FGF SIGNALING DURING GASTRULATION AND CARDIOGENESIS

By

Alexander Sebastian Bobbs

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF MOLECULAR AND CELLULAR BIOLOGY

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2012
As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Alexander S. Bobbs entitled "FGF Signaling during Gastrulation and Cardiogenesis" and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

Date: 11/6/12
Parker Antin

Date: 11/6/12
Ted Weinert

Date: 11/6/12
Ray Runyan

Date: 11/6/12
Lisa Nagy

Final approval and acceptance of this dissertation is contingent upon the candidate’s submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Date: 11/6/12
Dissertation Director: Parker Antin
STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED:  
Alexander Sebastian Bobbs
ACKNOWLEDGEMENTS

I must first thank my mentor, Dr. Parker Antin. Parker has been an invaluable help and a tremendously patient mentor throughout my projects in his lab. I would like to thank him for all that he's taught me, and all the opportunities he offered me. In addition, I acknowledge my committee members past and present, including Lisa Nagy, Ted Weinert, Ray Runyan, Brad Davidson, and Roy Parker. I would also like to thank my collaborators in the Antin Lab, particularly our lab manager Tatiana Yatskievych, who prepared most of the embryos used in my experiments. In addition, I received help with the molecular biology work from my assistant Aleksi Saarela and post-doctoral fellow James Cooley. Members of the GEISHA project, including Maricela Pier, Stella Zhang, and Terry Seseparasa, conducted in situ hybridizations that supplemented my project. I also received considerable instruction in bioinformatics and programming from Jay Konieska and Nirav Merchant.

Besides those who helped me on this dissertation project, I would like to acknowledge some previous mentors who helped my development as a molecular biologist: Linda Iverson and Sailesh Gopalakrishna-Pillai at City of Hope, Gene Gerner and Ning Qu at the University of Arizona, and Nancy Hamlett at Harvey Mudd College.

Lastly, I would like to acknowledge my family for their support and patience throughout the near-decade I spent in and out of grad school, and the late Michael Chrichton for his books that got me interested in molecular biology as a child.
TABLE OF CONTENTS

LIST OF FIGURES ......................................................................................................................... 7
LIST OF TABLES .............................................................................................................................. 8
ABSTRACT ......................................................................................................................................... 9

CHAPTER ONE: INTRODUCTION ....................................................................................................... 11
1.1: GASTRULATION AND CARDIOGENESIS IN THE AVIAN EMBRYO .................. 11
   1.1.1: Overview of early embryogenesis in Gallus gallus ........................................... 11
   1.1.2: Gastrulation in Gallus gallus .............................................................................. 14
   1.1.3: FGF signaling ......................................................................................................... 18
   1.1.4: Cardiac myogenesis .............................................................................................. 19
1.2: MICRORNAS IN EMBRYONIC DEVELOPMENT ...................................................... 23
   1.2.1: MicroRNA formation and function ..................................................................... 24
   1.2.2: The let-7 family of microRNAs ......................................................................... 27
   1.2.3: Summary of approach ......................................................................................... 29
1.3: COMBINATORIAL INPUTS OF BMP AND FGF SIGNALING IN CARDIAC 
   MYOGENESIS ......................................................................................................................... 32
   1.3.1: FGF and BMP regulatory bins ............................................................................. 32
   1.3.2: Computational Predictions of Transcription Factor Binding Motifs ................. 37
   1.3.3: Summary of approach ......................................................................................... 39
1.4: EXPLANATION OF DISSERTATION FORMAT ...................................................... 42
1.5: SUMMARY ......................................................................................................................... 44

CHAPTER TWO: FIBROBLAST GROWTH FACTOR (FGF) AND MICRORNAS IN 
AVIAN GASTRULATION ............................................................................................................. 46
2.1: PRESENT STUDY ............................................................................................................... 46
2.2: FUTURE DIRECTIONS ....................................................................................................... 49
CHAPTER THREE: COMPUTATIONAL AND EXPERIMENTAL ANALYSIS OF PROMOTER REGIONS CHARACTERIZING REGULATORY "BINS" IN FIBROBLAST GROWTH FACTOR (FGF) AND BONE MORPHOGENETIC PROTEIN (BMP) SIGNALING PATHWAYS.

.................................................................................................................................51

3.1: ABSTRACT ....................................................................................................................51

3.2: INTRODUCTION .........................................................................................................52

3.3: MATERIALS AND METHODS ..................................................................................54

3.3.1: Assembling promoterome files ..............................................................................54

3.3.2: Identifying Over-represented Motifs ......................................................................55

3.3.3: Condensing and Clustering Motifs .........................................................................56

3.3.4: Primary Cell Culture ...............................................................................................57

3.3.5: Transfections and Luciferase Assays ......................................................................59

3.4: RESULTS ......................................................................................................................61

3.4.1: Over-Represented Motifs .......................................................................................61

3.4.2: Clusters of motifs .....................................................................................................63

3.4.3: Testing enhancer elements .....................................................................................67

3.4.4: Analysis of components in functional clusters .......................................................71

3.4.5: Enhancer response to FGF signaling ....................................................................75

3.5: CONCLUSION ..............................................................................................................80

3.6: FUTURE DIRECTIONS ...............................................................................................83

CHAPTER FOUR: CONCLUDING THOUGHTS ....................................................................85

APPENDIX A: FIBROBLAST GROWTH FACTOR (FGF) SIGNALING NEGATIVELY REGULATES THE ABUNDANCE OF MICRORNAS THAT REGULATE PROTEINS REQUIRED FOR CELL MIGRATION AND EMBRYO PATTERNING DURING GASTRULATION ...................................................................................88

REFERENCES ..................................................................................................................127
LIST OF FIGURES

Figure 1: Primitive Streak Formation and Gastrulation in Early Chick Development......13
Figure 2: FGF Signaling is required for cell migration ..................................17
Figure 3: Location of Heart Precursor Cells..................................................22
Figure 4: Pathway for microRNA biogenesis.............................................25
Figure 5: Microarray Data Analysis Design...............................................34
Figure 6: Number of genes in each Regulatory Bin .....................................36
Figure 7: Work-flow diagram for identifying transcriptional modules in similarly-regulated genes.................................................................58
Figure 8: Enhancer-reporter Activity in Transcription Factor Binding Site Clusters Found in Genes in the BMP-upregulated Bin..................................................69
Figure 9: Enhancer-reporter Activity in Transcription Factor Binding Site Clusters Found in Genes in the FGF-upregulated Bin........................................70
Figure 10: Enhancer-Reporter Activity using fragments of the *Ing1* promoter........72
Figure 11: Enhancer-Reporter Activity using fragments of the *HMGCR* promoter....73
Figure 12: Enhancer-Reporter Activity using fragments of the *Eomes* promoter ....74
Figure 13: Enhancer element response to FGF signaling..................................78
Figure 14: Fragments of the *Ing1* promoter, Co-transfected with FGFR.............79
LIST OF TABLES

Table 1: Over-represented motifs in BMP upregulated and FGF upregulated bins ........65
Table 2: Top 20 Most Common Over-represented Transcription Factor Binding Site Pairs found in the Same Cluster .................................................................66
ABSTRACT

An early event in animal development is the formation of the three primary germ layers that define the body plan. During gastrulation, cells migrate through the primitive streak of the embryo and undergo vast changes in morphology and gene expression, thus creating the mesodermal and endodermal cell layers. Gastrulation requires expression of Fibroblast Growth Factor (FGF), Wnt, and Platelet-Derived Growth Factor (PDGF). Previous work in our lab found that embryos treated with FGF inhibitors failed to gastrulate, as cell migration was completely halted. Microarray data from preingression epiblast tissue showed that inhibition of FGF signaling caused significant upregulation and downregulation of thousands of genes.

During gastrulation, 44 microRNAs are expressed in the primitive streak of *G. gallus* embryos, and six (microRNAs -let7b, -9, -19b, -107, -130b, and -218) are strongly upregulated when FGF signaling is blocked. The abundance of these six FGF-regulated microRNAs is controlled at various stages of processing: most are regulated transcriptionally, and three of them (let7b, 9, and 130b) are blocked by the presence of Lin28B, an RNA-binding protein upregulated by FGF signaling. These microRNAs target various serine/threonine and tyrosine kinase receptors. We propose a novel pathway by which FGF signaling downregulates several key microRNAs (partially through Lin28B), upregulating gene targets such as PDGFRA, which permits and directs cell migration during gastrulation. These findings add new layers of complexity to the
role that FGF signaling plays during embryogenesis, and provide a deeper understanding of the genetic "hardwiring" that allows cells to behave appropriately during development.

FGF signaling is also required for the formation of the heartfields, and has an overlapping pattern of expression with BMP (Bone Morphogenetic Protein). A microarray experiment using inhibitors of FGF and BMP found that thousands of genes in pre-cardiac mesoderm are affected by FGF signaling, BMP signaling, or a cooperative effect of the two. The promoter regions of similarly regulated genes were queried for over-represented transcription factor binding sites or novel DNA motifs. Subsequent cluster analysis of over-represented sites determined candidate transcriptional modules that were tested in primary cardiac myocyte and fibroblast cultures. About 75% of predicted modules in FGF-upregulated genes proved to be functional enhancers or repressors. Functional enhancers among FGF-upregulated genes contained clusters of CdxA and NFY sites, and increased transcription in the presence of a constitutively active FGF receptor.

This work combined multiple computational methods to create a unique pipeline for identifying transcription factors that drive expression in response to cell-cell signaling. While it has long been known that common transcriptional modules are responsible for similarly regulated genes, this method allows a more specific investigation into which transcription factor binding sites are responsible for gene regulatory patterns.
CHAPTER ONE: INTRODUCTION

1.1: GASTRULATION AND CARDIOGENESIS IN THE AVIAN EMBRYO

Avian embryos are often used as model systems for amniote development, as the genes and cell movements involved in early embryogenesis are similar to those in human development. Eggs for the chicken (Gallus gallus) are available in large quantity, and the embryos can be grown to any desired stage using a humid, warm environment. Embryos can be extracted from eggs with relative ease and cultured outside the egg, allowing for drug treatment and injection of foreign DNA (Stern, 2005).

1.1.1: Overview of early embryogenesis in Gallus gallus

After fertilization of an egg within the oviduct of a chicken, cleavage occurs in the center of an area of cytoplasm known as the blastodisk. Subsequent cell divisions cause the formation of a 5-6 cell layer of blastoderm tissue. The blastoderm is separated from the yolk by a space known as the subgerminal cavity. Cell death in the middle of the blastoderm leaves behind a central thin layer, the Area pellucida, surrounded by the thicker Area opaca (Gilbert, 2006). Shortly after, cells displaced from the embryo form an extra-embryonic structure (the hypoblast), attached to the yolk (Rosenquist, 1966).

One of the first characteristics in both mammalian and avian embryos is the primitive streak, a thickening down the middle of the embryo that is the first visible manifestation of the organism's axes (Hamburger and Hamilton, 1951). Streak formation begins when cells increase in height at the posterior border (or "marginal zone") between the Area
pellucida and Area opaca (Lawson and Schoenwolf, 2001). Those cells involved in streak formation migrate towards the anterior. Meanwhile, a depression, the primitive groove, forms in the middle of the primitive streak (Figure 1A).

A key feature of early embryo development is gastrulation, when the body plan, axes and germ layers of the organism are established. During gastrulation, the embryo forms three primary germ layers: ectoderm, mesoderm, and endoderm. Ectoderm, the outermost layer, gives rise to the epidermis and central nervous system. Mesoderm, the middle layer, is the precursor of a variety of organs and tissues, including the skeleton, muscles, heart, blood, and kidneys. Lastly, the inner endoderm layer becomes the inside of the digestive system and attached organs (Kalthoff, 2001).

The three layers of the body plan are established when cells in the epiblast (the side of the embryo opposite the hypoblast) migrate through the primitive streak. Migrating cells which pass through Henson's node, the furthest anterior point of the primitive streak, continue migrating anteriorly and displace cells in the hypoblast, forming the endoderm (Lawson and Schoenwolf, 2001). Most subsequent migrating cells move laterally after ingressing into the primitive streak, forming the mesoderm layer (Figure 1B) and also some endoderm. This process of migration is directed by Fibroblast Growth Factor (FGF)
Figure 1. Primitive Streak Formation and Gastrulation in Early Chick Development. (A) Brightfield images show the flat embryo during stages 1-3, as defined by Hamburger and Hamilton, 1951. The thin central area, the area pellucida, is surrounded by the thicker area opaca. During stages 2 and 3, the primitive streak can be seen forming in the posterior side of the embryo. (B) Cross-sectional diagram showing migration of cells in the epiblast through the primitive streak, some replacing the hypoblast to become endoderm (shown in yellow) while others migrate laterally to become the mesoderm. Partially adapted from Gilbert Developmental Biology.
signaling, as cells in the streak are attracted to FGF4, and repelled by FGF8 (Yang et al., 2002). During this process, migratory cells undergo Epithelial-to-Mesenchymal Transition (EMT), characterized by a loss of cell-cell adherence and apico-basal polarity, along with an increase in individual cell mobility. This requires a change in the expression of many proteins, including downregulation of cell-cell adhesion proteins (Nakaya and Sheng, 2008).

Some of the cells that migrate through the primitive streak form the pre-cardiac mesoderm, bilateral regions in the anterior of the embryo that eventually become heart cells. These regions are specified by the presence of BMP and FGF signaling through expression of BMP2 and FGF8, and inhibited in the central region of the embryo by Noggin that is secreted from the notochord (Alsan and Schultheiss, 2002). Specified precardiac cells are marked by expression of transcription factors GATA4 (Laverriere et al., 1994) and Nkx2-5 (Komuro and Izumo, 1993). Following differentiation, they are capable of "beating" on their own (Harary and Farley, 1963). These dual precardiac regions fuse at about 29 hours of development, forming a tube that comprises the early heart (Gilbert, 2006). A process of looping transforms this tube into the more mature heart.

1.1.2: Gastrulation in Gallus gallus

While gastrulation is conceptually similar in all animals, the exact mechanics vary considerably. In particular, there is a large difference between amniotes (birds, reptiles,
and mammals), whose eggs develop on land in the presence of an amnion, versus anamniotes (fish and amphibians), whose eggs lack an amnion and therefore must be laid in water. Anamniote gastrulation involves the folding of continuous sheets of cells. Frog embryos, for example, undergo gastrulation whereby sheets of cells invaginate, a feat accomplished with flexible cell-cell adhesion (Winklbauer, 2012). By contrast, gastrulation in amniotes involves EMT, with individual cells leaving the epiblast sheet of cells to form endoderm or mesoderm. Gastrulation in mammals is very similar to gastrulation in birds, despite the fact that mammals do not have a yolky egg. In both mammals and birds, gastrulation occurs through the formation of a primitive streak, with subsequent primary germ layers created by cells migrating through the streak and undergoing EMT (Shook and Keller, 2003). Oddly enough, gastrulating human embryos bear closer resemblance to chicken embryos than they do to mice embryos, as rodent epiblasts are curved.

During EMT, adhesion proteins expressed in epiblast cells are lost or modified. As such, one of the most important requirements for EMT is the repression of E-cadherin by the Snail family of transcription factors (Barrallo-Gimeno, 2005). At the same time, the basal lamina (an extracellular matrix holding the epiblast together) is broken down (Chuai and Weijer, 2009), and epithelial cells lose their polarity as they undergo EMT. As migratory cells ingress through the streak, other epiblast cells move towards the streak, replacing the position of the ones that undergo EMT. Besides the loss of certain proteins as cells move through the streak, many others are upregulated to promote cell movement:
vimentin, actin, and fibronectin are likely candidates for this function (Hay, 2005). After
the gastrulation stage finishes, the primitive streak regresses back towards the posterior
region. Development of the head region will continue in the anterior, while cell migration
continues in the posterior regions (Spratt, 1947).

Various signaling pathways (including Wnt, FGF, and PDGF) dictate the presence,
direction, and timing of cell migration during EMT. Previous work in our lab by Kathy
Hardy showed that non-canonical Wnt signaling is required for normal cell migration,
with Wnt11b and Wnt5a providing redundant functions (Hardy et al., 2008). Hardy also
found that inhibition of FGF signaling, a pathway that modifies the expression of
thousands of genes, causes migration to stop and cells to remain in the epiblast (Figure 2)
(Hardy et al., 2011). As mentioned before, patterns of movement can be dictated by
different FGF ligands, with FGF4 acting as an attractant and FGF8 acting as a repellent
(Yang et al., 2002). However, it is also known that the attractant property of FGF4
depends on the presence of PDGFRA (Platelet-Derived Growth Factor) signaling.
PDGFRA itself provides directional cues and is necessary for cell migration through the
primitive streak. PDGFRA operates through a pathway that activates PI3 kinase and is
required for N-cadherin expression (Yang et al., 2008).
Figure 2. FGF signaling is required for cell migration. (A) Merged brightfield and fluorescence images of control and treated embryos, using the FGF-inhibitor SU5402. Embryos are electroporated with a GFP construct so that their movements can be tracked. Defects in migration can be observed, especially by comparing cross-sections. (B) Quantification of GFP-expressing cell locations 5 hours after treatment. There are almost no mesodermal cells with SU5402 treatment, indicating that FGF signaling is required for formation of the mesoderm. Adapted from Hardy et al, 2011.
1.1.3: Fibroblast Growth Factor Signaling

Fibroblast Growth Factors (FGFs) are a large family of signaling molecules found in many species. There are 22 unique FGF ligands in human, and 12 in chicken (Thisse and Thisse, 2005). FGF ligands contain a core domain of 120 amino acids that interact with four types of FGF receptors in a dose-dependent manner. The FGF receptors are transmembrane proteins that undergo dimerization and autophosphorylation when bound to a ligand on the extracellular domain. This autophosphorylation triggers tyrosine kinase activity that recruits signaling complexes that activate downstream signaling pathways (Böttcher and Niehrs, 2005).

Three signal transduction pathways are activated when FGF ligands bind to FGF receptors: the Ras/MAPK pathway, the PLCy/Ca$^{2+}$ pathway, and the PI3 kinase/Akt pathway. The Ras/MAPK pathway starts with tyrosine phosphorylation of FRS2α/β, which constitutively bind to the intracellular domain of FGFRs. This leads to the formation of a complex between adaptor protein Grb2 and guanine nucleotide exchange factor Sos, which activates Ras through GTP exchange. This reaction activates the Ras and MAP kinase cascades, which phosphorylate many transcription factors. (Lunn et al., 2007). The PLCy/Ca$^{2+}$ pathway involves the binding of PLCy to FGFRs, and the subsequent release of Ca$^{2+}$ and activation of protein kinase C. The PI3 kinase/Akt pathway can be activated by multiple inputs (Böttcher and Niehrs, 2005).
FGF signaling plays critical roles in embryo development of a wide variety of species. For example, studies in *Xenopus laevis* have found that FGF signaling regulates the expression of genes required for the function of the Spemann organizer (Branney et al., 2009). In chicken, different FGF receptors have varying spatial and temporal patterns in the developing embryo, and FGFR activity is necessary for cell differentiation (Lunn et al., 2007). As noted before, FGF ligands can act as cell attractants or repellents during gastrulation (Yang et al., 2002), and FGFR activity is required for cell migration during gastrulation (Hardy et al., 2011).

In this dissertation (Chapter 2 and Appendix A), I will describe work that further elucidates the underlying activity of FGF signaling during gastrulation, by looking at the effect it has on the microRNA abundance.

### 1.1.4: Cardiac Myogenesis

In birds, reptiles, and mammals, the cardiac progenitor cells come from anterior regions of the primitive streak during early stages of gastrulation (stages 3a-3c). By the time embryos reach stage 5, premyocardial cells have migrated away from the streak to form two regions of cardiogenic anterior lateral mesoderm known as the heart fields. By stage 7, the two heart fields have migrated anteriorly and medially to form the cardiac crescent. The heart fields then begin to fuse and by stage 10 have formed the primitive heart tube (Brand, 2003). FGF signaling and canonical Wnt signaling are required for the cell
movements that form the heart fields (Ciruna and Rossant, 2001). Figure 3 shows a diagram of the precardiac cell locations at different stages.

Cardiogenesis requires the activation of several key transcription factors. The first identified cardiogenic regulatory factor was tinman in the fruitfly Drosophila melanogaster, an NK-class homeodomain transcription factor. Tinman orthologs have been found in vertebrates, where it is known as Nkx2-5 (Harvey, 1996). During evolution the tinman gene has duplicated and amniotes contain several Nkx2-5 homologues. In Drosophila, loss of tinman leads to absence of the dorsal vessel, while in vertebrates lacking Nkx2-5 the heart forms but is highly abnormal (Bodmer, 1993; Lyons et al., 1995). Gene ablation studies have identified several other transcription factors that play an important role in heart development. These include Myocardin, a transcriptional co-activator of serum response factor (Bruneau, 2002), the basic helix-loop-helix transcription factors MesP1 and MesP2 (Saga et al., 1999), and the GATA family of transcription factors (Laverriere et al., 1994). These and other factors participate in a transcriptional network that regulates time and place of heart development. Expression of these transcription factors is regulated by growth factor signaling, and thus there has been much interest in the signaling pathways that are active during heart field formation and migration.

Besides FGF and Wnt, one of the key signaling families required for cardiogenesis is the TGF-β superfamily, which includes BMPs (Bone Morphogenetic Proteins). In the
cascade of heart development, BMPs are responsible for activating GATA4 and Nkx2-5 (Gilbert, 2006). The BMP ligand binds to a heterodimeric complex consisting of a Type 1 and Type 2 TGF-β receptor. When activated, the two receptors phosphorylate each other, causing downstream phosphorylation of Smads 1, 4, and 5, and activation of a specific transcriptional program (Massague, 2005). Chicken embryos show an overlapping pattern of FGF and BMP expression in the precardiac mesoderm. In this dissertation (Chapter 3) I will describe methods that examine genes affected by BMP and FGF in the heart field, and identify cis-acting transcriptional components that may connect each of these signaling pathways to the genes that they control.
Figure 3. Location of Heart Precursor Cells. At stage 3, heart precursors are located in the anterior of the primitive streak, just posterior to Hensen's Node. By stage 5, the cells have migrated to form the bilateral primary heart fields. By stage 7, the two heart fields have fused to form the cardiac crescent. Adapted from Brand, 2003.
1.2: MICRORNAS IN EMBRYONIC DEVELOPMENT

MicroRNAs are short, 21-23 nucleotide RNA molecules that attenuate gene expression through either translational inhibition or mRNA degradation. The first microRNA discovered was lin-4 in the nematode worm *Caenorhabditis elegans*. Lin-4 was noted to have a homology to the 3' untranslated regions of developmental genes lin-14 and lin-28 (Lee et al., 1993). Subsequent work identified a second *C. elegans* microRNA, let-7, and confirmed that microRNAs act as repressors of their mRNA targets. Since then, hundreds of microRNAs have been found in many different species. (Grosshans and Slack, 2002).

MicroRNAs are important in development and evolution. Expansion of microRNAs by genome duplication seems to have played a key role in vertebrate evolution (Gu et al., 2009). The actual sequences of microRNAs are highly conserved between different vertebrate species, with most of the variation seen in the abundance and timing of their expression. Species with highly divergent physiology are likely to contain large differences in the expression patterns of microRNAs during their development (Ason, 2006). Studies in mice have shown that microRNAs are critical for proper development, as loss of the microRNA processing enzyme Dicer causes early embryonic lethality (Bernstein et al., 2003). With regard to chicken embryos, previous work in our lab identified 84 microRNAs that are expressed prior to day 5 of embryogenesis. *In situ* hybridization analyses revealed that microRNAs let-7b, 130b, and 367 are expressed at or near the primitive streak in gastrulation-stage embryos (Darnell et al., 2006).
During a study of cell migration during gastrulation, our lab treated chicken embryos with the small molecule inhibitor SU5402 to block FGF signaling. Microarray and *in situ* hybridization analyses showed that hundreds of genes are upregulated or downregulated as a result of FGF inhibition (Hardy et al., 2011). One of the most prominently downregulated genes was Lin28B, an RNA-binding protein that regulates microRNA processing. This was our first clue of a link between gastrulation, FGF signaling, and microRNA function.

1.2.1: MicroRNA formation and function

Although microRNAs are very small (about 21 nucleotides), they come from much larger precursor molecules (See Figure 4). The "primary transcript" of a microRNA is a normal ~1kb RNA molecule transcribed by RNA Polymerase II (Pol II), and given the normal post-transcriptional modifications of a 5' guanosine cap and a 3' poly-adenine tail. Some of these primary transcripts are encoded within the intronic regions of other mRNAs (such as protein-coding genes) while others have their own promoters (Zeng, 2006).
Figure 4. Pathway for microRNA biogenesis. MicroRNAs arise as regular-sized RNAs transcribed by RNA Polymerase II. A 80-100 nucleotide structure is excised by the enzymes Drosha and Pasha (DGCR8 in humans), which is then exported into the cytoplasm by Exportin-5. This "pre-microRNA" structure is further processed by Dicer, which complexes with the RISC complex to become the active, final microRNA
The primary transcript for a microRNA always contains a stemloop secondary structure about 80-100 nucleotides in length. This stemloop structure is processed out of the primary transcript to become the "pre-microRNA" by a nuclear RNAse III enzyme called Drosha, acting alongside the regulatory subunit DGCR8 (in humans) or Pasha (in flies) (Lee, 2003). Following this step, the pre-microRNA is exported from the nucleus by an Exportin 5/Ran-GTP transporter, and then processed to its final 21-23 nucleotide length by the enzyme Dicer and joined with the RISC complex (Zeng, 2006). Stability of the microRNA at this point is guaranteed by binding to the Argonaute protein, while release from Argonaute leaves the microRNA available for degradation by XRN-2 or SDN exonucleases (Kai and Pasquinelli, 2010).

MicroRNA processing can be blocked by a protein called Lin28. Lin28 binds to microRNA precursors, preventing them from being further processed into mature microRNAs. This leads to reduced abundance of the mature microRNA. Lin28 binding appears to be highly specific to the let-7 family of microRNAs (Lehrbach et al., 2009; Newman et al., 2008; Viswanathan, 2008). However, more recent work showed an effect of Lin28 on microRNA-1 (Rau et al., 2011) and this dissertation will present evidence that Lin28 also affects other families of microRNAs.

As microRNAs are too small to code for proteins, their function instead relies on "silencing" the protein-coding mRNAs at appropriate times and places. There are two methods by which microRNAs can achieve this: (1) by repressing translation of mRNAs,
or (2) by inducing cleavage and therefore destruction of the target mRNA. The latter mechanism (cleavage) is more common in plants than in animals, though some animal microRNAs such as microRNA-1 and microRNA-124 have been shown to degrade hundreds of target mRNAs (Lim et al., 2005). To suppress the production of particular proteins, microRNAs bind to specific "seed" sites corresponding to the first eight nucleotides of the microRNA (Bartel, 2009; Mourelatos, 2008). However, other considerations determine whether a seed site is functional, such as gene context, positioning, and matching bases at nucleotides 13-16 of the microRNA (Grimson, 2007).

1.2.2: The let-7 family of microRNAs.

There is great interest in the let-7 family of microRNAs due to their profound impact on both development and human disease. MicroRNA-let7 was the first human microRNA discovered, and let-7 family members have since been identified in all eukaryotes. As of this writing, there are 11 let-7 microRNAs in human, and 12 let-7 microRNAs in chicken (including one novel chicken let-7 gene discovered in this dissertation) (Kozomara and Griffiths-Jones, 2011). The name "let-7" is short for "lethal 7", as knockouts of the gene were found to be lethal in C. elegans (Roush and Slack, 2008).

The general function of let-7 is to promote cell differentiation. Increased abundance of let-7 causes cells to lose pluripotency and become specialized, whereas loss of let-7 (and a corresponding increase of let-7 antagonizers, such as Lin28) causes cells to retain undifferentiated stem-cell-like attributes (Slack, 2010). Other microRNAs oppose let-7's function. In mouse embryonic stem cells, let-7 had no capacity to cause differentiation
unless the innate stem cell microRNAs were depleted (Melton et al., 2010). Let-7 also has a role in vascular diseases. Let-7f, along with microRNA-27b and microRNA-130a are pro-angiogenic factors (Urbich et al., 2008). Consistent with the paradigm that let-7 maintains a differentiated state, many human cancers contain excessively low levels of let-7 (Boyerinas et al., 2010).

As discussed above, let-7 biogenesis can be blocked by the presence of the RNA-binding protein Lin28. This interaction comprises a double-negative feedback loop, as let-7 in turn blocks the expression of Lin28 (Guo, 2006; Yang et al., 2010). Due to the known association of let-7 and Lin28, it is not surprising that Lin28 is also associated with cancer. Lin28 is a proliferative factor in liver cancer (Guo, 2006), a necessary factor in breast cancer metastasis (Dangi-Garimella et al., 2009), and overall is overexpressed in 15% of human cancer cell lines (Viswanathan et al., 2009). By the same token, Lin28 promotes stem cell formation (West, 2009), and is expressed in developing limb buds (Yokoyama et al., 2008).

The exact method by which Lin28 prevents let-7 maturation seems to vary. Some studies showed that Lin28 functions by blocking the interaction between primary let-7 transcripts and Drosha, preventing the initial processing step (Newman et al., 2008; Viswanathan, 2008). However, other studies found that Lin28 can bind the processed pre-microRNAs (Piskounova et al., 2008) and prevent interaction with Dicer (Heo, 2008; Rybak et al., 2008). This slight discrepancy could be explained by differences between the let-7
microRNA family members (Roush and Slack, 2008). In addition, there are two Lin28 homologs, Lin28A and Lin28B, which inhibit microRNA formation by distinct mechanisms. Lin28A blocks let-7 precursors from being chopped up by Dicer in the cytoplasm, while Lin28B sequesters primary transcripts in the nucleus (Piskounova et al., 2011).

1.2.3: Summary of Approach

The "big picture" purpose of this project is to understand how embryos are formed. Previous work has identified cell movements, signaling pathways, and genes that are required for an early embryo to undergo gastrulation and form its basic body plan. However, we are still attempting to decipher the underlying "circuitry" of animal development, which links the signaling molecules to their affected genes and ultimately to cell behaviors and phenotypes. Since each microRNA has the potential to impact hundreds of genes, any effect of FGF signaling that modifies microRNA abundance may explain the effects of this pathway on gene expression and cell behaviors during gastrulation (Hardy et al., 2011).

Thus, the goal of my project is to understand the relationship between FGF signaling, gastrulation, and microRNA abundance. Previous work in the field has explored the connection between FGF signaling and microRNAs: in chicken, it was found that FGF 4 negatively regulated the expression of microRNA 206, which is restricted to developing somites (Sweetman et al., 2006). Meanwhile, studies in zebrafish fin regeneration found
that FGFR1 knockout caused the downregulation of 22 microRNAs, and the upregulation of 34 (Yin et al., 2008). Many microRNA studies do not address the question of how microRNAs are regulated, although work in mouse embryonic fibroblasts showed that FGF2 (acting through NDY1) could repress microRNA-101 at a transcriptional level (Kottakis et al., 2011). Overall, however, there is not much known about the how FGF signaling might regulate microRNA abundance and function, or how this might impact gastrulation.

This study answers several questions: (1) Which microRNAs are expressed at the primitive streak during gastrulation? (2) Which microRNAs are regulated by FGF signaling in gastrulation-stage embryos? (3) How are these microRNAs regulated? (4) Is there a connection between Lin28B (which is upregulated by FGF signaling) and microRNA abundance? (5) What are the mRNA targets of these FGF-controlled microRNAs? (6) Do any of these mRNA targets have biological significance, possibly affecting gastrulation or other developmental processes?

To address these questions, I first present microarray and qPCR data establishing which microRNAs are expressed at the primitive streak, and which are affected by the FGF inhibitor SU5402. To assess which forms of each microRNA are regulated (primary transcript, stem-loop precursor, or mature), I used separate cDNAs and PCR primers that distinguish between these forms (Schmittgen et al., 2004). To study the effect of Lin28B on microRNA abundance, experimental embryos were electroporated (Voiculescu et al.,
2008) with either plasmids that overexpress Lin28B or splice-blocking morpholinos (GeneTools) that knock down Lin28B.

We are interested not just in how FGF signaling influences microRNA abundance, but also which mRNA transcripts are likely targeted by FGF-regulated microRNAs. To accomplish this, we employed prediction algorithms searching for microRNA target sites conserved by evolution (Bartel, 2009), and tested the most promising sites using a luciferase UTR-reporter assay. The effect of microRNA targeting was also tested in vivo, using electroporated microRNA mimics and a Lin28B morpholino.

This work includes novel discoveries relating to embryogenesis. This study will demonstrate that FGF signaling suppresses the abundance of 6 differentially-regulated microRNAs which leads to increased levels of key target genes such as PDGFRA, permitting cell migration through the primitive streak.
1.3: COMBINATORIAL INPUTS OF BMP AND FGF SIGNALING IN CARDIAC MYOGENESIS

Previously, I discussed formation of the heart fields, bilateral regions of mesodermal tissue that migrate anteriorly to merge and form the cardiac crescent, which becomes the animal's heart. Specification of the precardiac tissue in the heart fields depends on overlapping patterns of BMP and FGF signaling. In this subchapter, I will discuss a project that assessed the transcriptional impact of BMP and FGF signaling in the heart fields, and concepts in bioinformatics used to explore this process further.

1.3.1: FGF and BMP Regulatory Bins

In a previous project in our lab, stage 4 chicken embryos were treated with small molecule inhibitors that block either FGF signaling, BMP signaling, or both. The embryos developed for several hours until they reached stage 7, at which point cardiogenic mesoderm was dissected from the two heart fields. This created four separate experimental groups: untreated, FGF-inhibited (treated with SU5402), BMP-inhibited (treated with Dorsomorphin), and FGF/BMP-inhibited (treated with both SU5402 and Dorsomorphin). Labeled RNA from this tissue was hybridized to a custom 20,500 feature long oligo microarray, allowing comparison of the four different experimental groups (Konieczka, 2012).

The data from the microarrays was adjusted for dye-swap replicates and normalized to account for location on the array (Futschik and Crompton, 2005). To compare the
experimental groups in a meaningful way, a "wheel" design matrix was developed (based on Capaldi et al., 2008) that allows use of linear regression to extract the contributions of FGF signaling, the contributions of BMP signaling, and the Co-operative effect in the presence of both pathways (Figure 5).

For example, contributions of FGF signaling alone can be determined by comparison of the Dorsomorphin-treated sample to embryos treated with both Dorsomorphin and SU5402. Since BMP is inhibited in both cases, any differences cannot be due to BMP signaling or a Cooperative effect, but only FGF signaling. Conversely, the contribution of BMP signaling can be determined by comparing samples in which both have been treated with SU5402 (thus negating FGF signaling from the comparison) but only one has been treated with Dorsomorphin. Other comparisons contain contributions from multiple sources. For example, comparing samples from untreated embryos to that treated with both inhibitors derives the contributions from FGF signaling, BMP signaling, and Cooperative (Co) effects. As a result of this analysis, we can identify which genes are upregulated, downregulated, or unaffected by FGF, BMP, or Cooperative signaling.
Figure 5. Microarray Data Analysis Design. This diagram shows the comparisons that are made between RNA samples from tissue treated (or untreated) with inhibitors for Fibroblast Growth Factor (FGF), Bone Morphogenetic Protein (BMP), or Cooperative (Co) signaling pathways. Based on each comparison, the contributions of individual signaling pathways are measured multiple times, along with any cooperative effect. Concept adapted from Capaldi et al. 2008.
To systematically categorize the influence of BMP or FGF signaling pathways on all genes expressed in the precardiac mesoderm, changes in observed gene expression were converted to integers based up on a statistical cutoff as follows: "1" for upregulation, "-1" for downregulation, and "0" for unaffected. Given that each gene can be affected by 3 signaling pathways (BMP, FGF, or Co), and each has 3 potential effects (1, -1, or 0), 27 (3^3) combinations are possible for describing how each gene is regulated. We termed each of these possibilities "regulatory bins", reflecting a common regulatory pattern with respect to FGF and BMP signaling. Bins are described by the aforementioned number code. For example, Bin "1,0,0" describes genes upregulated by BMP, Bin "0,-1,0" describes genes downregulated by FGF, and so on. Figure 6 shows how many genes qualify for each Bin.

Based on the number of genes in each regulatory bin, several patterns were noticed: (1) The number of genes upregulated by any combination of signaling inputs is similar to the number of genes downregulated in the same way. This is apparent by the symmetry of Figure 6 (symmetry is conserved with any systematic ordering of the graph). (2) FGF and BMP signaling rarely act in opposite directions on the same gene. (3) Individual signaling pathways (FGF or BMP) rarely act in the same direction as they do Cooperatively. (4) A plurality of genes regulated by FGF and/or BMP signaling in cardiac myogenesis are regulated cooperatively, rather than by individual components.
Figure 6. **Number of genes in each Regulatory Bin.** Using RNA from embryos treated with inhibitors for BMP and/or FGF, we determined the number of genes upregulated, downregulated, or unaffected by BMP, FGF, or a cooperative effect of both pathways (Co). The number of genes that match each possible combination of regulatory effects is displayed above. However, the (0,0,0) bin, those genes unaffected by any treatment, is not displayed for graphical reasons. This bin contains more than 11,000 genes.
The above paradigm regarding Regulatory Bins ties into the larger idea of genomic regulatory systems in development. That is, transcription target sites in certain genes create an underlying "hardwiring" of the cell, allowing it to respond appropriately to the correct signals at the correct time (Arnone and Davidson, 1997). The promoter regions of these genes contain a number of cis-acting transcriptional target sites, that, when present in the correct combination, create "transcriptional modules" that define the group of co-expressed genes (Leyfer and Weng, 2005). Understanding these cis-acting transcriptional elements and how they enable a gene to respond to signaling molecules is key to understanding embryonic development at a genetic level.

Thus, in this study, I will describe approaches to identify transcription target sites that characterize each regulatory bin, assemble these sites into candidate transcriptional modules, and test these predictions. There are two main approaches used to find cis-acting transcriptional elements. The first is to use previous information on transcription factor binding sites, searching for strong matches to already-known sites. The second is to search for any DNA motif that appears to characterize a Regulatory Bin, without prior bias to whether it resembles a known transcriptional element.

1.3.2: Computational Predictions of Transcription Factor Binding Motifs

One of the most common methods to describe DNA motifs is with a position-specific scoring matrix (PSSM), alternatively known as a position weight matrix (PSW) (Bailey and Elkan, 1994). A PSSM appears as a matrix of numbers, with a variable number of
rows representing each nucleotide in the motif, and 4 columns representing the likelihood of each nucleotide base (Adenine, Cytosine, Guanine, and Thymidine, in that exact order). Each value in the matrix assigns the likelihood of each base appearing at that particular position in the DNA motif. For example, if Row #1 in the matrix is (100,0,0,0), then there is a 100% chance that the first nucleotide base in the motif is Adenine. If Row #2 is (0,70,0,30), then there is a 70% chance that the second nucleotide base is Cytosine, and 30% chance that the base is Thymidine. In this way, the PSSM allows for more flexibility than a simple string of characters, allowing for ambiguity in the sequence of a transcription factor binding site.

By assigning odds of specific nucleotides appearing at specific positions within a DNA motif, a PSSM returns a log-likelihood score, assessing the probability that a given sequence of DNA contains a particular transcription factor binding site. Adjustments are usually made based on the background frequency of nucleotide bases. Software such as Match (Kel et al., 2003) or MAST (Motif Alignment and Search Tool) (Bailey and Gribskov, 1998) are capable of scanning large regions of DNA for high-scoring motif matches. Previous work has verified the usefulness of computationally-predicted transcription factor binding sites. A study in *E. coli* found that Bayesian prediction models using PSSMs correctly identified 238 experimentally verified sigma-70 binding sites (Ben-Gal et al., 2005). In a study of cell-cycle regulatory genes, Elkon et al. used computationally-predicted motifs to identify 8 transcription factors that controlled cell-cycle genes (Elkon et al., 2003).
To create PSSMs for known transcription factor binding sites, various databases have produced alignments of verified sites, often combining multi-species data. By aligning verified sites from different sources, we can clearly identify which nucleotides in a binding site are always the same, and which contain ambiguity. This is reflected in the PSSM, by assigning 100% odds to unchanging nucleotides and lower odds to more ambiguous positions. A large repository of this information has been collected in the TRANSFAC database (Wingender et al., 1996). In addition, PSSMs can be obtained simply by analyzing similarly-regulated sequences, and identifying over-represented motifs (Bailey et al., 2006).

**1.3.3: Summary of Approach**

The goal of this project is to identify regulatory modules of transcription factor binding sites that cause genes to respond to FGF signaling, BMP signaling, or some combination of the two. This project will combine our knowledge of FGF and BMP regulatory bins during cardiac myogenesis (Section 1.3.1) and paradigms of computationally-predicted transcription factor binding sites (Section 1.3.2). As part of this dissertation, I will describe the identification of transcriptional modules *in silico*, and subsequent experiments that test the functionality of these predictions.

Functional transcription factor binding sites are primarily (though not exclusively) located upstream of the transcription start site (Linhart et al., 2008). Therefore, we
defined the relevant "promoter region" of each gene to be 10 kilobases 5' of the transcription start site. In order to predict promoter elements that respond to FGF and BMP signaling, we identified motifs that are over-represented among the promoters of certain regulatory bins, as compared to the frequency of the same motifs in every other promoter region in the genome. With regard to known transcription factor binding sites, we referred to the hypergeometric distribution, defined as \( P = \frac{\binom{a}{b} \binom{N-a}{n-b}}{\binom{N}{n}} \), where "P" is the probability of the null hypothesis (i.e. not over-represented), "N" is the total number of promoters in the genome, "n" is the size of the regulatory bin, "a" is the number of times a transcription factor binding site appears in the genome, "b" is the number of times a binding site appears in a regulatory bin, and \( \binom{x}{y} \) is the binomial coefficient of x and y (Rice, 2007).

A second approach is to locate over-represented motifs that may be novel (previously-unknown) sites. The Amadeus platform is able to preform such analysis (Linhart et al., 2008). Amadeus is included in a user-friendly software package called Allegro (A Log-Likelihood based Engine for Gene expression Regulatory motifs Over-representation). Allegro finds novel overrepresented motifs by breaking input sequences into candidate k-mers (where k is a specified motif length) and determining log-likelihood over-representation scores using a hypergeometric distribution (Halperin et al., 2009).
The above methods for identifying over-represented motifs in gene promoters are likely to generate much redundancy, producing motifs that are highly similar or overlapping. To condense a set of candidate motifs, Naomi Habib has developed the BliC algorithm, which clusters and merges similar motifs in a regulatory set (Habib et al., 2008). After clustering DNA sequence motifs is complete, a final list of transcription factor binding sites can be compiled. However, this list would only comprise individual elements of a transcriptional module that allows a gene to be regulated (Leyfer and Weng, 2005). To recognize combinations of DNA binding sites in close proximity, the Cluster Buster algorithm evaluates the strength of PSSM matches in a given sequence and the proximity of proposed DNA-sequence motifs to each other, compared to a model of independent, random nucleotides (Frith et al., 2003). This identifies sequences of DNA that are strong candidates as functional transcriptional modules.

Finally, we used primary cell cultures of cardiac myocytes and fibroblasts to test predicted transcriptional modules. Predicted clusters of over-represented transcription factors were cloned into a secreted luciferase enhancer-reporter vector, then transfected into cardiac cells. This enables us to test whether our predicted transcriptional modules cause upregulation, downregulation, or have no effect. In addition, co-transfections with a constitutively active or dominant-negative FGFR determined whether these transcriptional modules respond to FGF signaling.
1.4: EXPLANATION OF DISSERTATION FORMAT

I present my dissertation in four chapters and one appendix. In Chapter 1, I provide a detailed summary of the research questions and their context, including a review of previous work and the literature on this subject.

In Chapter 2, I present a summary of major results that establish a link between FGF signaling, microRNAs, and the regulation of gastrulation. Aside from development of the \textit{in situ} hybridization images (done by the GEISHA group), this experimental work was organized, conducted, and analyzed by myself with guidance from Dr. Parker Antin and significant technical assistance from my undergraduate assistant Alesksi Saarela and laboratory manager Tatiana Yatskievych. The four of us are co-authors on a publication in the Journal of Biological Chemistry entitled "Fibroblast Growth Factor (FGF) signaling negatively regulates the abundance of microRNAs that regulate proteins required for cell migration and embryo patterning during gastrulation" (Bobbs et al., 2012). This article, along with all figures and supplemental materials, is provided as Appendix A of this dissertation to serve as a comprehensive description of the project.

Chapter 3 of this dissertation provides a complete description of a second project, investigating the regulatory inputs of FGF and BMP signaling on heart field development, and characterization of cis-acting transcriptional elements. Working under Dr. Antin, I preformed all of the bioinformatics and computational work, and the majority of the experimental work. Postdoctoral fellow James Cooley assisted by constructing
many of the DNA vectors used in these experiments, and Tatiana Yatskievych and Dr. Cooley dissected the embryos that I used to prepare primary cell cultures. Unlike the previous project concerning microRNAs, this work has not yet been published, and is thus presented entirely in the body of this dissertation without the inclusion of a journal article.

Chapter 4 provides some concluding thoughts on the general principles explored in both projects.
1.5: SUMMARY

We are using chicken (*Gallus gallus*) embryos as a model for human development. Two vital processes in embryo development are gastrulation (the formation of the three basic layers of the body plan) and the formation of precardiac mesoderm in the heart fields. Both of these developments depend on proper cell-cell signaling, with gastrulation depending on FGF, PDGFRA, and Wnt signaling, while heart field development relies on FGF and BMP signaling. To further understand this process, I have completed two projects that are described in the next two chapters this dissertation.

The first project explores the role of microRNAs during gastrulation. MicroRNAs are small 21-nucleotide molecules that are processed from larger transcripts, and can suppress the translation of messenger RNAs. Our lab has previously established that certain microRNAs such as let-7b (which has a known role in cell differentiation) are expressed in gastrulating chicken embryos. Furthermore, we know that FGF signaling upregulates an RNA-binding protein called Lin28B that blocks microRNA formation. Based on these facts, I sought to understand which microRNAs are affected by FGF signaling, how they are affected, at what the downstream consequence of this regulation may be.

The second project builds off of a previous dissertation by Jay Konieska, in which he discovered that hundreds of genes in pre-cardiac tissue are regulated by a complex combination of FGF and BMP signaling. Taking these findings, I applied computational
techniques employing numerical representations of transcription factor binding sites, called Position-Specific Scoring Matrices (PSSM) in order to find over-represented motifs that characterize genes that are either up-regulated or downregulated by BMP, FGF, or a cooperative component of the two. This approach involves looking for known transcription factor binding sites, as well as discovering novel, previously uncharacterized ones. In addition, we designed experiments to test these predictions in vitro, determining functional significance of transcriptional modules, as well as cell-type specificity and responsiveness to FGF signaling.
CHAPTER TWO: FIBROBLAST GROWTH FACTOR (FGF) SIGNALING AND MICRORNAs IN AVIAN GASTRULATION

2.1: PRESENT STUDY

The methods, results and conclusions of this study are described in a journal publication appended to this dissertation. The following is a summary of the most important findings of this study:

FIBROBLAST GROWTH FACTOR (FGF) SIGNALING NEGATIVELY REGULATES THE ABUNDANCE OF MICRORNAs THAT TARGET MOLECULES REQUIRED FOR CELL MIGRATION AND EMBRYO PATTERNING DURING GASTRULATION (full details in Appendix A)

The basic body plan is established by the process of gastrulation, during which three primary cell layers are formed (endoderm, mesoderm, and ectoderm). In animals that develop on land (amniotes), this occurs by migration of cells through a region known as the primitive streak, followed by subsequent epithelial-mesenchymal-transition. Previous work has shown repeatedly that FGF and PDGFRA signaling are required for both the presence and direction of cell migration during gastrulation. Inhibition of these pathways causes upregulation and downregulation of hundreds of genes. Among those genes most strongly regulated by FGF signaling is an RNA-binding protein called Lin28B, which selectively inhibits microRNA formation.

Using a multi-species microarray, we detected 44 known chicken microRNAs near the primitive streak region of gastrulation-stage chick embryos. In the same experiment, we also found 4 microRNAs that had been discovered in other species but were previously
not known to exist in chicken. Real-time quantitative PCR determined that six microRNAs (microRNAs -let-7b, -9, -19b, -107, -130b, -218) were highly upregulated in samples treated with SU5402, a small molecule inhibitor of FGF signaling. The most strongly affected microRNA was let-7b, with almost a 50-fold change in response to inhibition of FGF signaling. The other five affected microRNAs changed between 6.5 and 20-fold.

To further characterize microRNA regulation through FGF signaling, we measured each of the six FGF-regulated microRNAs at different stages of processing. Most are likely regulated transcriptionally, as treatment with SU5402 caused an increase in primary transcript levels. However, two microRNAs (130b and 19b) showed significant changes in their intermediate products, the 80-100 nucleotide stem loop structures known as pre-microRNAs. Gain and loss of function studies showed that these microRNAs are also regulated in part by Lin28B. MicroRNA let7-b, a known target of Lin28, was also verified to be responsive to Lin28B manipulations in vivo.

Various computational methods designed around the TargetScan algorithm identified approximately 580 mRNAs that are likely targets of the six microRNAs downregulated by FGF signaling. The 3’UTRs of fifteen predicted target mRNAs, including PDGFRA, ACVR1, ACVR2B, LIN28B, ABL2, LRIG1, CPEB1, CPEB3, and CPEB4 were tested in a cell culture assay and found to contain functional microRNA binding sites capable of reducing expression of a heterologous luciferase mRNA. Of all those tested, only one computationally predicted target (COL3A) was not functional. Follow-up measurements
determined that, in general, genes targeted by FGF-regulated microRNAs were repressed in translation, not transcription. PDGFRA was an exception, as the transcript is degraded by microRNA-218 both \textit{in vitro} and \textit{in vivo}.

From these results we proposed a model by which FGF signaling, partially through the activity of Lin28B, represses the abundance of specific microRNAs. This leads to an increase in expression of proteins encoded by target mRNAs, including PDGFRA, which is necessary for cell movements during gastrulation. This complete story adds another layer of complexity to the regulatory network controlled by FGF signaling during gastrulation.


2.2: FUTURE DIRECTIONS

In the present study, we established a role for FGF signaling in regulating microRNAs that regulate expression proteins important for gastrulation, including PDGFRA. In addition to PDGFRA, FGF-regulated microRNAs also affect a number of other serine/tyrosine kinases, such as TGFβRs (tumor growth factor beta receptors) and ACVRs (Activin A Receptor). A future project may further explore these receptors, asking what role (if any) they play in gastrulation, as determined by knockout studies, and forced expression of the microRNAs that target these genes.

A connection was also identified between microRNA-let-7b and members of the CPEB (cytoplasmic polyadenylation binding) family. CPEB proteins act as adenylation factors, increasing the length of poly-A tails on certain messenger RNAs and thereby enhancing their stability and translation. An interesting aspect of CPEB activity is that CPEB recognizes a target sequence within the 3’ UTR of mRNAs, similar in concept to the mechanism by which microRNAs target CPEB itself (Richter, 2007). Thus, using methods similar to TargetScan, we could identify genes that contain conserved CPEB binding sites. Functionality of CPEB binding sites could be tested by force-expressing or knocking down CPEB, and observing the poly-A tail length of candidate target molecules (Kim, 2008). I hypothesize the following model: FGF signaling downregulates let-7b, which upregulates CPEB 1/3/4, which activates a set of unknown target genes. The net affect would be that any genes activated by CPEB would ultimately be upregulated by FGF signaling. However, this view is
complicated in light of the fact that CPEB can also act as a repressor in some circumstances (de Moor, 1999).

Finally, there is still the issue of the many microRNAs present in the primitive streak of gastrulating embryos that are not regulated by FGF signaling. While the work in this dissertation focused on the impact of FGF signaling, nearly 40 microRNAs are expressed at the same time and place, and may play a role in early embryo development that is independent of FGFR. In particular, members of the microRNA-302 family are expressed at extremely high levels (Appendix A, Supplementary Table 1). Future studies may consider the probable targets of these molecules, and any impact they may have on embryo development and gastrulation in particular.
CHAPTER THREE: COMPUTATIONAL AND EXPERIMENTAL ANALYSIS OF PROMOTER REGIONS CHARACTERIZING REGULATORY "BINS" IN FIBROBLAST GROWTH FACTOR (FGF) AND BONE MORPHOGENETIC PROTEIN (BMP) SIGNALING PATHWAYS.

3.1: ABSTRACT

The primary and secondary heart fields of G. gallus embryos are established and maintained by overlapping patterns of FGF and BMP signaling. Based on previous microarray studies, we defined 27 "regulatory bins" that represent all possible responses any gene may have to Fibroblast Growth Factor (FGF), Bone Morphogenetic Protein (BMP), or a cooperative effect of the two. Based on a genome-wide survey of known transcription factor binding sites in the promoter regions of chicken genes, all of the sizeable regulatory bins (>100 genes) contain a statistically significant set of known over-represented DNA motifs. This data was combined with over-represented novel motifs discovered by the Allegro software, and used to predict clusters of transcription factor binding sites that form potential transcriptional modules. Between 66% and 75% of these predicted modules proved functional in enhancer test assays using cultured primary cardiac myocytes. We also show that computationally-derived enhancer elements identified in promoter regions of genes predicted to be FGF responsive were upregulated in response to a constitutively-active FGF receptor. This work provides a computational approach for explaining how large groups of genes can respond differentially to complex signaling inputs.
3.2: INTRODUCTION

As heart disease is a leading cause of death among Americans, there has been much interest in understanding how cardiac tissue is formed. The amniote heart develops from bilateral regions of mesodermal tissue known as the heart fields. These regions of the developing embryo consist of pre-cardiac cells that originate from the primitive streak during gastrulation. These cells are committed to a cardiac fate by a combination of signaling pathways, such as Wnt, FGF, and BMP (Gilbert, 2006).

A previous project in our lab used microarray technology to measure the expression of genes in the heart fields of chicken embryos. Embryos were treated with small molecule inhibitors that block FGF signaling, BMP signaling, or both. Based on the microarray data, all genes in the chicken genome were assigned a "regulatory bin" describing how they respond to FGF and BMP. The regulatory bins are categorized according to the signaling pathways that affect them (FGF, BMP, or a Cooperative effect), and how they are regulated (upregulated, downregulated, or unaffected). Based on this criteria, 27 regulatory bins were created, each of which contained from a few to hundreds of genes (Konieczka, 2012).

The heart fields in the developing embryo have been experimentally divided into two regions called the first and second heart fields, though in fact they are contiguous. The primary heart field resides more laterally and forms the primitive heart tube and left
ventricle. The secondary heart field represents the medial region of the overall heart field and differentiates into dorsal mesocardium. The dorsal mesocardium differentiates slowly over several days and forms the right ventricle and the outflow and inflow tracts (Gilbert, 2006). BMP2 is expressed in the (lateral) primary heart field mesoderm and endoderm, while FGF8 is expressed in the pharyngeal endoderm immediately adjacent to the secondary heart field. A central region shows overlapping expression of FGF8 and BMP2. Studies in our lab mapped the expression of selected genes from the FGF-upregulated (0,1,0) and BMP-upregulated (1,0,0) bins, and found that they correlate with the geographic expression of FGF8 and BMP2 (Konieczka, 2012).

The work described in this chapter investigates how BMP and FGF signaling regulate gene expression in the primary and secondary heart fields. These studies analyzed the composition of the FGF and BMP regulatory bins, specifically with regard to transcriptional elements. This project is based around the question: "What characterizes the promoter regions of genes in a particular regulatory bin, making them different from those under a different regulatory program?" To accomplish this, I used the following strategy: (1) Find novel and previously-known transcription factor binding sites in the chicken genome, calculating which are over-represented in each regulatory bin. (2) For the FGF- and BMP- upregulated bins, locate regions of promoter sequence that contain notable "clusters" of over-represented transcription factor binding sites. (3) Test the functionality of binding site clusters using an enhancer-reporter system transfected into embryonic cardiac myocytes in culture.
3.3: MATERIALS AND METHODS

3.3.1: Assembling "Promoterome" Files

Transcription factor binding sites generally appear in close proximity to the transcription start site of the gene they regulate. However, binding sites can also appear in introns. For purposes of this study, we define "promoter" as 10 kilobases 5' of the transcription start site, largely to obtain a consistent sequence length for later steps of analysis. Promoter sequence was created in FASTA format using a custom-built Perl-script called "promqueen.pl", which takes a list of Genbank ID numbers as input, calculates the positions of sequence 10kb upstream of the transcription start site, and downloads these sequences using the E-Utils service for the NCBI database maintained by the National Library of Medicine (Sayers, 2008).

All promoter sequences were masked by a local installation of Repeatmasker (Smit, 2010) to eliminate repetitive elements and low-complexity sequence. The final result is a FASTA file containing "masked" sequences for all chicken genes, with the letter 'N' substituting for any regions of sequence that should not be considered for further analysis. Collectively, this set of sequences is referred to as the "promoterome."
3.3.2: Identifying Over-Represented Motifs

DNA sequence motifs were identified using two strategies: (1) searching for previously known transcription factor binding sites, and (2) finding novel motifs without prior bias. For the first approach, a set of 1,365 Position-Specific Scoring Matrices (PSSMs) was obtained from the TRANSFAC database (Wingender et al., 1996). Of this set, 1,062 "high quality" matrices were selected to avoid trivial or non-reproducible results.

A background model of nucleotide base distribution was created, and the promoterome was scanned for the presence of each of the 1,062 "high quality" transcription factor binding sites, using a Perl-script ('ScanGenomeAndPrint.pl') running on the University of Arizona High Performance Computing systems. Briefly, the scanning algorithm reports the highest "score" for each matrix in each sequence, based on a system of assigning positive values to nucleotide matches and negative values to mismatches. The 20 highest positive scores for each PSSM are averaged, and any sequence with a positive score that is at least 70% of this "Top 20 Average" is considered to contain that particular binding site.

A hypergeometric distribution (Rice, 2007) determined whether any given transcription factor binding site is over-represented in a regulatory bin, using custom-built statistical analysis software written for the R platform ('hyperbin.R'). The probabilities returned by the hyper geometric test for each of the binding sites were adjusted for multiple
hypothesis testing using the Benjamini-Yekutieli method (Benjamini and Yekutieli, 2001).

To identify novel motifs without prior bias, Amadeus software running within the Allegro package was used (Halperin et al., 2009; Linhart et al., 2008). Masked promoterome sequence was inputted along with lists of gene IDs describing each regulatory bin, with motifs identified based on enrichment. For each regulatory bin, Allegro was run 20 times, combining 10 different promoter sizes (1kb-10kb, in 1kb increments) and 2 different motif sizes (10 or 12).

3.3.3: Condensing and Clustering Motifs

The "BLiC" algorithm developed by Naomi Habib (Habib et al., 2008) eliminated redundancies among over-represented motifs found in previous steps. BLiC compares and merges motifs in a recursive manner, combining motifs that are identical, highly similar, or overlapping. Since Amadeus found over-represented motifs using 20 different sets of parameters for each regulatory bin, we expect that this would generate a large amount of redundancy that could be eliminated using BLiC. Conversely, any motifs found by Amadeus that could not be merged were considered irreproducible results (since no similar motif had been found with alternate parameters) and discarded.

The final, refined set of over-represented motifs was mapped onto the original promoters in each regulatory bin using a local installation of MAST (Motif and Alignment Search...
Tool) (Bailey and Gribskov, 1998), with a Perl-script ('MAST_batch.pl') to reformat the data and batch-process it. "Clusters" of motifs were determined using Cluster Buster, which locates regions of binding sites in close proximity that can not be accounted for by a neutral model of nucleotide distribution (Frith et al., 2003). A local installation of Cluster Buster searched for clusters using three different "gap parameters" (average expected distance between binding sites): 20, 35, and 50. The outputs using these three parameters were parsed by a Perl script ('CBUST_analysis.pl') and statistical analysis software written in R ('motif_pairwise.R'), creating two outputs: (1) a non-redundant list of clusters using the longest sequence produced by multiple gap parameters. (2) A list of binding sites in each cluster, defined as those sites that contribute 5% or more to the total Cluster Score. A summary of these computational methods can be seen in Figure 7.

### 3.3.4: Primary Cell Culture

Myocyte and fibroblast cell cultures were prepared according to standard protocols from chicken embryos grown for six days in a warm, humid environment. Embryos were removed from eggs, and hearts were dissected into Hanks Balanced Salt Solution (Invitrogen). The hearts were dissociated by rinsing 4 times in a mixture of Hanks Salt Solution and Trypsin (0.6 mg/mL), with 5-minute incubations at 37º C between each rinse. Myocytes were resuspended in CCM (Invitrogen or Cellgro MEM with 5% inactive FBS, 50 Units/mL Penicillin, and 50 Units/mL Streptomycin) and filtered through a 70 µm nylon cell strainer to remove clumps. A "pre-plating" incubation allowed 30 minutes for the fibroblasts to adhere to the plate. After this incubation, the
Figure 7. Work-flow diagram for identifying transcriptional modules in similarly-regulated genes. Predicting over-represented transcriptional elements involves several steps: (1) Creating a "promoterome" representing all 5' sequences for every gene in the model system. (2) Determining which promoters contain each transcription factor binding site or novel motif. (3) Calculating which binding sites are over-represented among regulatory bins. (4) Refining the over-represented motifs into a final, non-redundant set. (5) Determining where clusters of over-represented motifs occur in promoter regions of similarly-regulated genes. Partially adapted from Habib, 2008.
suspension (now containing mostly myocytes) was removed, leaving behind a flask of fibroblasts. After this separation, cells in suspension were counted using a standard hemocytometer, and then brought to a concentration of 500,000 cells/mL in CMM with Glutamine added. 250,000 cells were added to each well of a 24-well plate, and allowed to grow for 3-4 days.

3.3.5: Transfections and Luciferase Assays

Sequences containing significant clusters of transcription factor binding sites (as described in section 3.3.3) were amplified by PCR from chicken genomic DNA using Platinum Taq Polymerase (Invitrogen) according to standard protocols. Primers used for PCR contained restriction sites for EcoRI or HindIII on the forward primer, and KpnI or BamHI on the reverse primer. PCR amplicons were digested with the appropriate enzymes overnight, purified, and ligated into the Multiple Cloning Site of pCluc Mini-TK2 (New England Biolabs), an enhancer-reporter luciferase vector containing a minimal TK promoter.

Experimental enhancer vectors were transfected into primary myocyte or fibroblast culture using Effectene (Invitrogen). Each well on a 24-well cell culture plate was transfected with 500 ng of pCLuc Mini-TK2 construct per well, 400 ng of pCMV-Gluc control plasmid (New England Biolabs), 3.125 µg of Enhancer, and 5 µg of Effectene. Transfection reactions were prepared in "Buffer EC", according to the standard Effectene
protocol. Reactions involving co-transfection with a dominant negative or constitutively-active FGF receptor used 250 ng of FGF receptor plasmid, 250 ng of pCLuc Mini-TK2 construct, and 200 ng pCMV-GLuc. As these luciferase vectors code for a secreted luciferase, samples were taken from the cell culture media without the need for cell collection or lysis, collected from each experimental well at 24 hours after transfection.

Luciferase measurements were obtained with a Turner Designs TD-20/20 Luminometer, using the Biolux Cypridina and Gaussia luciferase assay kits (New England Biolabs). 20 µL of media from each experimental well was mixed with 50 µL of Cypridinia luciferase substrate to allow measurement of luciferase activity. The same method was repeated using 50 µL of Gaussia luciferase substrate. All experimental conditions were run in triplicate, normalized using Gaussia luciferase measurements, and compared to luciferase activity using the unmodified pCLuc mini-TK2.
3.4: RESULTS

3.4.1: Over-represented Motifs

Of the 27 regulatory bins concerning gene response to FGF, BMP, or Cooperative signaling, we disregarded the unregulated bin (0,0,0) along with any bin that contained less than 50 genes. This cut-off was determined based on the distribution of bin sizes (Figure 6), as small collections of genes are deemed less likely to represent a prominent trend in signaling regulation, and more likely to yield false positives in a search for over-represented motifs. Fifteen regulatory bins besides the unregulated set contained more than 50 genes: FGF upregulated (0,1,0); BMP upregulated (1,0,0); FGF and BMP upregulated (1,1,0); FGF and BMP upregulated, Cooperatively downregulated (1,1,-1); FGF upregulated, Cooperatively downregulated (0,1,-1); BMP upregulated, FGF downregulated (1,-1,0); BMP upregulated, Cooperatively downregulated (1,0,-1); BMP downregulated (0,-1,0), BMP downregulated, Cooperatively upregulated (-1,0,1); BMP downregulated, FGF downregulated (-1,-1,0); FGF downregulated (0,-1,0); FGF downregulated, Cooperative upregulated (0,-1,1); Cooperative upregulated (0,0,1); and Cooperatively downregulated (0,0,-1).

All 15 of the above regulatory bins were tested for statistically over-represented transcription factor binding sites. Three bins did not contain any over-represented known binding sites (1,1,0; 1,-1,0; -1,-1,0). It is notable that these are the three smallest bins, with two of them containing only slightly more than the requisite 50 genes. This
reinforces our notion that a large number of genes are required to find statistically over-represented transcriptional elements. The remaining 12 regulatory bins contained at least three previously known over-represented binding sites, with a median of 21.5 sites per bin. The bin containing the largest number of over-represented sites is BMP-upregulated-Cooperatively Downregulated (1,0,-1), which contains 554 genes with 61 over-represented motifs. However, some of these motifs proved to be redundant, essentially different PSSMs for the same transcription factor.

As seen in Table 1, the BMP-upregulated bin (1,0,0) is characterized by 29 previously known transcription factor binding sites (from the TRANSFAC database), 7 motifs produced from a merging of multiple known sites or a known site and a novel site, and 30 completely novel sites. A few notable motifs in the BMP bin include Nkx2-5, a transcription factor required for heart development (Komuro and Izumo, 1993), and Mef-2, a myocyte-specific factor involved in muscle differentiation (Gossett et al., 1989). The FGF-upregulated bin (0,1,0) is characterized by 11 previously known binding sites, 2 merged ones, and 27 completely novel motifs. Some notable motifs that characterize the FGF bin include NFY_C, a CCAAT box promoter element that is bound by Nuclear Transcription Factor Y (Dolfini et al., 2012), and CdxA, the binding site for Caudal Type Homeobox 1, expressed at the primitive streak and in the lateral mesoderm (Frumkin et al., 1991).
3.4.2: Clusters of Motifs

Having found many transcription factor binding sites that characterize each regulatory bin, we sought to assemble "transcriptional modules", combinations of binding sites that form functional enhancer elements. Cluster Buster (Frith et al., 2003) identified clusters of the motifs chosen as described above. A critical parameter in cluster assembly is the "gap", or the average distance in base pairs between motifs. In order to cover all likely possibilities, we observed the results of clusters assembled using a variety of gap values: 20 bp, 35 bp, 50 bp, 100 bp, and 200 bp (the default recommended gap value is 35). We determined that gap values of 20 bp, 35 bp, and 50 bp all produced clusters of reasonable length and content, and produced similar results. The larger gap values (100 bp and 200 bp) tended to consider the entire promoter region a "cluster." Thus, we combined the clusters calculated using the three smaller gap values, taking the largest version of each cluster in cases of redundancy (such as the common occurrence of all three gap parameters finding clusters in the same region of promoter sequence). After such manipulations, the BMP-positive regulatory bin contained 51 non-redundant clusters of binding sites, with a log-likelihood score of 5 or better. The FGF-positive bin had fewer results, with 23 clusters meeting the same criteria.

The cluster data was checked to ensure that we would not obtain similar, trivial results with any random collection of DNA motifs. To test this idea, promoter sequences that yielded significant clusters in the BMP-positive set were queried for clusters using motifs over-represented in the FGF-positive bin. This did not yield any significant clusters, and
neither did the reverse experiment, suggesting that the clusters we obtained could not be created with an alternate set of motifs.

To determine which pairs of sites most often appeared together, a list of motifs contributing at least 5% to each Cluster Score was produced, and a cross-referencing algorithm calculated how many times each motif appeared in the same cluster with every other motif. The results of this analysis can be seen in Table 2. For clusters in the BMP-upregulated set, we see that the MATH1 motif is present in most clusters, often pairing with INO4, STAT3, Nkx2-5, and Stat6/HSF (which also are often seen with each other, since they are in the same clusters with MATH1). In the FGF-upregulated set, we see that many clusters involve Sp1 in collaboration with Wt1, Alfin, CdxA, CEBP, and the novel site "FGF.node7".
Table 1. Over-represented motifs in BMP upregulated and FGF upregulated bins. All DNA-sequence motifs statistically over-represented in either the BMP-upregulated (1,0,0) or FGF-upregulated (0,1,0) bins appear above, broken into three categories: (1) A motif that appears in the Transfac database and represents a known transcription factor binding site. (2) A motif that is produced by merging 2 known binding sites, or combining a known one with a novel site. (3) A novel site, listed as "node" per BliC notation.
<table>
<thead>
<tr>
<th>Rank</th>
<th>BMP upregulated motif pairs</th>
<th>FGF upregulated motif pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>INO4 MATH1</td>
<td>Sp1 WT1</td>
</tr>
<tr>
<td>2</td>
<td>STAT3 MATH1</td>
<td>Sp1 Alfin</td>
</tr>
<tr>
<td>3</td>
<td>Stat6/HSF MATH1</td>
<td>Sp1 CdxA</td>
</tr>
<tr>
<td>4</td>
<td>Nkx2-5 MATH1</td>
<td>Sp1 FGF.node7</td>
</tr>
<tr>
<td>5</td>
<td>MATH1 ARR1</td>
<td>Wt1 Alfin</td>
</tr>
<tr>
<td>6</td>
<td>DOF2/DOF3 Stat6/HSF</td>
<td>CEBP FGF.node7</td>
</tr>
<tr>
<td>7</td>
<td>INO4 STAT3</td>
<td>Sp1 CEBP</td>
</tr>
<tr>
<td>8</td>
<td>ERR2/SF1 MATH1</td>
<td>Sp1 ETF</td>
</tr>
<tr>
<td>9</td>
<td>DOF2/DOF3 MATH1</td>
<td>WT1 ETF</td>
</tr>
<tr>
<td>10</td>
<td>FOXM1 MATH1</td>
<td>WT1 FGF.node7</td>
</tr>
<tr>
<td>11</td>
<td>MET4 MATH1</td>
<td>FGF.node15 FGF.node7</td>
</tr>
<tr>
<td>12</td>
<td>Stat6/HSF INO4</td>
<td>BRCA (merged) CEBP</td>
</tr>
<tr>
<td>13</td>
<td>Tra-1 MATH1</td>
<td>BRCA (merged) FGF.node7</td>
</tr>
<tr>
<td>14</td>
<td>CKROX MATH1</td>
<td>NFY ETF</td>
</tr>
<tr>
<td>15</td>
<td>DOF2/DOF3 Nkx2-5</td>
<td>Sp1 FGF.node3</td>
</tr>
<tr>
<td>16</td>
<td>Stat6/HSF Nkx2-5</td>
<td>BRCA (merged) Sp1</td>
</tr>
<tr>
<td>17</td>
<td>MET4 Stat6/HSF</td>
<td>Sp1 NFY</td>
</tr>
<tr>
<td>18</td>
<td>MEF2 MATH1</td>
<td>Stat5A Alfin</td>
</tr>
<tr>
<td>19</td>
<td>INO4 Nkx2-5</td>
<td>Stat5A CEBP</td>
</tr>
<tr>
<td>20</td>
<td>BMP.node1 MATH1</td>
<td>Stat5A FGF.node7</td>
</tr>
</tbody>
</table>

Table 2. 20 Most Commonly Over-represented Transcription Factor Binding Site Pairs found in the Same Cluster. Based on 10 kb promoter regions from genes upregulated by either BMP or FGF signaling, a set of over-represented transcription factor binding sites was calculated, then further analyzed to identify pairs of sites in close proximity. The above table shows which pairs of binding sites most often are clustered together.
3.4.3: Testing Enhancer Elements for Functionality

In previous parts of this project, we found a number of transcription factor binding sites that were over-represented in the promoter regions of genes having certain regulatory responses, and identified clusters of these sites in close proximity. The functionality of these clusters was tested using cultured primary embryonic cardiac myocytes and fibroblasts. Enhancer-reporter constructs were generated, that contained various binding site clusters directly upstream of a minimal TK-2 promoter driving transcription of secreted *Cypridina* luciferase. The luciferase activity of each enhancer vector was compared to that of the parental vector containing only the minimal promoter.

Out of 51 clusters identified in the BMP-upregulated bin, 39 were successfully cloned and tested in myocytes and fibroblasts (Figure 8). In myocytes, 15 of the 39 clusters significantly enhanced transcription compared to the parental vector, 11 repressed transcription, and 13 had no effect on luciferase activity. The average level of upregulation in enhancers was about 2-7 fold, with a median value of 2.8. The strongest enhancement in myocytes came from sequence surrounding a cluster of MEF-2 and CDX-2 sites in the promoter of the gene Glud1. The same experiments were repeated in primary fibroblast cultures to determine whether these results are cell-type specific. All of the promoter regions that repressed transcription in primary myocyte experiments also repressed transcription in cardiac fibroblasts. However, promoter regions that increased transcription tended to be cell-type specific: 10 of the 15 clusters that produced enhancement in myocytes showed no enhancement in fibroblasts, one cluster (Glud1) had
weaker enhancement in fibroblasts, and three clusters showed approximately the same enhancement in both myocytes and fibroblasts.

Among clusters identified in the FGF-upregulated bin, 21 of the 23 were successfully cloned and tested (Figure 9). Nine enhanced luciferase activity, seven repressed expression, and five had no effect relative to the parental vector containing the TK minimal promoter alone. The transcriptional modules that increased luciferase activity generally did so by a much higher margin than functional enhancers in the BMP-upregulated bin. Clusters in the FGF-upregulated bin that produced enhancement increased expression 3-25 fold, with a median value of 6.6, compared to a median value of 2.8 in the BMP-upregulated bin. As with clusters in the BMP-upregulated bin, repressor activity was not cell-type-specific whereas enhancers were generally active in myocytes and not active or less so in fibroblasts. Based on these results, we concluded that our prediction algorithms successfully identified functional transcriptional modules with an acceptable rate of success.
Figure 8. Enhancer-reporter Activity in Transcription Factor Binding Site Clusters Found in Genes in the BMP-upregulated Bin. Normalized luciferase activity reported for proposed transcriptional modules from the promoter regions of various genes in the BMP-upregulated bin (1,0,0), with all values reported as a fold change over the parental vector TK2.
Figure 9. Enhancer-reporter Activity in Transcription Factor Binding Site Clusters Found in Genes in the FGF-upregulated Bin. Normalized luciferase activity reported for proposed transcriptional modules from the promoter regions of various genes in the FGF-upregulated bin (0,1,0), with all values reported as a fold change over the parental vector TK2.
3.4.4: Analysis of components in functional enhancers

Enhancer elements identified above varied in size from less than 100 bp to 3 kb. Given the large size of some of these enhancer elements, it is difficult to conclude which parts of the sequence confers increased transcriptional activity. In a very large sequence, the active element could potentially be an unrecognized binding site. Thus, to more precisely define elements conferring transcriptional enhancement, motif clusters were subdivided into smaller fragments according to the predicted location of transcription factor binding motifs. As shown in Figure 10, a ~600 bp sequence in the promoter for ING1 (inhibitor of growth) that conferred 12-13 fold increase in transcription was divided into four segments based on the location of transcription factor binding sites. The first 78 nucleotides, containing CdxA, Sp1, and various unknown motifs, produced 5-fold enhancement. The next two segments, one ~200 bp, and another ~300 bp, each provided the same enhancement as the full cluster, suggesting that elements within these segments are sufficient to confer the enhancement observed with the intact ING1 cluster. Each of these segments contain an NFY binding site, an element that is common in our high-preforming enhancers. A fourth segment, covering bases 1-279, likewise showed similar enhancement as the entire cluster.
Figure 10: Enhancer-Reporter Activity using fragments of the ING1 promoter. The full-length 600-bp cluster of transcription factor binding sites is mapped out according to the location of DNA motifs. Enhancer-reporter vectors with fragments of this cluster placed in front of a minimal promoter are normalized relative to the parental vector and compared to the full 600-bp transcriptional module. These results show that two separate sections of the module, both containing a NFY_C site, produce the same ~13-16 fold enhancement as the entire enhancer element. However, a fragment containing the first 78 base-pairs shows ~4.5 fold enhancement enhancement.
Figure 11: Enhancer-Reporter Activity using fragments of the HMGCR promoter.
Enhancer-reporter tests using the 37-bp cluster of transcription factor binding sites in the HMGCR promoter yielded a 5.8-fold enhancement over a minimal promoter. However, this enhancement depends on the presence of the entire cluster; neither half of the cluster is sufficient to provide any increase in transcription. Mutation of the NFY_C site in the full-length cluster also knocked out enhancer activity. Capital letters in sequence show the predicted HMGCR site in the cluster, with mutated bases in red.
Figure 12: Enhancer-Reporter Activity using fragments of the EOMES promoter. The full 152-bp cluster of transcription factor binding sites produces nearly 12-fold enhancement of luciferase activity in this experiment. However a subfragment containing only a cluster of novel and Sp1 motifs has no effect, while inclusion of a subfragment containing CdxA, Stat5a, and node8 (novel) sites causes a significant enhancement, albeit less than the full cluster. This same subfragment provides no enhancement if the CdxA site is mutated. Capital letters in sequence show the predicted HMGCR site in the cluster, with mutated bases in red.
A similar experiment tested subfragments of the HMGCR promoter region. A 37-bp cluster of transcription factor binding sites conferred nearly 6-fold transcriptional activity compared to a minimal promoter. However, neither half of the 37 bp cluster was sufficient for any enhancement whatsoever. Mutation of the predicted NFY binding site within the intact HMGCR enhancer abolished enhancer activity (Figure 11).

Analysis of a cluster from the EOMES promoter (Figure 12) revealed that while the entire enhancer provided a 12-fold increase in transcription, the first 60 nucleotides showed no enhancer activity. The 3’ most 60 nucleotides, containing binding sites for CdxA and Stat5a in addition to a novel motif, gave a 4-fold greater enhancement compared to the minimal TK promoter. This effect is smaller than that of the full enhancer, but still significant, suggesting that the CdxA/Stat5A sites contribute to transcription upregulation, but require the rest of the cluster for full effect. When the CdxA site was mutated in this same fragment, no enhancement was observed, further suggesting that the CdxA site is functional and necessary.

3.4.5: Enhancer response to FGF signaling.

The transcription factor binding sites that characterize each regulatory bin were identified because they are overrepresented in the promoters of genes controlled by certain combinations of FGF and BMP signaling. Experiments described above show that several clusters of sites can function as enhancer elements, but did not determine whether they are responsive to FGF or BMP signaling. To investigate FGF responsiveness of
enhancers identified in FGF regulatory bins, enhancer test constructs containing the EOMES, HMGCR, ELMOD, or ING1 promoter fragments were co-transfected into embryonic cardiac myocytes with dominant active or dominant negative versions of FGFR. As seen in Figure 13, luciferase activity is increased approximately 18 fold when the 152-bp EOMES promoter fragment is co transfected with the constitutively active FGFR, versus a significantly smaller 13-fold increase when co-transfected with a control vector (pBE), and a 10-fold enhancement in cells co-transfected with dominant negative FGFR. This result suggests that the EOMES enhancer element is FGF-responsive. A similar result is observed with a 600 bp fragment from the ING1 promoter, which causes an 8-9 fold increase over a minimal promoter when co-transfected with either a control vector or one expressing dominant negative FGFR, and a much stronger 16-fold enhancement when co-transfected with a constitutively active FGFR. A 300 bp cluster containing NFY sites in the ELMOD promoter revealed a similar effect, with the constitutively active FGFR doubling the strength of the enhancer element, while a 37 bp cluster from the HMGCR promoter showed a small but statistically significant response to FGF signaling.

Finally, we further analyzed the ING1 promoter by co-transfecting various smaller fragments of the ING1 cluster with a constitutively active or dominant-negative FGF receptor. At time of this writing, this data is in need of further refinement, as

1 It may be noted that other experiments showed a larger effect of the ING1 cluster. This discrepancy is likely the result of different transfection conditions as well as small differences in the purity of myocyte cell culture. All values in Figure 13 are under the similar conditions, and thus should be compared to each other.
experimental error prevents statistical significance. However, I suspect that the middle fragment of the ING1 cluster, from bases 78-279, is responsible for FGF-responsiveness. This fragment resulted in ~15-fold enhancement in response to either a control vector or dominant negative FGF receptor, while co-transfection with constitutively-active FGF showed a ~31-fold enhancement, albeit with large experimental error.
Figure 13. Enhancer element response to FGF signaling. Values are determined by an enhancer-reporter luciferase assay, normalized relative to a minimal TK2 promoter vector transfected under the same conditions. We observe that co-transfection with a constitutively active FGF receptor causes an increase in enhancer activity in promoter fragments of EOMES, ING1, HMGCR, and ELMOD. Asterisks indicate a statistically significant difference from the Control (pBE) values.
Figure 14. Fragments of the ING1 promoter, Co-transfected with FGFR. This preliminary data shows the enhancer activity of fragments of ING1 (discussed in Figure 10) in the presence of a constitutively active or dominant negative FGFR. The "middle" fragment (78-279 nt) has a higher average response with constitutively active FGFR, though unusually high experimental error prevents us from assigning statistical significance.
3.5: CONCLUSION

In this chapter, I have outlined methods for deciphering how large groups of genes can share similar regulatory patterns. Building off of Jay Koneiska's concept of "regulatory bins," I calculated over-represented transcription factor binding sites and novel DNA motifs that characterize each bin. In addition, these over-represented sequences were assembled into clusters in the promoter regions of the genes in each regulatory bin, creating transcriptional modules. Finally, these modules were tested in primary myocyte culture to determine if they have an effect on transcription, and if they are affected by FGF signaling.

The over-represented transcription factor binding sites are likely to be responsible for any particular gene response. Meanwhile, experiments testing individual clusters provide more specific view. Only a fraction of genes in the FGF- and BMP-upregulated bins contained significant clusters at all. Many of the rest of the genes still contain certain bin-specific transcription factor binding sites that may explain their regulatory patterns, and it is possible that there are important combinations of binding sites that act over longer distances. Most of the clusters proved to be functional enhancer or repressor elements, independent of their context in the original gene's promoter. This speaks to the strength of our computational methods in identifying functional transcriptional elements. In addition, enhancers in the FGF-upregulated bin showed that they were responsive to FGF signaling. These examples show that our search has definitively located and characterized fragments of the genes' promoter regions that cause known regulatory responses.
In the FGF-upregulated bin, binding sites for NFY and CdxA characterize a positive response to FGF signaling. This was first determined statistically, by simply observing that these sites are over-represented in the promoter regions of genes upregulated by FGF signaling. Many of the strongest enhancers among the clusters contained these two sites. These sites were tested directly by mutating or excluding them from the transcriptional modules they were contained in, and it was found that all enhancement disappeared in their absence.

In the FGF-positive regulatory bin (581 genes), 86 genes contain at least one NFY site in the promoter and 33 genes contain at least one CdxA site. The NFY site is bound by Nuclear Transcription Factor Y (NF-Y), a complex of three proteins that are required for cell-cycle activation. NF-Y sites are known to be partnered with GC box factors (such as Sp1, another over-represented motif in the FGF bin), E2Fs, and transcription factors associated with stress response (Dolfini et al., 2012). The CdxA site is bound by Caudal Type Homeobox 1 (Cdx1), a transcription factor expressed at the primitive streak and in the lateral mesoderm (Frumkin et al., 1991). Cdx genes are involved in anteroposterior patterning (Young and Deschamps, 2009), and are required for hematopoietic development (Lengerke and Daley, 2012).

In conclusion, we have provided a broad view of transcriptional elements that explain the varying patterns of genes that respond to FGF and/or BMP signaling during heart field
development, and a more specific explanation for several FGF-responsive genes. This work explores the genetic "hardwiring" that allows cells to differentiate at the right time and place in embryonic development. In addition, we have laid out a pipeline, combining a variety of different computational methods, algorithms, and data sources: E-Utils, Allegro, Transfac, BLiC, Cluster Buster, and a set of custom-built analysis and data-curating programs written in Perl and R. Together, these methods enable a sophisticated method for discovering testable transcriptional modules.
3.6: FUTURE DIRECTIONS

Our studies of FGF regulated gene promoters have found transcription modules containing CdxA and NFY sites that are strong FGF-responsive enhancers. While we have tested nearly every cluster in the FGF-upregulated bin containing CdxA and NFY sites, there remain many other genes that contain CdxA and/or NFY sites, but do not form significant clusters. It is possible that these other regions containing CdxA and NFY sites are functional enhancers, a possibility that could be tested in the same way as the clusters were tested. To further examine whether these sites are actually bound by their predicted transcription factors (Cdx1 and NF-Y, respectively), Chromatin Immunoprecipitation (ChIP) assays could be used.

These promoter elements were studied in cell culture, and an obvious next step would be to test constructs in vivo. For an in vivo test, the same promoter enhancer elements could be linked to a GFP reporter and electroporated into the heart fields of G. gallus embryos. We could then observe (1) whether embryos electroporated with enhancer-containing vectors show stronger GFP expression than those with a control GFP vector, and (2) whether increases in GFP signal are observed in the second versus first heart fields.

In this study, detailed analysis of individual enhancers focused on those identified in the FGF-regulated bin. Future experiments could focus on enhancers identified in the BMP regulated genes. In addition, cluster analysis and enhancer tests could be preformed on
genes in any of the other 10 regulatory bins that contained over-represented transcription factor binding sites.

A more direct biological approach could determine which of the FGF- and BMP-regulated genes are actually necessary for heart field formation. Force-expression (with electroporation of an expression vector) or knock-down (using RNAi) would allow us to observe the effects of misregulation among individual genes. A promising candidate for such an experiment is Eomesodermin, which is known to induce Mesp1 expression and cardiac cell differentiation (van den Ameele et al., 2012). This would provide a concrete functional link between the genes in the regulatory bins and their biological importance in directing the formation of heart fields.

Finally, the pipeline created for this project is based on general statistical and biological principles, and thus could be applied to other areas of study in embryonic development. Our lab has also examined genes regulated by FGF signaling at the primitive streak during gastrulation (Hardy et al., 2011). The same methods used to profile transcriptional responses within the heart field could be applied to changes in gene expression during gastrulation, either independently or by cross-referencing the set of genes upregulated during gastrulation with those upregulated in the heart fields. Our computational methods are not limited to development, and thus could be applied to a wide variety of projects, such as changes in gene expression in certain types of cancers.
Throughout the previous two chapters, I described two separate projects that comprised my dissertation work in Parker Antin's lab. Though dealing with separate events in chicken embryogenesis, both projects attempt to elucidate the genetic interactions that cause cells to migrate and differentiate in the complex patterns required to form a functional organism. One could think of these processes as a computer program, in which the inputs (signals) and outputs (cell movements and transformations) are known, but the underlying program is only partially understood. Through this work, I present new information detailing the complex "circuitry" that controls cell behavior in response to a similar input: signaling from the Fibroblast Growth Factor (FGF).

In the first project, it was shown that FGF regulates the abundance of microRNAs at multiple steps, including transcription and selective blockade of RNA processing. Six microRNAs were drastically upregulated in pre-ingression epiblast when FGF signaling was inhibited. These six microRNAs targeted serine-tyrosine kinase receptors including Platelet-Derived Growth Factor (PDGF), a required signaling pathway for cell migration during gastrulation. Thus, there appears to be a hierarchy of signaling pathways involved in gastrulation, with FGF required to enable PDGF.

In the second project, computational analysis determined which cis-acting elements distinguish FGF-responsive genes from those regulated by BMP, a combination of FGF
and BMP, and unregulated genes. Both novel DNA motifs and established transcription factor binding sites were found to characterize co-regulated genes. A satisfying result was that clusters of predicted transcription factor binding sites were usually functional in cardiac myocytes, and a selection of the most functional enhancers was consistently responsive to manipulations in FGF signaling. This showed that our computationally determined promoter elements could cause the gene regulation observed in the heart fields.

Both of these projects illustrate the versatility of the chicken embryo as a model system, as it proved suitable for all the drug treatments, cell cultures, and genetic manipulations that were required to fully explore the experimental questions. However, work in the chicken system was also hindered by a lack of available antibodies, which prevented protein detection.

Throughout this dissertation, I have referred to both computational methods as well as biological lab work, as interaction between both methods was required. Some of this interaction was traditional, such as the use of computer algorithms to analyze biological data from microarrays. A more ambitious use of bioinformatics required predictions based on sequence data, expression patterns, and groups of co-regulated genes. Here, we see the synergy between biology and computer science: biological data is used to form computational predictions, which leads to new hypotheses that can tested in further experiments. New technologies increase the amount of information that can be obtained
from experiments, and increased knowledge allows more sophisticated computational predictions. One can only expect that the field of biology will rely more and more on bioinformatics in the future.
APPENDIX A: FIBROBLAST GROWTH FACTOR (FGF) SIGNALING NEGATIVELY REGULATES THE ABUNDANCE OF MICRORNAS THAT REGULATE PROTEINS REQUIRED FOR CELL MIGRATION AND EMBRYO PATTERNING DURING GASTRULATION
FGF signaling during gastrulation negatively modulates the abundance of microRNAs that regulate proteins required for cell migration and embryo patterning *

Alexander S. Bobbs¹, Aleksi V. Saarela¹, Tatiana A. Yatskievych², and Parker B. Antin¹²

¹From the Department of Molecular and Cellular Biology, University of Arizona, Tucson, Arizona, 85724 USA
²Department of Cellular and Molecular Medicine, University of Arizona, Tucson, Arizona, 85724 USA

*Running title: FGF Signaling Regulates MicroRNA Abundance

To whom correspondence should be addressed: Dr. Parker B. Antin, Department of Cellular and Molecular Medicine, University of Arizona, PO Box 245217, 1656 E. Mabel St., Tucson, AZ 85724-5217 USA, Tel: (520) 626-6382; Fax: (520) 626-7600; E-mail: pba@email.arizona.edu

Keywords: Gastrulation, Chicken Embryo, FGF Signaling, Lin28B, MicroRNAs

Background: FGF signaling is a critical pathway regulating cell migration during gastrulation.

Result: Inhibition of FGF signaling upregulates six microRNAs, partially through loss of LIN28B. FGF-regulated microRNAs target proteins important for gastrulation, including serine/threonine and tyrosine kinase receptors.

Conclusion: FGF signaling regulates microRNA abundance to control gastrulation.

Significance: This identifies a new mechanism by which FGF signaling regulates gene expression.

SUMMARY

FGF Signaling plays a pivotal role in regulating cell movements and lineage induction during gastrulation. Here we identify 44 microRNAs that are expressed in the primitive streak region of gastrula stage chicken embryos. We show that the primary effect of FGF signaling on microRNA abundance is to negatively regulate the levels of miRs -let-7b, -9, -19b, -107, -130b, and -218. LIN28B inhibits microRNA processing and is positively regulated by FGF signaling. Gain and loss of function experiments show that LIN28B negatively regulates the expression of miRs-19b, -130b, and let-7b, while negative modulation of miRs -9, -107 and -218 appears to be independent of LIN28B function. Predicted mRNA targets of the FGF regulated microRNAs are over represented in serine/threonine and tyrosine kinase receptors, including ACVR1, ACVR2B, PDGFRα, TGFR1 and TGFR3. Luciferase assays show that these and other candidates are targeted by FGF-regulated microRNAs. PDGFRα, a receptor whose activity is required for cell migration through the primitive streak, is a target of miRs-130b and -218 in vivo. These results identify a novel mechanism by which FGF signaling regulates gene expression by negatively modulating microRNA abundance through both LIN28B dependent and LIN28B independent pathways.

The body plan of vertebrates arises during gastrulation, when through complex morphogenetic movements the three primary germ layers and the major body axes are formed. A key feature of gastrulation in amniotes is the ingestion of cells from the epithelial epiblast through the primitive streak to form the mesoderm and the endoderm cell layers (1). Cells leaving the epiblast undergo an epithelial-to-mesenchymal transition (EMT), in which the epithelial phenotype is downregulated and a migratory phenotype is activated (2). Mesoderm cells initially maintain a migratory phenotype, while cells comprising the endoderm re-establish cell-cell junctions and form a contiguous layer at the base of the primitive streak. Transcriptional profiling has shown that gastrulation and germ layer formation involve the upregulation and downregulation of several
pluripotency
Lin28

specific aspects of EMT in culture
investigating
MET in vivo remains enigmatic.

miRs
though the

stem
general,

processing
while Lin28B blocks

the let

negatively regulate
signaling.

primitive streak region of avian embryos
thousand genes (3,4). Genes upregulated in the
primitive streak region of avian embryos include
members of the FGF, NOTCH, PDGF, EPH and
canonical and non canonical WNT pathways,
negative pathway modulators, and a large number
of transcription factors. The FGF, PDGF and
WNT pathways have been shown to regulate
primitive streak formation and/or cell migration
through the streak (5-8). FGF signaling is required
for expression of members from all of these
pathways, suggesting that it lies at the top of a
hierarchy of signaling cascades.

MicroRNAs (miRs) regulate crucial though
still relatively undefined processes during early
embryogenesis. Their importance is underscored
by the early and severe phenotypes observed when
miR processing is inactivated. In sea urchin, loss of
Drosha or Dicer blocks gastrulation and greatly
reduces expression of endoderm and mesoderm-
specific genes (9). These defects were rescued by
introduction of four miRs. Dicer and DGRC8 null
mouse embryos are reduced in size, fail to
gastrulate or express mesoderm markers, and
exhibit a defect in development of ES cells that
form the inner mass (10,11). Ablation of Dicer in
zebrafish leads to cell movement defects during
gastrulation and abnormal heart and neural
development (12). miRs play an important role in
regulating pluripotency in embryonic stem cells
(13), however this function alone is unlikely to
account for the embryonic phenotypes observed
when dicer or DGRC8 are ablated.

Individual miRs have been shown to regulate
specific aspects of EMT in cultured cells (14-18),
though the potential function of these and other
miRs in regulating EMT and the reverse process
MET in vivo remains enigmatic. While
investigating gene regulation during avian
gastrulation (4), we observed that LIN28B is
highly downregulated following inhibition of FGF
signaling. LIN28B and its paralog LIN28A
negatively regulate miR processing, in particular
the let-7 family of miRs (19,20). Lin28A inhibits
processing of pre-let-7 by Dicer in the cytoplasm,
while Lin28B blocks the earlier step of pri-let-7
processing in the nucleus by the Drosha-
containing Microprocessor complex (21). In
general, studies in cancer cell lines and cultured
stem cells have shown that repression of let-7 by
Lin28A and/or Lin28B is associated with
pluripotency and tumorigenesis (22,23), while
upregulation of miR-let-7 is associated with cell
differentiation and reduced cell proliferation (13).
Almost all of these findings have arisen from
tissue culture studies of stem cells or immortalized
cell lines. Little information is available about the
functions of Lin28 and its relationship to miR
processing during early embryogenesis. Here we
investigate the relationship between FGF
signaling, LIN28B expression, and miR function
during avian gastrulation.

EXPERIMENTAL PROCEDURES

Chicken Embryo culture and manipulation -
Fertile chicken (Gallus gallus) eggs from Hy-Line
International (Spencer, IA) were incubated at 37°C
with high humidity until the embryos reached
Hamburger-Hamilton (HH) stage 4 (24,25). For
studies of FGF-signaling inhibition, embryos
placed in New Culture (4,26) were exposed to
100µM SU5402 (Pfizer, New York, NY) or
DMSO carrier diluted in cell culture medium
supplemented with penicillin and streptomycin.
For studies of Lin28B knockdown or forced-
expression, embryo electroporations were
performed as previously described (4,7) using a
plasmid containing the entire coding region cDNA
of LIN28B plus a 3’ terminal FLAG epitope tag
under control of the chicken beta-actin promoter
(pBE-Lin28B), or morpholinos
(AAGACAGAAAAAGGTTACTCTTGGT)
antisense to the 3’ exon-intron boundary of
LIN28B exon 2, or the Standard Control
Morfolino (GeneTools, Philomath, OR).
Embryos were incubated for 5 hours at 37°C.
Embryos analyzed by in situ hybridization (ISH)
were fixed in 4% paraformaldehyde in PBS (PFA),
while those used for real-time PCR analysis were
dissected into lysis buffer (Life Technologies,
Grand Island, NY).

Real-time PCR Analysis - RNA was isolated
from embryos using the mirVana miR Isolation
Kit (Ambion). First strand cDNA was prepared
using the Ncode miR First-Strand cDNA Synthesis
kit (Invitrogen). To generate cDNA suitable for
measuring miR precursors, a custom primer
corresponding to the 3’ end of the pre-miR
sequence was used in the Reverse Transcriptase
reaction in lieu of random hexamers. Real-time
PCR analyses were carried out with Maxima
SYBR Green qPCR Master Mix
(Fermentas/Thermo Scientific, Glen Burnie, MD)
in a Rotorgene RG6000, using standard protocols and the Rotorgene statistical analysis software. Mature miRs were amplified using primers corresponding to the exact sequence of the miRs, along with the "Universal" primer supplied with the Ncode kit (Invitrogen). miR precursors were amplified using primers corresponding to their stem-loop structures, while primary miR transcripts were amplified using the 3′ primer of the stem-loop structure and a 5′ primer positioned approximately 100 bases upstream of the stem-loop structure.

All real-time PCR assays were performed in triplicate and included a no-template control. To normalize the data, each real-time experiment also included measurement of a control gene; ribosomal protein RPL4 for measuring primary miRs and pre-miRs, or Beta-Galactosidase (β-gal) for cDNA prepared from transfected HeLa cells.

**MicroArray Analysis** - Samples of RNA isolated from the primitive streak region of HH stage 4 chicken embryos were mixed with spike-in controls and fluorescently labeled with Hy3 and Hy5 dyes using the miRCURY LNA microRNA Power Labeling Kit (Exiqon, Woburn, MA). Samples were hybridized according to the protocols of the miRCURY microRNA Array Kit, using Multispecies Array based on miRBase version 11. Hybridized slides were processed and scanned according to standard protocols. Each slide contained four array replicates, which were compared against negative probe controls using an unpaired Student's t-test and adjusted for multiple hypothesis testing using custom-built software written in the R statistical computing language.

**Whole-mount ISH** - Control or experimentally manipulated embryos were fixed in 4% PFA overnight at 4°C. ISH were performed as described (28) with some modifications (29). Detection of miR transcripts was accomplished using LNA probes (Exiqon) labeled with Digoxigenin at the 5′ and 3′ ends. ISH using LNA probes was performed as described (30).

**HeLa Cell Culture and Luciferase Assays** - HeLa cells were grown and passaged in DMEM media (Invitrogen) plus 10% Fetal Bovine Serum, supplemented with penicillin and streptomycin. Luciferase reporter constructs were generated from the vector pmiR-Report (Ambion). Constructs contained either the full-length 3′ UTR from the candidate target mRNA, or a 50-60 bp miR recognition element (MRE). 12-well culture plates of HeLa Cells were grown to 70-80% confluence in 500 ul of antibiotic-free media, and then transfected using standard protocols with 1ug of luciferase construct and 0.5 ug of a β-gal expression vector per well, using either Lipofectamine 2000 (Life Technologies) or TransPass transfection reagent (New England Biolabs, Ipswitch, MA). Cells were harvested 48 hours after transfection into Tropix Lysis Buffer (Life Technologies). Luciferase and β-gal activities were measured using the Dual Light Reporter Gene Assay System (Applied Biosystems). All luciferase assays were run in triplicate, normalized relative to β-gal, and compared to transfections using the parental vector (unmodified pmiR-Report). For experiments using miR mimics (Dharmacon), 100nM of either a G. gallus miR mimic or a C. elegans miR-167 control mimic, were used. The target protector miR-130b morpholino (GeneTools) was transfected at 5μM. Cells were harvested 30 hours after transfection, according to Yekta et al. (31).

**Computational Analysis of miR Targets** - Lists of evolutionarily conserved miR targets were obtained using a local installation of TargetScan (32), and filtered against mRNA expression data (4) to identify predicted miR mRNA targets expressed in the primitive streak.

Gene Ontology (GO) terms for G. gallus proteins were obtained from NCBI. Over represented GO terms among the targets of each FGF-regulated miR (let-7b, miR-9, miR-19b, miR-107, miR-130b, and miR-218) were identified using a chi-squared test, comparing GO term frequency in the target set versus the frequency of these terms across all proteins, adjusted for multiple hypothesis testing.

**RESULTS**

The relationship between FGF signaling and microRNA (miR) function, miR expression in the primitive streak was assessed using a multi-species miR microarray (33). Of the 44 chicken miRs detected significantly above background on the microarray (p<.05) after adjustments for multiple
hypothesis testing), the expression of 40 miRs was confirmed by real-time PCR analyses (Supplemental Table 1). Hybridization signal was detected for four miRs not previously identified in chicken. Expression of the minor (star) sequences for miRs-18a and -193b was verified by PCR. The miR-193b strand we detected is designated the major strand in other species. However, on miRBase this is shown as the minor strand. Expression of sequences corresponding to chicken miRs-let7e and -1030i was confirmed by PCR. Since these genes were not annotated on the chicken genome, BLAST analysis was used to identify potential gene locations. A putative miR-let7e gene was identified on chromosome 26, and a putative miR-1030i gene was identified on chromosome 1. These locations are syntenic with their locations in other species. Secondary structure analysis (34) showed that these regions can form stem-loop structures characteristic of miR genes that would permit the processing of their primary transcripts into mature microRNAs (data not shown).

To determine whether FGF signaling influences miR abundance in the primitive streak, miR expression levels were compared in RNA samples isolated from the primitive streak regions of control embryos versus embryos exposed for five hours to the FGFR inhibitor SU5402 (4,35). The abundance of six miRs was significantly increased following inhibition of FGF signaling, while only two miRs showed modest downregulation (Figure 1). miR let-7b levels increased almost 50-fold in response to FGF signaling inhibition, while the levels of miRs -9, -19b, -107, -130b, and -218 increased 6.5 to 20-fold. miR-let-7e, identified in our microarray screen, was expressed at very low levels that were unchanged by SU5402 treatment. Other let-7 family miRs were expressed at undetectable levels in the gastrula stage chicken embryo (data not shown). ISH detection of miR expression showed that miRs-let-7b, -9, -19b, -130b and -218 were expressed in the epiblast and primitive streak, but at reduced levels in the mesoderm and endoderm (Figure 1B). miR-107 was not consistently detected by ISH.

miR abundance is regulated by FGF signaling at multiple processing steps - miR abundance can be regulated transcriptionally and at multiple post transcriptional processing steps (36). To determine how FGF signaling influences miR expression, the relative abundances of primary (pri-miRs), hairpin loop intermediates (pre-miRs), and mature miR transcripts were assessed by real-time RT-PCR using primer pairs designed to amplify each type of miR transcript (Figure 2A). Comparing RNA samples from primitive streak regions of control versus SU5402 treated embryos, pri-miRs -9, let-7b, -107 and -218 were increased following inhibition of FGF signaling, with reduced or no change in pre-miR abundance (Figure 2B). In contrast, pri- and pre-miR levels for -19b were increased, while only the pre-miR levels of -130b were increased.

Previous studies have shown that Lin28B can negatively regulate processing of let-7 primary transcripts (21,37,38). Since LIN28B is downregulated following inhibition of FGF signaling in the gastrula stage chicken embryo [Figure 3A; (4)], gain and loss of function experiments were performed to determine whether FGF signaling acts through LIN28B to reduce miR-let-7b abundance. LIN28B knockdown was accomplished using a splice-blocking morpholino that deletes exon 2, introducing a frame-shift and a premature stop codon eleven amino acids into exon 3 (Figure 3B). Electroporation of the LIN28B morpholino into the primitive streak region of control embryos increased mature miR-let-7b levels two-fold relative to a control morpholino (Figure 3C). Although LIN28B forced expression had no effect on miR-let-7b levels in untreated embryos (Figure 3C), forced expression in SU5402 treated embryos significantly reduced miR-let-7b abundance towards control levels (Figure 3D). These results indicate that LIN28B can modulate miR-let-7b levels in control embryos, and that the dramatic (50 fold) increase in mature miR-let-7b following inhibition of FGF signaling is at least partially due to reduced LIN28B expression.

The effects of LIN28B over expression and knockdown on the levels of other FGF-regulated miR transcripts were also assessed. LIN28B morpholino knockdown decreased, and LIN28B over expression increased, the levels of miRs -130b and -19b, while the levels of miRs -9, -107 and -218 were unaffected (Figure 4). Together, these results show that FGF signaling negatively regulates miR abundance through LIN28B dependent and LIN28B independent mechanisms.
Predicted mRNA targets of FGF-regulated miRs - To investigate the function of miRs negatively regulated by FGF signaling, bioinformatic analyses were performed to identify their predicted mRNA targets. Transcriptional profiling had previously identified 11,516 genes that are expressed in the primitive streak region of HH stage 4 chicken embryos (4). The 3' UTRs of these mRNAs were analyzed using the TargetScan algorithm (32,39,40) to identify mRNAs containing evolutionarily conserved seed sites for miRs -let-7b, -9, -19b and -107 that are identical in sequence to their chicken orthologs (44-46). Human miR-130b differs from the chicken ortholog by several nucleotides and miR-218 is not expressed in HeLa cells, so miR mimics corresponding to chicken miRs-130b and -218 were transfected into cells along with the appropriate luciferase reporter and β-gal control vectors to adjust for transfection efficiency.

The chicken PDGFRA 3'UTR contains predicted seed sites for miRs -130b and -218. The reporter vector containing the 3' UTR of PDGFRA was significantly downregulated by co-transfection of miR-130b and/or miR-218 (Figure 5A). While the luciferase vector containing the PDGFRA miR-218 MRE showed a 35% reduction when co-transfected with miR-218, luciferase expression for the vector containing the miR-130b MRE was not reduced following co-transfection with miR-130b. However, a miR-130b seed site target protector morpholino abolished the reduction in luciferase activity observed with the full length PDGFRA 3'UTR, demonstrating functionality of the miR-130b seed site (Figure 5A).

A luciferase vector containing a miR-130b MRE sequence from the ACVR1 3'UTR showed a 50% reduction following co transfection with miR-130 (Figure 5A). The ACVR2B 3'UTR contains two predicted seed sites each for miRs-107 and -let-7b. The human and chicken orthologs of these miRs have identical sequences, and both miRs are expressed in HeLa cells. Addition of the ACVR2B 3' UTR to the parental luciferase expression vector reduced luciferase activity by approximately 40%, and each of the individual MREs reduced luciferase activity relative to the parental vector (Figure 5B). UTRs of several additional mRNAs showing putative seed sites for miR-let-7b were also tested (Figure 5C). Addition of 3'UTRs or miR-let-7b MREs from LIN28B, ABL2, LRIG1, CPEB1, CPEB3, and CPEB4, but not from COL3A, significantly reduced luciferase activity relative to the control vector (Figure 5C). Together, these results demonstrate direct targeting of numerous mRNAs by FGF-regulated miRs.

FGF-regulated miR target mRNAs containing predicted seed sites - To determine if mRNAs containing predicted miR seed sites are targeted by FGF-regulated miRs, the 3'UTRs or the miR recognition sites (MRE: miR seed site plus 60 nt of surrounding sequence) from candidate mRNAs were cloned downstream of a luciferase reporter and assayed for repression in HeLa cells. HeLa cells express human miRs-let-7b, -9, -19b and -107 that are identical in sequence to their chicken orthologs (44-46). Human miR-130b differs from the chicken ortholog by several nucleotides and miR-218 is not expressed in HeLa cells, so miR mimics corresponding to chicken miRs-130b and -218 were transfected into cells along with the appropriate luciferase reporter and β-gal control vectors to adjust for transfection efficiency.

The chicken PDGFRA 3'UTR contains predicted seed sites for miRs -130b and -218. The reporter vector containing the 3' UTR of PDGFRA was significantly downregulated by co-transfection of miR-130b and/or miR-218 (Figure 5A). While the luciferase vector containing the PDGFRA miR-218 MRE showed a 35% reduction when co-transfected with miR-218, luciferase expression for the vector containing the miR-130b MRE was not reduced following co-transfection with miR-130b. However, a miR-130b seed site target protector morpholino abolished the reduction in luciferase activity observed with the full length PDGFRA 3'UTR, demonstrating functionality of the miR-130b seed site (Figure 5A).

A luciferase vector containing a miR-130b MRE sequence from the ACVR1 3'UTR showed a 50% reduction following co transfection with miR-130 (Figure 5A). The ACVR2B 3'UTR contains two predicted seed sites each for miRs-107 and -let-7b. The human and chicken orthologs of these miRs have identical sequences, and both miRs are expressed in HeLa cells. Addition of the ACVR2B 3' UTR to the parental luciferase expression vector reduced luciferase activity by approximately 40%, and each of the individual MREs reduced luciferase activity relative to the parental vector (Figure 5B). UTRs of several additional mRNAs showing putative seed sites for miR-let-7b were also tested (Figure 5C). Addition of 3'UTRs or miR-let-7b MREs from LIN28B, ABL2, LRIG1, CPEB1, CPEB3, and CPEB4, but not from COL3A, significantly reduced luciferase activity relative to the control vector (Figure 5C). Together, these results demonstrate direct targeting of numerous mRNAs by FGF-regulated miRs.
To determine whether changes in luciferase activity were due to inhibition of translation or degradation of target mRNAs, transfection assays were repeated and luciferase mRNA levels determined by real-time RTPCR. As shown in Figure 5D, the addition of miRs -130b and -218 led to a reduction of luciferase transcripts which contained the PDGFRα 3'UTR. Among targets of miR-let-7b, only the 3' UTRs of CPEB1 and CPEB3 appeared to induce luciferase transcript degradation. Addition of ACVR1, LRIG1 and CPEB4 3' UTRs increased luciferase mRNA abundance, even though the final protein levels (measured by luciferase activity) were decreased (Figure 5D).

Analysis of FGF-regulated miRs in vivo - To investigate whether FGF-regulated miRs target specific mRNAs in vivo, miR over expression experiments were performed in HH stage 4 embryos. Despite extensive effort, antibodies recognizing the chicken proteins encoded by targeted mRNAs could not be identified, so the levels of targeted mRNAs were assessed. When HH stage 4 embryos were electroporated with mimics for miRs -let-7b, -130b and -218, mature miR levels increased 1350, 18, and 60 fold, respectively, as determined by QPCR. Of the mRNA targets and miR mimics combinations tested, PDGFRα mRNA levels were significantly downregulated by mimics for miRs -130b and 218, and CPEB4 mRNA levels were reduced following electroporation of let-7b mimics (Figure 6A). Although CPEB4 mRNAs are not detectable by whole mount ISH, PDGFRα mRNAs are readily detectable in the primitive streak (4,8). Reduction in PDGFA mRNA levels by miRs -130b and -218 was also evident by ISH, while a control miR had no effect on PDGFRα ISH staining intensity (Figure 6B-E). Although electroporation of mimics for miRs -130b and 218 did not detectably alter migration of cells through the primitive streak (data not shown), electroporation of the LIN28B morpholino led to altered cell migration and abnormal embryo development (Figure 7). Cells containing the LIN28B morpholino were frequently observed aggregated above the primitive streak, with Lin28B-positive few cells in the mesoderm. The mesoderm layer also appeared to contain fewer cells, although cells containing the LIN28B morpholino were observed in the endoderm. Embryos electroporated with the control morpholino were dispersed throughout the mesodermal and endodermal layers.

DISCUSSION

We have shown that FGF signaling regulates gene expression during gastrulation by negatively modulating the abundance of miRs -let-7b, -9, -19b, -107 and -218. FGF signaling is required for expression of LIN28B, which gain and loss of function studies have shown inhibits the processing of miRs-let-7b, 19b and 130b. miR seed site prediction in 3'UTRs of mRNAs expressed in the primitive streak identifies approximately 580 potential targets of FGF-regulated miRs. Receptor kinases, including BMP Type I and Type II receptors, TGFBR1, and PDGFRα, are over represented among mRNAs targeted by FGF-regulated miRs. Cell culture transfection assays and in vivo experiments confirm the targeting of these mRNAs. These findings show that FGF signaling can regulate signaling pathways and other processes during gastrulation by negatively modulating miR abundance through mechanisms that include upregulating LIN28B expression.

Previous studies have indirectly linked FGF signaling and LIN28 expression to the generation and self-renewal of cultured stem cells (47-50). Here we find that inhibition of FGF signaling in gastrulating chicken embryos causes a 10-fold decrease of Lin28B mRNA levels, while Lin28A mRNA levels are unaffected. Reduced LIN28B levels are at least partially responsible for increased miR abundance, since knock-down of Lin28B causes an increase in miRs -let-7b, 19b, and -130b, and forced expression of LIN28B in SU5402 treated embryos reduced miR-let-7b abundance towards control levels. Lin28A can block the processing of miR-let7 family members at the level of intermediate formation or final maturation (51), while Lin28B can inhibit processing of miR-let7 by sequestering the primary transcripts (21). Our results show that LIN28B can also regulate processing of miRs -19b and -130b. Given the complexity of the miR processing pathway and the unknown potential interplay between LIN28A and LIN28B function, the variety of patterns observed in the relative abundance of primary transcripts versus miR intermediates when comparing control versus SU5402 treated embryos is not surprising. miRs-
FGF Signaling Regulates MicroRNA Abundance

9, -107, and -218 show increased levels following inhibition of FGF signaling without apparent involvement of LIN28B, suggesting that FGF signaling affects their expression through direct or indirect transcriptional repression. A few studies have similarly linked FGF signaling to the downregulation of miR expression. Inhibition of FGFR1 during zebrafish fin regeneration led to the downregulation of 22 and upregulation of 34 miRs (52). In chicken embryos, miR-206 is repressed by FGF4 in somatic skeletal muscle cells (53). It is not known whether the FGF dependent reduction in miR levels in these two studies is mediated by increased expression of Lin28A or Lin28B. A striking finding in the present study is that the primary effect of FGF signaling on miR expression in gastrulating chicken embryos is to negatively modulate their abundance.

Computational methods identified almost 600 potential targets of FGF-regulated miRs. Predicted targets include several miRNAs (CPEB1, CPEB4, CNOT6 and CNOT7L) that themselves code for proteins that regulate mRNA polyadenylation and deadenylation and mRNA stability. Luciferase assays confirmed the targeting of CPEB1 and CPEB4 by miR-let-7b. A second group of overrepresented mRNAs code for serine/threonine and tyrosine kinase receptors that include the BMP type I receptors ACVR1 and ACVR1C, Type II receptors ACVR2A and ACVR2B, plus TGFBR1, TGFBR3 and PDGFRA. Targeting of ACVR1, ACVR2B and PDGFRA was confirmed in cell culture assays, and the ability of miRs -130b and -218 to reduce mRNA levels of PDGFRA was confirmed in vivo. Although the ligand specificities in chicken of ACVR1, ACVR1C, ACVR2A and ACVR2B receptors have not been determined, in other organisms they can bind nodal and BMP growth factors. In avian embryos, Nodal is expressed along the anterior portion of the primitive streak and BMPs are expressed posteriorly, creating an antagonistic signaling gradient that patterns the emerging mesoderm. A similar expression gradient is observed in most vertebrate embryos (2). The potential targeting of these receptors by FGF regulated miRs suggests that FGF signaling can influence dorsal ventral patterning. Unfortunately, the lack of antibodies recognizing these receptors in chicken has precluded a more in-depth analysis of this possibility.

PDGFRA controls N-cadherin protein expression through PI3 kinase and is required for migration of epiblast cells through the primitive streak (8). We have shown both in vitro and in vivo that miR-218 causes a substantial downregulation of PDGFRA mRNA. Together, this data suggests a mechanism by which FGF signaling-mediated downregulation of miR-218 expression leads to increased PDGFRA protein levels, permitting cell migration through the primitive streak. The finding that FGF signaling regulates miR abundance during gastrulation adds another layer of complexity to the pivotal role that this pathway plays during embryogenesis.

REFERENCES

FGF Signaling Regulates MicroRNA Abundance


FOOTNOTES
*This work was supported by NIH grant 1 P41 HD064559 and an American Heart Association grant to PBA.
1Department of Molecular and Cellular Biology, University of Arizona, Tucson, Arizona, 85724 USA
2Department of Cellular and Molecular Medicine, University of Arizona, Tucson, Arizona, 85724 USA
3To whom correspondence should be addressed: Dr. Parker B. Antin, Department of Cellular and Molecular Medicine, University of Arizona, PO Box 245217, 1656 E. Mabel St., Tucson, AZ 85724-5217 USA, Tel: (520) 626-6382; Fax: (520) 626-7600; E-mail: pba@email.arizona.edu
4The abbreviations used are: EMT, epithelial-mesenchymal-transition; HH, Hamburger-Hamilton; miR, microRNA.

FIGURE LEGENDS
FIGURE 1. Changes in miR expression following inhibition of FGF signaling. A, Comparison of miR expression levels in RNA samples from control versus SU5402 treated primitive streak regions of embryos, determined by real-time PCR. All miRs detected in the primitive streak region of gastrula stage embryos (Supplemental File 1) were assessed for changes in expression levels. Only those miRs with a statistically significant change in abundance are shown. Statistical significance was determined using a non-paired t-test. Inset shows ISH detection of miR-let-7b in control versus SU5402 treated embryos. miR-let-7b expression was increased throughout the embryo, particularly in the primitive streak and more anteriorly. B, ISH analysis of miR expression HH stage 4 embryos. miR-let-7b staining in the primitive streak was variable, with some embryos showing robust streak staining (Figure 1B), while others showed somewhat reduced expression in the streak itself (e.g., control embryo in (A)).

FIGURE 2. FGF signaling regulates miR abundance through transcriptional and post transcriptional mechanisms. A, Simplified pathway of miR processing, showing the location of primers for detecting pri- or pre- miR transcripts. B, Fold change in abundance of pri-, pre-miRs, and mature miRs in control versus SU5402 treated embryos. Pri-miR levels are increased for all FGF-regulated miRs except miR-130b. Only miRs -130b and -19b show increased pre-miRs levels, while miR-9 shows a decrease in pre-miRs.

FIGURE 3. Lin28B regulation of miR abundance. A, Real-time RT-PCR comparison of LIN28A and LIN28B mRNA levels in samples from primitive streak regions of control versus SU5402 treated embryos. LIN28B mRNA levels are reduced ten-fold by inhibition of FGF signaling while LIN28A mRNA levels are unchanged. LIN28B ISH analysis of control versus SU5402 treated embryos. B, Top, exon-intron structure of the LIN28B gene, showing the location of the splice blocking morpholino at the exon2-intron-2 boundary that is predicted to cause skipping of exon 2 and introduction of a premature in frame translation stop codon. Bottom, RT-PCR analysis of RNA isolated from HH stage 4 embryos electroporated with a control (C) or Lin28B (M) morpholino. The PCR product generated using primers spanning exons1-3 show the expected 382nt product for the control morpholinos sample, and 194nt product for the sample electroporated with the Lin28B morpholino, reflecting absence of exon 2. C, Normalized levels of miR-let-7b in samples from embryos electroporated with control or Lin28B targeting morpholinos. The Lin28B morpholino increases let-7b levels two-fold. Forced expression of Lin28B has no effect on let-7b levels. D, Forced expression of Lin28B in SU5402 treated embryos (SU5402 + LIN28B) reduces let-7b levels relative to SU5402 treated embryos electroporated with a control plasmid.

FIGURE 4. Fold change in levels of miRs -9, -130b, 19b, -107, -218 following Lin28B morpholino knockdown or Lin28B forced expression. Levels of miRs -130b and -19b increased with Lin28B knockdown and decreased following Lin28B forced expression. Levels of other miRs were unchanged.

FIGURE 5. 3'UTR luciferase reporter assays. A, 3' UTR reporter analysis of the vectors containing the PDGFRA or ACVR1 3'UTR, or a MRE containing miR seed sites plus 50-60 nt of flanking sequence, with or without exogenous miR-130b and/or miR-218. Addition of miRs significantly reduced luciferase activity. Luciferase reporter vectors containing the PDGFRA 3' UTR or microRNA recognition elements show reduced luciferase activity, except for the vector containing the miR-130b MRE. Addition of a morpholino targeting the miR-130b seed site (target protector) restores luciferase activity of the PDGFRA 3'UTR vector to control levels. B, 3' UTR reporter analysis of the vectors containing the ACVR2B 3'UTR or MREs for each of the miR-let-7b or miR-107 binding sites. C, Of seven strongly-predicted mRNA targets of miR-let-7b, only COL3A failed to show a reduction in luciferase activity. D, Effects of 3'UTRs and MREs on luciferase mRNA levels. Although all of the 3' UTs tested reduce luciferase activity and are targeted by miRs (A,B,C), only 3'UTRs from PDGFRA, and the CPEB1 and CPEB3 MREs, reduced luciferase mRNA levels. Addition of 3'UTRs for ACVR1, LRG1 and CPEB4 increased luciferase mRNA levels.
FIGURE 6. PDGFRA and CPEB4 mRNAs are targeted for degradation by miRs *in vivo*. A, Real-time PCR analysis comparing mRNA levels of predicted miR targets in samples from primitive streak regions of embryos electroporated with a control miR (cel-167) versus embryos electroporated with miRs -let-7b, -130b and/or -218. Asterisks indicate changes in mRNA levels that are statistically significant (p<0.05). B-E, Visualization of PDGFRA mRNA reduction in embryos by over expression of miRs. B,C, Images showing the same embryo electroporated with a control miR mimic (cel-167) or with mimics for miRs-130b and -218, visualizing the location of miR mimics (green fluorescence in B, D), and PDGFRA mRNAs by ISH (C, E). Introduction of miRs -130b and -218 leads to a reduction of PDGFRA mRNAs in the primitive streak.

FIGURE 7. LIN28B knockdown impairs cell migration and embryo development. Fluorescein conjugated Control (cel-167) or LIN28B morpholinos were electroporated into the primitive streak region epiblast of HH stage 4 embryos. Following 14 hours incubation, morpholino containing cells were visualized using anti-Flourescein-HRP. Embryos electroporated with the LIN28B morpholino (D-F, F’) showed impaired cell migration and development relative to embryos receiving the Control morpholino (A-C, C’).
Figure 2

A

B

- **Primary Transcript**
- **Precursor**
- **Mature miRNA**

* = p-value < 0.05
** = p-value < 0.01
*** = p-value < 0.005

Fold Change SU5402 Versus Control
Figure 3

A

B

C

D

miR-let7b Levels

miR-let7b levels

FGF Signaling Regulates MicroRNA Abundance
Figure 4
Figure 5
Figure 6

[Bar chart and images of in situ hybridization showing regulation of microRNA abundance by FGF signaling]
Figure 7

*FGF Signaling Regulates MicroRNA Abundance*
### Supplemental Table 1

This Table shows all microRNAs detected using a multispecies microarray, along with relative signal above background, adjusted p-value, and whether expression was verified by PCR.

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Signal Above Background</th>
<th>Adjusted p-value</th>
<th>PCR Verified</th>
</tr>
</thead>
<tbody>
<tr>
<td>gga-miR-1</td>
<td>4705</td>
<td>0</td>
<td>NO</td>
</tr>
<tr>
<td>gga-miR-101</td>
<td>72</td>
<td>0</td>
<td>YES</td>
</tr>
<tr>
<td>gga-miR-103</td>
<td>186</td>
<td>0</td>
<td>YES</td>
</tr>
<tr>
<td>gga-miR-124b</td>
<td>139</td>
<td>0</td>
<td>YES</td>
</tr>
<tr>
<td>gga-miR-130a</td>
<td>671</td>
<td>0</td>
<td>YES</td>
</tr>
<tr>
<td>gga-miR-130b</td>
<td>184</td>
<td>0</td>
<td>YES</td>
</tr>
<tr>
<td>gga-miR-15b</td>
<td>208</td>
<td>0</td>
<td>YES</td>
</tr>
<tr>
<td>gga-miR-17-5p</td>
<td>1399</td>
<td>0</td>
<td>YES</td>
</tr>
<tr>
<td>gga-miR-190</td>
<td>136</td>
<td>0</td>
<td>NO</td>
</tr>
<tr>
<td>gga-miR-19a</td>
<td>861</td>
<td>0</td>
<td>YES</td>
</tr>
<tr>
<td>gga-miR-203</td>
<td>194</td>
<td>0</td>
<td>YES</td>
</tr>
<tr>
<td>gga-miR-20b</td>
<td>1235</td>
<td>0</td>
<td>YES</td>
</tr>
<tr>
<td>gga-miR-302b</td>
<td>1687</td>
<td>0</td>
<td>YES</td>
</tr>
<tr>
<td>gga-miR-302b*</td>
<td>1846</td>
<td>0</td>
<td>YES</td>
</tr>
<tr>
<td>gga-miR-302d</td>
<td>31216</td>
<td>0</td>
<td>YES</td>
</tr>
<tr>
<td>gga-miR-30a-5p</td>
<td>156</td>
<td>0</td>
<td>YES</td>
</tr>
<tr>
<td>gga-miR-551</td>
<td>1209</td>
<td>0</td>
<td>YES</td>
</tr>
<tr>
<td>gga-miR-106</td>
<td>1067</td>
<td>0.01</td>
<td>YES</td>
</tr>
<tr>
<td>gga-miR-16</td>
<td>185</td>
<td>0.01</td>
<td>YES</td>
</tr>
<tr>
<td>gga-miR-181a</td>
<td>49</td>
<td>0.01</td>
<td>YES</td>
</tr>
<tr>
<td>gga-miR-18b</td>
<td>1136</td>
<td>0.01</td>
<td>YES</td>
</tr>
<tr>
<td>gga-miR-26a</td>
<td>413</td>
<td>0.01</td>
<td>YES</td>
</tr>
<tr>
<td>gga-miR-29a</td>
<td>437</td>
<td>0.01</td>
<td>YES</td>
</tr>
<tr>
<td>gga-miR-30c</td>
<td>237</td>
<td>0.01</td>
<td>YES</td>
</tr>
<tr>
<td>gga-miR-455</td>
<td>254</td>
<td>0.01</td>
<td>NO</td>
</tr>
<tr>
<td>gga-miR-92</td>
<td>237</td>
<td>0.01</td>
<td>YES</td>
</tr>
<tr>
<td>gga-let-7b</td>
<td>190</td>
<td>0.02</td>
<td>YES</td>
</tr>
<tr>
<td>gga-miR-10b</td>
<td>169</td>
<td>0.02</td>
<td>NO</td>
</tr>
<tr>
<td>gga-miR-20a</td>
<td>1520</td>
<td>0.02</td>
<td>YES</td>
</tr>
<tr>
<td>gga-miR-21</td>
<td>734</td>
<td>0.02</td>
<td>YES</td>
</tr>
<tr>
<td>gga-miR-23b</td>
<td>338</td>
<td>0.02</td>
<td>YES</td>
</tr>
<tr>
<td>gga-miR-107</td>
<td>113</td>
<td>0.03</td>
<td>YES</td>
</tr>
<tr>
<td>miRNA</td>
<td>FPKM</td>
<td>P-value</td>
<td>Expression</td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>gga-miR-133b</td>
<td>425</td>
<td>0.03</td>
<td>YES</td>
</tr>
<tr>
<td>gga-miR-19b</td>
<td>2539</td>
<td>0.03</td>
<td>YES</td>
</tr>
<tr>
<td>gga-miR-202*</td>
<td>229</td>
<td>0.03</td>
<td>YES</td>
</tr>
<tr>
<td>gga-miR-216</td>
<td>190</td>
<td>0.03</td>
<td>YES</td>
</tr>
<tr>
<td>gga-miR-489</td>
<td>135</td>
<td>0.03</td>
<td>YES</td>
</tr>
<tr>
<td>gga-miR-138</td>
<td>955</td>
<td>0.04</td>
<td>YES</td>
</tr>
<tr>
<td>gga-miR-148a</td>
<td>151</td>
<td>0.04</td>
<td>YES</td>
</tr>
<tr>
<td>gga-miR-15a</td>
<td>443</td>
<td>0.04</td>
<td>YES</td>
</tr>
<tr>
<td>gga-miR-194</td>
<td>117</td>
<td>0.04</td>
<td>YES</td>
</tr>
<tr>
<td>gga-miR-218</td>
<td>153</td>
<td>0.04</td>
<td>YES</td>
</tr>
<tr>
<td>gga-miR-9</td>
<td>135</td>
<td>0.04</td>
<td>YES</td>
</tr>
<tr>
<td>gga-miR-199</td>
<td>74</td>
<td>0.05</td>
<td>YES</td>
</tr>
</tbody>
</table>
Supplemental Table 2

This Table shows the number of miR seed sites identified in the 3' UTRs of mRNAs expressed in the primitive streaks of HH Stage 4 chicken embryos using a database acquired from TargetScan. Expression data was obtained from Hardy et al., 2011.

<table>
<thead>
<tr>
<th>NCBI Gene Name</th>
<th>NCBI Gene ID</th>
<th>miR-let7b</th>
<th>miR-9</th>
<th>miR-19b</th>
<th>miR-107</th>
<th>miR-130b</th>
<th>miR-218</th>
<th># of Sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>USP32</td>
<td>417646</td>
<td>X</td>
<td></td>
<td>X (2)</td>
<td>X</td>
<td>X (3)</td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>CPEB4</td>
<td>416363</td>
<td>X</td>
<td>X</td>
<td>X (2)</td>
<td></td>
<td>X (2)</td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>LIN28B</td>
<td>421786</td>
<td>X (4)</td>
<td>X</td>
<td></td>
<td>X (2)</td>
<td></td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>PTEN</td>
<td>423675</td>
<td></td>
<td></td>
<td>X (3)</td>
<td>X (3)</td>
<td></td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>CPEB1</td>
<td>415326</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X (2)</td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>CREB5</td>
<td>428435</td>
<td>X (2)</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>ACVR2B</td>
<td>374213</td>
<td>X (2)</td>
<td></td>
<td></td>
<td>X (2)</td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>BACH2</td>
<td>421809</td>
<td>X (2)</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>BTBD7</td>
<td>423424</td>
<td>X</td>
<td>X (2)</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>CNOT6L</td>
<td>422508</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>GABRB2</td>
<td>414890</td>
<td>X (2)</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>HIC2</td>
<td>416768</td>
<td>X</td>
<td>X (2)</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>HIPK2</td>
<td>374138</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>HLF</td>
<td>417395</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>4</td>
</tr>
<tr>
<td>LRIG1</td>
<td>427592</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>MYBL1</td>
<td>396158</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>PRKCE</td>
<td>421409</td>
<td>X</td>
<td>X</td>
<td>X (3)</td>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>RBM9</td>
<td>418055</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>SOCS5</td>
<td>428570</td>
<td>X (2)</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>TNRC6B</td>
<td>418010</td>
<td>X (2)</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>ZMYND11</td>
<td>420455</td>
<td>X (4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>ACSL1</td>
<td>422547</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>ADIPOR2</td>
<td>418155</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>ATRX</td>
<td>422331</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>BAHD1</td>
<td>423010</td>
<td>X (2)</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>BAI3</td>
<td>421875</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>3</td>
</tr>
<tr>
<td>CCND1</td>
<td>396341</td>
<td>X (2)</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>CLOCK</td>
<td>373991</td>
<td>X (2)</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>DLG5</td>
<td>423728</td>
<td>X (2)</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>EPC2</td>
<td>424309</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X (2)</td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>ETNK1</td>
<td>418196</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>FIGN</td>
<td>429028</td>
<td>X (3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>G3BP2</td>
<td>422576</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>GALNT1</td>
<td>420955</td>
<td>X (2)</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>ITSN1</td>
<td>418506</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Gene</td>
<td>Symbol</td>
<td>X (2)</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>--------</td>
<td>-------</td>
<td>-----</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KCNJ2</td>
<td>396328</td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KLF13</td>
<td>427493</td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KPNA4</td>
<td>425012</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MKL2</td>
<td>416421</td>
<td>X</td>
<td></td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NAV2</td>
<td>422977</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NMT2</td>
<td>420529</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLEKHA1</td>
<td>423940</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRKX</td>
<td>418656</td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PTPRG</td>
<td>374208</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ROBO2</td>
<td>395291</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCN2A</td>
<td>395945</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLC35F1</td>
<td>770279</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLC6A1</td>
<td>415960</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLITRK2</td>
<td>422396</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMARCD2</td>
<td>428275</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOCS6</td>
<td>421020</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPEN</td>
<td>419460</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPIRE1</td>
<td>421033</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ST18</td>
<td>421115</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STARD13</td>
<td>418908</td>
<td>X</td>
<td></td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SYT1</td>
<td>396083</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TANC2</td>
<td>419951</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TESK2</td>
<td>424596</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TNC</td>
<td>396440</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UBFD1</td>
<td>771251</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ULK2</td>
<td>417614</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VAV3</td>
<td>373939</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>YAF2</td>
<td>417790</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZNF711</td>
<td>422269</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZPLD1</td>
<td>418405</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABL2</td>
<td>395410</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACCN2</td>
<td>426883</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACVR1</td>
<td>395246</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADAM12</td>
<td>425843</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AFF1</td>
<td>422590</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AFF4</td>
<td>416337</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AKAP1</td>
<td>417668</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALCAM</td>
<td>396092</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AP1S2</td>
<td>418621</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>APPL1</td>
<td>416002</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARH GAP1</td>
<td>420473</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARH GAP5</td>
<td>428883</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATG16L1</td>
<td>424746</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene</td>
<td>Location</td>
<td>Status</td>
<td>Count</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>----------</td>
<td>--------</td>
<td>-------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATP11A</td>
<td>418749</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATP2A2</td>
<td>396446</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATP2B2</td>
<td>415958</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATP2B4</td>
<td>419934</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AXIN2</td>
<td>395152</td>
<td>X (2)</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B4GALT5</td>
<td>419358</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAZ2B</td>
<td>395400</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BCAT1</td>
<td>418193</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIRC6</td>
<td>421463</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMP3</td>
<td>422581</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BRPF3</td>
<td>421182</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BZW1</td>
<td>424073</td>
<td>X (2)</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BZW2</td>
<td>420594</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CACNA1C</td>
<td>395891</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAPRIN2</td>
<td>418141</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBFB</td>
<td>395257</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHD6</td>
<td>419127</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CIT</td>
<td>416984</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CLTC</td>
<td>395272</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CNOT7</td>
<td>422733</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CNR1</td>
<td>428633</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COBLL1</td>
<td>424182</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COL15A1</td>
<td>420803</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRTAP</td>
<td>395992</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CUL5</td>
<td>418969</td>
<td>X (2)</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CYLD</td>
<td>415725</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DCUN1D3</td>
<td>427002</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DLL1</td>
<td>395820</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DSEL</td>
<td>421028</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EIF4G2</td>
<td>395905</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ERBB2IP</td>
<td>431583</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ERC2</td>
<td>769533</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FLRT3</td>
<td>428552</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FZD6</td>
<td>378788</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GABRG2</td>
<td>396289</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GOLSYN</td>
<td>420277</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GOLT1B</td>
<td>418192</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRB10</td>
<td>420948</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HIPK1</td>
<td>419880</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HIVEP1</td>
<td>420856</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HOXA10</td>
<td>776143</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HPRT1</td>
<td>395653</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INOC1</td>
<td>423210</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ITPKB</td>
<td>416743</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene</td>
<td>Entrez ID</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>-----------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JAKMIP2</td>
<td>416352</td>
<td>X</td>
<td>X</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KIF13A</td>
<td>420832</td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KIF21A</td>
<td>417692</td>
<td>X</td>
<td>X</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KIF5C</td>
<td>408046</td>
<td>X</td>
<td>X</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LARP5</td>
<td>420457</td>
<td></td>
<td></td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LIN28</td>
<td>428206</td>
<td>X</td>
<td></td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LNPEP</td>
<td>427279</td>
<td>X</td>
<td>X</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LRRG3</td>
<td>417813</td>
<td>X</td>
<td>X</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LRP1B</td>
<td>424301</td>
<td></td>
<td></td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAGI2</td>
<td>417729</td>
<td>X</td>
<td>X</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAP4K4</td>
<td>418714</td>
<td></td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAPK1</td>
<td>373953</td>
<td></td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEGF11</td>
<td>415545</td>
<td>X</td>
<td></td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MLLT10</td>
<td>420509</td>
<td></td>
<td></td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MMP16</td>
<td>396116</td>
<td>X (2)</td>
<td></td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MPHOSPH 9</td>
<td>416829</td>
<td>X</td>
<td>X</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MYH9</td>
<td>396469</td>
<td>X</td>
<td>X</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NARG1</td>
<td>422440</td>
<td></td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NAT12</td>
<td>423547</td>
<td>X</td>
<td>X</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NID2</td>
<td>423583</td>
<td>X</td>
<td>X</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NLK</td>
<td>417669</td>
<td>X</td>
<td>X</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NOVA1</td>
<td>423300</td>
<td>X</td>
<td>X</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NPTN</td>
<td>415316</td>
<td>X</td>
<td>X</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NRPI1</td>
<td>395560</td>
<td></td>
<td></td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ODZ4</td>
<td>395406</td>
<td>X</td>
<td>X</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OSBPL10</td>
<td>420663</td>
<td>X</td>
<td></td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCDH10</td>
<td>408048</td>
<td>X</td>
<td>X</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCNX</td>
<td>426230</td>
<td></td>
<td></td>
<td>X (2)</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PDGFRα</td>
<td>395509</td>
<td></td>
<td></td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PDKI1L</td>
<td>419604</td>
<td>X</td>
<td>X</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PFKFB3</td>
<td>418247</td>
<td>X</td>
<td>X</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PHF13</td>
<td>428174</td>
<td></td>
<td></td>
<td>X (2)</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPARGC1B</td>
<td>416148</td>
<td>X (2)</td>
<td></td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRKACB</td>
<td>424542</td>
<td>X</td>
<td>X</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PSD</td>
<td>423760</td>
<td>X</td>
<td>X</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PT4A1</td>
<td>421877</td>
<td>X</td>
<td>X</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PTPRK</td>
<td>421711</td>
<td>X</td>
<td>X</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAI14</td>
<td>427431</td>
<td>X</td>
<td>X</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAP1B</td>
<td>417840</td>
<td>X</td>
<td>X</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAPGGEF4</td>
<td>429022</td>
<td>X</td>
<td>X</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RDCR1</td>
<td>428911</td>
<td>X</td>
<td>X</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RDX</td>
<td>395511</td>
<td>X</td>
<td>X</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RET</td>
<td>396107</td>
<td></td>
<td></td>
<td>X (2)</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene</td>
<td>Chromosome</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RFXDC2</td>
<td>431530</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RNF111</td>
<td>374021</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RNF216</td>
<td>416486</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RPS6KA3</td>
<td>418605</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RTN1</td>
<td>378791</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SASH1</td>
<td>421621</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCML2</td>
<td>418613</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEC14L1</td>
<td>417350</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEMA6A</td>
<td>427387</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEMA6D</td>
<td>415430</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SFRS12</td>
<td>427168</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SFRS2IP</td>
<td>417805</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SIPA1L2</td>
<td>421523</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLC05A1</td>
<td>420174</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLMAP</td>
<td>396360</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMPD3</td>
<td>415619</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNF1LK2</td>
<td>768557</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOS2</td>
<td>423572</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOX11</td>
<td>396104</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOX4</td>
<td>395143</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPRED2</td>
<td>421271</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPSB1</td>
<td>419441</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUZ12</td>
<td>417406</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAF4</td>
<td>419223</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAF4B</td>
<td>421086</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TBC1D14</td>
<td>422862</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TBC1D8</td>
<td>428012</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TBPL1</td>
<td>395517</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TGFBR1</td>
<td>374094</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>THOC2</td>
<td>422361</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TLK1</td>
<td>429024</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TNRC6A</td>
<td>416570</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRIM37</td>
<td>417628</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TSC1</td>
<td>417166</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UBE3A</td>
<td>418686</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UHRF1</td>
<td>420154</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>USP33</td>
<td>424553</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>USP47</td>
<td>423032</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VASH2</td>
<td>421364</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VAV2</td>
<td>373938</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VPS4B</td>
<td>420901</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WAC</td>
<td>428414</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WASL</td>
<td>417750</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WDFY3</td>
<td>422613</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene</td>
<td>Chromosome</td>
<td>X</td>
<td>Y</td>
<td>Z</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WDR47</td>
<td>424344</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZEB2</td>
<td>424306</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZNF516</td>
<td>421009</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABCF2</td>
<td>426038</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACVR1C</td>
<td>424325</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADAM10</td>
<td>374113</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADAMTS1</td>
<td>418479</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADAMTS1</td>
<td>415515</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADAMTS1</td>
<td>415600</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADAMTS3</td>
<td>428760</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADAMTS6</td>
<td>427160</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADCYAP1</td>
<td>408251</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADNP</td>
<td>419350</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADPGK</td>
<td>415317</td>
<td>X</td>
<td>X</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AFF3</td>
<td>418705</td>
<td></td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AFTPH</td>
<td>421277</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMOTL2</td>
<td>424863</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANK2</td>
<td>422689</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANK3</td>
<td>396309</td>
<td></td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANKRD46</td>
<td>426304</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANXA7</td>
<td>423747</td>
<td></td>
<td>X</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>APBA1</td>
<td>427246</td>
<td></td>
<td>X</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR</td>
<td>422165</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARFGEF1</td>
<td>420170</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARHGEF1</td>
<td>419752</td>
<td></td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARID1A</td>
<td>419542</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARL6IP1</td>
<td>416604</td>
<td></td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARPC1A</td>
<td>770269</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARRDC4</td>
<td>415509</td>
<td></td>
<td>X</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASH1L</td>
<td>425064</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATF1</td>
<td>426873</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATG2B</td>
<td>423441</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATP13A3</td>
<td>424898</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATP1B1</td>
<td>396529</td>
<td></td>
<td>X</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATXN2</td>
<td>416877</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B3GAT1</td>
<td>373949</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAMBI</td>
<td>428413</td>
<td></td>
<td>X</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BICD1</td>
<td>427919</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIVM</td>
<td>374099</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BRUNOL6</td>
<td>427479</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BRWD1</td>
<td>418521</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BTBD9</td>
<td>421429</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene</td>
<td>Chromosome</td>
<td>Location</td>
<td>X Chromosome</td>
<td>Value</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>------------</td>
<td>----------</td>
<td>--------------</td>
<td>-------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BTG2</td>
<td>116</td>
<td>419932</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C15orf24</td>
<td>112</td>
<td>423288</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CACNA1D</td>
<td>112</td>
<td>395895</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CACNA2D1</td>
<td>112</td>
<td>768444</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CADM2</td>
<td>112</td>
<td>418463</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CALD1</td>
<td>112</td>
<td>373965</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAMSAP1L1</td>
<td>112</td>
<td>424345</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAPRIN1</td>
<td>112</td>
<td>426151</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAPZA2</td>
<td>112</td>
<td>417771</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAST</td>
<td>112</td>
<td>427121</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CASZ1</td>
<td>112</td>
<td>419452</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCNE1</td>
<td>112</td>
<td>426117</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCNJ</td>
<td>112</td>
<td>423821</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCNJL</td>
<td>112</td>
<td>416155</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCNL1</td>
<td>112</td>
<td>425024</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDH2</td>
<td>112</td>
<td>414745</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDK6</td>
<td>112</td>
<td>420558</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDK8</td>
<td>112</td>
<td>428080</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDYL</td>
<td>112</td>
<td>420877</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CEP68</td>
<td>112</td>
<td>421274</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHD1</td>
<td>112</td>
<td>374195</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHD7</td>
<td>112</td>
<td>421140</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CLK3</td>
<td>112</td>
<td>415295</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CNIH</td>
<td>112</td>
<td>423561</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CNNM2</td>
<td>112</td>
<td>428975</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTN3</td>
<td>112</td>
<td>416101</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTN4</td>
<td>112</td>
<td>396149</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COL1A2</td>
<td>112</td>
<td>396243</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COL3A1</td>
<td>112</td>
<td>396340</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COL4A1</td>
<td>112</td>
<td>395530</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COL4A2</td>
<td>112</td>
<td>418752</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COL4A5</td>
<td>112</td>
<td>422348</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COL5A2</td>
<td>112</td>
<td>423986</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COL6A3</td>
<td>112</td>
<td>396548</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COLEC12</td>
<td>112</td>
<td>421061</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPNE8</td>
<td>112</td>
<td>417690</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRMP1</td>
<td>112</td>
<td>395156</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSNK1A1</td>
<td>112</td>
<td>395924</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSNK1G1</td>
<td>112</td>
<td>415328</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTDPI</td>
<td>112</td>
<td>420813</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTGF</td>
<td>112</td>
<td>374135</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTNNA1</td>
<td>112</td>
<td>416184</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTNNA2</td>
<td>112</td>
<td>396035</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene</td>
<td>Chromosome</td>
<td>Start</td>
<td>End</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>------------</td>
<td>-------</td>
<td>-------</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CUL3</td>
<td></td>
<td>424804</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CXCL12</td>
<td></td>
<td>395180</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DCP1A</td>
<td></td>
<td>415992</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DDEF2</td>
<td></td>
<td>768599</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DD2</td>
<td></td>
<td>425541</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DDX19B</td>
<td></td>
<td>426504</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DDX26B</td>
<td></td>
<td>422239</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DHR3</td>
<td></td>
<td>419480</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIP2A</td>
<td></td>
<td>771400</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DLX1</td>
<td></td>
<td>429283</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DMD</td>
<td></td>
<td>396236</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DOCK3</td>
<td></td>
<td>415919</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DOCK4</td>
<td></td>
<td>417779</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DPP6</td>
<td></td>
<td>420440</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DR1</td>
<td></td>
<td>424496</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DSCAM</td>
<td></td>
<td>418525</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DUS1L</td>
<td></td>
<td>417381</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DUSP1</td>
<td></td>
<td>374192</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DUSP16</td>
<td></td>
<td>431336</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E2F7</td>
<td></td>
<td>417868</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EEA1</td>
<td></td>
<td>417896</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EIF4A2</td>
<td></td>
<td>395232</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EIF4E</td>
<td></td>
<td>422703</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EIF5A2</td>
<td></td>
<td>396545</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ELL2</td>
<td></td>
<td>427119</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ELOVL5</td>
<td></td>
<td>428646</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ENC1</td>
<td></td>
<td>427203</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ERC1</td>
<td></td>
<td>418153</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ERRFI1</td>
<td></td>
<td>419364</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ETL3</td>
<td></td>
<td>422022</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EZH2</td>
<td></td>
<td>420784</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FAM3C</td>
<td></td>
<td>417758</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FBN1</td>
<td></td>
<td>373992</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FBN2</td>
<td></td>
<td>374203</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FBXW7</td>
<td></td>
<td>422481</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FGF18</td>
<td></td>
<td>395453</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FGF7</td>
<td></td>
<td>415439</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FNBP1L</td>
<td></td>
<td>424495</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FRMD4B</td>
<td></td>
<td>426071</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FRZB</td>
<td></td>
<td>395545</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GALNT2</td>
<td></td>
<td>421540</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GIT2</td>
<td></td>
<td>374035</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GLI2</td>
<td></td>
<td>395956</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GNAL</td>
<td></td>
<td>474379</td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene</td>
<td>Chromosome</td>
<td>EXON</td>
<td>AFFX</td>
<td>PMOD</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>------------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>----</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GNPTAB</td>
<td>418096</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GNS</td>
<td>417829</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GOLGA4</td>
<td>420758</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GOLGA7</td>
<td>426787</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GOT1</td>
<td>396261</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GPC6</td>
<td>418794</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRID1</td>
<td>428934</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GTF2H1</td>
<td>423081</td>
<td></td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GUCY1A2</td>
<td>771678</td>
<td></td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GULP1</td>
<td>423988</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HAS2</td>
<td>395594</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HDBP</td>
<td>395984</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HECTD1</td>
<td>423310</td>
<td></td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HELZ</td>
<td>417427</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HERC2</td>
<td>418682</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HERPUD2</td>
<td>420742</td>
<td></td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HES1</td>
<td>395128</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HIPK3</td>
<td>421597</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HMGB1</td>
<td>395724</td>
<td></td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HS2ST1</td>
<td>395140</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HS3ST2</td>
<td>416581</td>
<td></td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HSPA4L</td>
<td>422496</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HSPA8</td>
<td>395853</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ID4</td>
<td>374150</td>
<td></td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IDH3A</td>
<td>415362</td>
<td></td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IMPAD1</td>
<td>421132</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ING1</td>
<td>418759</td>
<td></td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INPP5A</td>
<td>423973</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INTS6</td>
<td>418876</td>
<td></td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ITCH</td>
<td>419145</td>
<td></td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ITGA6</td>
<td>396226</td>
<td></td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IVNS1ABP</td>
<td>426862</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KALRN</td>
<td>424253</td>
<td></td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KBTBD8</td>
<td>416085</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KCNAB1</td>
<td>395730</td>
<td></td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KCNC1</td>
<td>423076</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KCND2</td>
<td>374143</td>
<td></td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KCTD16</td>
<td>427644</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KCTD2</td>
<td>422116</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KIF1B</td>
<td>419449</td>
<td></td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KIF26A</td>
<td>423489</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KLF10</td>
<td>420255</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KLF3</td>
<td>429811</td>
<td></td>
<td></td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KRAS</td>
<td>418207</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene</td>
<td>CGID</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>----------</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LANCL3</td>
<td>427992</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LDLRAP1</td>
<td>419589</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LMNB1</td>
<td>396223</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LMO3</td>
<td>418179</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LMO7</td>
<td>418815</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LMTK2</td>
<td>769200</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LRRC8D</td>
<td>424513</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LRRN1</td>
<td>771599</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAEA</td>
<td>426024</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAFG</td>
<td>769355</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAP3K7IP3</td>
<td>418588</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAP3K9</td>
<td>776822</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAPK8IP3</td>
<td>426986</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MARK3</td>
<td>423480</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAT2B</td>
<td>416165</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCC</td>
<td>415603</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEF2A</td>
<td>395670</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEIS2</td>
<td>395588</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEIS2</td>
<td>395588</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MESDC1</td>
<td>415473</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MITF</td>
<td>395886</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MKRN1</td>
<td>395500</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MLL</td>
<td>414897</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MLL3</td>
<td>420437</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTMR10</td>
<td>415382</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTMR4</td>
<td>417472</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTMR6</td>
<td>418938</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MXRA5</td>
<td>418657</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MYO10</td>
<td>420913</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MYO1C</td>
<td>417555</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MYT1</td>
<td>419249</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NDRG1</td>
<td>420321</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NEBL</td>
<td>395148</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NEFM</td>
<td>396206</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NEO1</td>
<td>395822</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NF1</td>
<td>396085</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NFASC</td>
<td>419824</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NKD1</td>
<td>415729</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NOX4</td>
<td>419012</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NR5A2</td>
<td>395961</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NRB1</td>
<td>421985</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NUFIP2</td>
<td>417582</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NUP35</td>
<td>423998</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OBFC2A</td>
<td>424046</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene</td>
<td>Entrez ID</td>
<td>X</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>-----------</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ODZ3</td>
<td>422557</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OSBPL3</td>
<td>428432</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P4HA2</td>
<td>416326</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PAFAH1B1</td>
<td>374224</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PARD6B</td>
<td>419352</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PBRM1</td>
<td>396074</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCDHA8</td>
<td>768760</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PDCD6</td>
<td>420988</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PDE8A</td>
<td>415455</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PDS5A</td>
<td>422787</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PDXDC1</td>
<td>416413</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PHF15</td>
<td>416315</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PHF21A</td>
<td>423199</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PHLD2B</td>
<td>418418</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PHYHIPL</td>
<td>423603</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PKN2</td>
<td>424519</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PKNOX1</td>
<td>418542</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PKP4</td>
<td>424328</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLC1</td>
<td>424060</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLEKHA7</td>
<td>423069</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PNRC1</td>
<td>421817</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>POGZ</td>
<td>429985</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPAP2B</td>
<td>424666</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPARA</td>
<td>374120</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPARG</td>
<td>373928</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPP2R5C</td>
<td>423460</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPP2R5E</td>
<td>423512</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRTG</td>
<td>415407</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PTBP2</td>
<td>424479</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PTPRA</td>
<td>396060</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PUM2</td>
<td>421961</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAB11FIP4</td>
<td>427808</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAB18</td>
<td>420483</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAB1A</td>
<td>421273</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAB39</td>
<td>428085</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAB40B</td>
<td>417338</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAB5A</td>
<td>420649</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RABEP1</td>
<td>395930</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RABGAP1</td>
<td>417106</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAN</td>
<td>396193</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAP1A</td>
<td>419867</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAPH1</td>
<td>424107</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RGMB</td>
<td>427277</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene</td>
<td>GeneID</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>---------</td>
<td>----</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RHOB</td>
<td>395734</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RLBP1L1</td>
<td>421141</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RNF165</td>
<td>416382</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RNF19B</td>
<td>419669</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ROCK2</td>
<td>374182</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RRAGD</td>
<td>421814</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RSPO3</td>
<td>421715</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SACS</td>
<td>418944</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SALL1</td>
<td>395446</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCYL3</td>
<td>424396</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEC61A2</td>
<td>426369</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SFMBT1</td>
<td>415902</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SFRS2</td>
<td>396195</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SGMS2</td>
<td>422529</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SHMT1</td>
<td>416520</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SHROOM3</td>
<td>422636</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SIN3A</td>
<td>415307</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SIX4</td>
<td>395601</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLC16A12</td>
<td>423791</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLC24A4</td>
<td>772279</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLC25A6</td>
<td>374072</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLC35B3</td>
<td>420863</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLC38A1</td>
<td>417806</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLC39A10</td>
<td>429005</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMEK2</td>
<td>426680</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNX16</td>
<td>420202</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNX2</td>
<td>426797</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNX4</td>
<td>424260</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOCS1</td>
<td>416630</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOX5</td>
<td>418195</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOX6</td>
<td>423068</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPG20</td>
<td>418902</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPON1</td>
<td>395657</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPRED1</td>
<td>423292</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SRF</td>
<td>396103</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ST3GAL5</td>
<td>407775</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STAC</td>
<td>428460</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STIM2</td>
<td>422799</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STK35</td>
<td>771724</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STK39</td>
<td>424173</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STK40</td>
<td>395120</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUV420H1</td>
<td>425119</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV2B</td>
<td>427508</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SYNJ1</td>
<td>418498</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene</td>
<td>Reference</td>
<td>Status</td>
<td>Chromosome</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>-----------</td>
<td>--------</td>
<td>------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SYT14</td>
<td>428566</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TANC1</td>
<td>424330</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAOK3</td>
<td>395499</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TBC1D8B</td>
<td>422383</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TCF12</td>
<td>396334</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TCF20</td>
<td>417980</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TDG</td>
<td>395510</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TFDP1</td>
<td>418739</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TFRC</td>
<td>396191</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TGDS</td>
<td>418793</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TGFBR3</td>
<td>374251</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TGF1</td>
<td>396339</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>THBS2</td>
<td>414837</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TIAM1</td>
<td>418491</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TLE4</td>
<td>374080</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TM9SF3</td>
<td>395460</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TMED5</td>
<td>424499</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TMEM32</td>
<td>422240</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TMEM55A</td>
<td>420221</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TMOD1</td>
<td>395883</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TNS1</td>
<td>396439</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOB1</td>
<td>395225</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOB2</td>
<td>417992</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOR1B</td>
<td>417191</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOX3</td>
<td>415724</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRAM1</td>
<td>378918</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRHDE</td>
<td>427868</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRIM33</td>
<td>419883</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRRAP</td>
<td>416440</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TSHZ1</td>
<td>428513</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TXNDC4</td>
<td>420994</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TXNDC5</td>
<td>420867</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UBE2A</td>
<td>395672</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UBE2D1</td>
<td>423641</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UBE2D2</td>
<td>416137</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UBXD8</td>
<td>416225</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UGT8</td>
<td>374033</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UNC84A</td>
<td>416449</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UPF2</td>
<td>419100</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>USP24</td>
<td>424665</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>USP34</td>
<td>421193</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>USP44</td>
<td>417913</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>USP8</td>
<td>415451</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UXS1</td>
<td>418728</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene</td>
<td>ID</td>
<td>Chromosome</td>
<td>Gene</td>
<td>ID</td>
<td>Chromosome</td>
<td>Gene</td>
<td>ID</td>
<td>Chromosome</td>
</tr>
<tr>
<td>--------</td>
<td>----------</td>
<td>------------</td>
<td>------</td>
<td>----------</td>
<td>------------</td>
<td>------</td>
<td>----------</td>
<td>------------</td>
</tr>
<tr>
<td>VCL</td>
<td>396422</td>
<td>X</td>
<td>VSNL1</td>
<td>396189</td>
<td>X</td>
<td>VT11A</td>
<td>770885</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WNK1</td>
<td>427925</td>
<td>X</td>
<td>WNT7B</td>
<td>427937</td>
<td>X</td>
<td>WSB1</td>
<td>395685</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>XPR1</td>
<td>424415</td>
<td>X</td>
<td>YES1</td>
<td>396238</td>
<td>X</td>
<td>YPEL2</td>
<td>417632</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZADH2</td>
<td>421010</td>
<td>X</td>
<td>ZAK</td>
<td>424149</td>
<td>X</td>
<td>ZBTB44</td>
<td>419732</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZFHX3</td>
<td>395682</td>
<td>X</td>
<td>ZC3H7B</td>
<td>417996</td>
<td>X</td>
<td>ZDHHCC7</td>
<td>415821</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZNF654</td>
<td>418428</td>
<td>X</td>
<td>ZNF654</td>
<td>418428</td>
<td>X</td>
<td>ZNRF2</td>
<td>420678</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZNRF3</td>
<td>416919</td>
<td>X</td>
<td>ZRANB1</td>
<td>423953</td>
<td>X</td>
<td>ZRANB1</td>
<td>423953</td>
<td>X</td>
</tr>
</tbody>
</table>
Supplemental Table 3

This table shows over-represented Gene Ontology terms for proteins encoded by mRNAs with predicted target sites for FGF regulated miRs.

<table>
<thead>
<tr>
<th>Gene Ontology Terms</th>
<th>Number of genes</th>
<th>Adjusted p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>miR-let-7b</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Molecular Function</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>transmembrane receptor protein</td>
<td>2</td>
<td>1.33E-07</td>
</tr>
<tr>
<td>serine/threonine kinase activity</td>
<td>2</td>
<td>1.33E-07</td>
</tr>
<tr>
<td>platelet-derived growth factor binding</td>
<td>2</td>
<td>5.12E-07</td>
</tr>
<tr>
<td>transforming growth factor beta receptor activity</td>
<td>2</td>
<td>1.36E-06</td>
</tr>
<tr>
<td>receptor signaling protein serine/threonine kinase activity</td>
<td>2</td>
<td>2.07E-05</td>
</tr>
<tr>
<td>manganese ion binding</td>
<td>3</td>
<td>2.14E-05</td>
</tr>
<tr>
<td>SMAD binding</td>
<td>2</td>
<td>0.001933223</td>
</tr>
<tr>
<td>extracellular matrix structural constituent</td>
<td>2</td>
<td>0.013059435</td>
</tr>
<tr>
<td><strong>miR-9</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Molecular Function</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>actin binding</td>
<td>5</td>
<td>0.007829822</td>
</tr>
<tr>
<td>mRNA binding</td>
<td>2</td>
<td>0.007829822</td>
</tr>
<tr>
<td>vascular endothelial growth factor receptor activity</td>
<td>3</td>
<td>0.012885294</td>
</tr>
<tr>
<td><strong>Biological Process</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miRNA catabolic process</td>
<td>2</td>
<td>7.14E-06</td>
</tr>
<tr>
<td>cell adhesion</td>
<td>13</td>
<td>7.06E-05</td>
</tr>
<tr>
<td>pre-microRNA processing</td>
<td>2</td>
<td>0.000680037</td>
</tr>
<tr>
<td>RNA 3'-end processing</td>
<td>2</td>
<td>0.002842119</td>
</tr>
<tr>
<td><strong>miR-19b</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Molecular Function</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>voltage-gated ion channel activity</td>
<td>2</td>
<td>0.051678371</td>
</tr>
<tr>
<td><strong>miR-107</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Molecular Function</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>transforming growth factor beta receptor activity</td>
<td>2</td>
<td>9.35E-07</td>
</tr>
<tr>
<td>transcription corepressor activity</td>
<td>3</td>
<td>0.000476577</td>
</tr>
<tr>
<td>SMAD binding</td>
<td>2</td>
<td>0.002195363</td>
</tr>
<tr>
<td><strong>Biological Process</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>palate development</td>
<td>3</td>
<td>0.027029621</td>
</tr>
<tr>
<td>negative regulation of osteoblast differentiation</td>
<td>2</td>
<td>0.048262035</td>
</tr>
</tbody>
</table>
### miR-130b

**Molecular Function**
- transforming growth factor beta receptor activity, type I: 2, 9.94E-11
- transforming growth factor beta binding: 2, 2.14E-09
- transmembrane receptor protein serine/threonine kinase activity: 2, 1.51E-05
- transforming growth factor beta receptor activity: 2, 4.63E-05
- receptor signaling protein serine/threonine kinase activity: 2, 0.000120546
- kinase activity: 6, 0.001151488
- SMAD binding: 2, 0.021052872
- receptor activity: 9, 0.024076191
- metal ion binding: 9, 0.025489948

**Biological Process**
- acrosome assembly: 2, 0.000327585
- pharyngeal system development: 2, 0.015459557

### miR-218

**Molecular Function**
- protein heterodimerization activity: 4, 0.001139013

**Biological Process**
- positive regulation of DNA replication: 2, 0.025600114
Supplemental Figure 1.

This figure shows the structure of proposed novel miRs detected using a multi-species microRNA microarray. Predicted secondary structures are shown.
REFERENCES


Heo, I., Joo, C., Choo, J., Ha, M., Han, J., et al., 2008. Lin28 mediates the terminal uridylation of let-7 precursor microRNA. Molecular Cell 32, 276-284.


