4078 (2005).

92   Pyeon, D. *et al.* Fundamental differences in cell cycle deregulation in human papillomavirus-positive and human papillomavirus-negative head/neck and cervical cancers. *Cancer Res* **67**, 4605-4619, doi:10.1158/0008-5472.can-06-3619 (2007).

93   Compagno, M. *et al.* Mutations of multiple genes cause deregulation of NF-kappaB in diffuse large B-cell lymphoma. *Nature* **459**, 717-721, doi:10.1038/nature07968 (2009).

94   Storm, H. H. The Danish Cancer Registry, a self-reporting national cancer registration system with elements of active data collection. *IARC Sci Publ*, 220-236 (1991).

95   Liu, C., Xia, Y., Jiang, W., Liu, Y. & Yu, L. Low expression of GABARAPL1 is associated with a poor outcome for patients with hepatocellular carcinoma. *Oncology reports*, doi:10.3892/or.2014.3096 (2014).

96   Ciafre, S. A. & Galardi, S. microRNAs and RNA-binding proteins: a complex network of interactions and reciprocal regulations in cancer. *RNA biology* **10**, 935-942, doi:10.4161/rna.24641 (2013).

97   Gubin, M. M. *et al.* Overexpression of the RNA binding protein HuR impairs tumor growth in triple negative breast cancer associated with deficient

angiogenesis. *Cell cycle (Georgetown, Tex.)* **9**, 3337-3346, doi:10.4161/cc.9.16.12711 (2010).

98   Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **57**, 289-300 (1995).

99   Calvano, S. E. *et al.* A network-based analysis of systemic inflammation in humans. *Nature* **437**, 1032-1037, doi:nature03985 [pii]

10.1038/nature03985 (2005).

100   Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *J Roy Stat Soc B Met* **57**, 289-300 (1995).

101   Reuter, J. S. M., D.H. RNAstructure: software for RNA secondary structure prediciton and analysis. *BMC Bioinformatics* **11** (2010).