

ETD Archive

---

2010

## Characterization of a MicroRNA Harboring Intron for Pre-mRNA Splicing and MicroRNA Processing

Neha Aggarwal  
*Cleveland State University*

Follow this and additional works at: <https://engagedscholarship.csuohio.edu/etdarchive>

 Part of the [Biology Commons](#)

[How does access to this work benefit you? Let us know!](#)

---

### Recommended Citation

Aggarwal, Neha, "Characterization of a MicroRNA Harboring Intron for Pre-mRNA Splicing and MicroRNA Processing" (2010). *ETD Archive*. 674.

<https://engagedscholarship.csuohio.edu/etdarchive/674>

This Thesis is brought to you for free and open access by EngagedScholarship@CSU. It has been accepted for inclusion in ETD Archive by an authorized administrator of EngagedScholarship@CSU. For more information, please contact [library.es@csuohio.edu](mailto:library.es@csuohio.edu).

**CHARACTERIZATION OF A MICRORNA  
HARBORING INTRON FOR PRE-MRNA SPLICING  
AND MICRORNA PROCESSING**

**NEHA AGGARWAL**

Master of Science in Biochemistry

Panjab University

August, 2007

Submitted in partial fulfillment of requirements for the degree

**MASTER of SCIENCE in BIOLOGY**

at the

Cleveland State University

May, 2010

This thesis has been approved for  
The Department of Biological, Geological and Environmental Sciences and for  
The College of Graduate Studies of  
Cleveland State University  
by

\_\_\_\_\_ Date: \_\_\_\_\_

Girish Shukla, Ph.D., BGES, CSU  
Major Advisor

\_\_\_\_\_ Date: \_\_\_\_\_

Crystal M. Weyman, Ph.D., BGES, CSU  
Advisory Committee Member

\_\_\_\_\_ Date: \_\_\_\_\_

Barsanjit Mazumder, Ph.D., BGES, CSU  
Advisory Committee Member

## ACKNOWLEDGEMENTS

*None of us got where we are solely by pulling ourselves up by our bootstraps. We got here because somebody - a parent, a teacher, a guide - bent down and helped us pick up our boots.*

~Thurgood Marshall

First of all, I would like to thank the Almighty for all His blessings without which nothing would have been possible.

It seems simple to name all the people who helped me to get this done, but it is tough to thank them enough. It is with a deep sense of gratitude that I thank my mentor and guide, Dr. Girish Shukla, for his close supervision and precious guidance.

I extend my sincere thanks to Dr. Kavleen Sikand, who always spared her precious time to share valuable insights. Her support and guidance were indeed of great help and strength. The help and support I received from my dear lab members and friends, their love and care always kept me going.

I would also like to extend thanks to Dr. Richard Padgett and Ms. Rosemary Dietrich for help and support in performing the in vitro splicing assay. I want to pay special thanks to my committee members, Dr. Crystal Weyman and Dr. Barsanjit Mazumder.

I fall short of words to express my thanks to my parents and family. Their boundless love has always been a source of inspiration to me and reach where I stand today.

# CHARACTERIZATION OF A MICRORNA HARBORING INTRON FOR PRE-MRNA SPLICING AND MICRORNA PROCESSING

NEHA AGGARWAL

## ABSTRACT

Nuclear pre-mRNA splicing is an important step in eukaryotic RNA processing and can yield a variety of transcripts. Further, it is known that nuclear pre-mRNA introns are of two types, U2-dependent and U12-dependent. Recently discovered, microRNAs (miRNA) are short, endogenous, 20-22 nucleotide long, non-coding RNA molecules that base pair with target mRNAs to modulate translation. Current evidence suggests that processing of intronic miRNA does not affect the splicing. Since, intronic miRNAs are involved in essential cellular processes, we surmised if splice sites of miRNA coding introns are flexible to support productive precursor and mature miRNA processing.

We constructed a minigene reporter consisting exons 26-29 including introns from MYH6 gene and an intronic miRNA, miR-208 harbored in intron 27. Using *in vivo* and *in vitro* methods, we examined splicing of MYH6 intron 27 and processing of miR-208. The predicted stem-loop structure of miR-208 and the 5' splice site of the intron 27 were mutated. Stem-loop mutations had no recordable splicing defect of intron 27. The 5'SS mutant activated a cryptic splicing event using a sub-optimal splice site located 50 nucleotides upstream of wild-type splice site. Interestingly, conversion of U2-type 5' splice site to consensus U12-type 5' splice site did not affect *in vivo* splicing of the intron, suggesting potential conversion of intron to U12-type, albeit without a consensus U12-

type branch site sequence. However, miRNA processing was affected as established by in vitro splicing assay and qRT-PCR analysis of miR-208.

We mutated nucleotides from position 3 to 7 in intron 27 to study if changes in splice site will allow miRNA processing to occur and if splicing will also be affected. Our data indicates significant changes in spliced phenotype as well as in miR-208 expression level.

Since, all known intronic miRNAs are harbored in U2-dependent introns of mammalian genes, we wanted to study if miRNA processing will be compatible with U12-dependent intron. With this aim, we incorporated miR-208 in a U12-dependent intron. However, it appears that spliceosome and microprocessor cannot act together in a U12-dependent intron. This is first time that a miRNA has been incorporated in a U12-dependent intron to study its processing and also if splice sites are plastic for miRNA processing.

## TABLE OF CONTENTS

ABSTRACT .....	iv
LIST OF FIGURES .....	ix
LIST OF TABLES .....	xi
LIST OF ABBRAVIATIONS .....	xii
CHAPTERS	
I. INTRODUCTION	
1.1 Gene Expression .....	1
1.2 Nuclear Precursor-mRNA (pre-mRNA) Splicing .....	4
1.2.1 Intron: Types, Sequence Conservation and Splicing .....	5
1.2.2 Process of Splicing .....	7
1.2.3 Methods to Study Pre-mRNA Splicing .....	10
1.3 MicroRNAs .....	12
1.3.1 miRNA Biogenesis .....	12
1.3.2 Genomic Locations of miRNA .....	14
1.3.3 Intronic miRNAs .....	15
1.4 Model System: Myosin Heavy Chain6 and miR-208 .....	16

## II. MATERIALS AND METHODS

2.1	Minigene Construction .....	21
2.2	CHO Cell Maintenance and Transfection .....	24
2.3	Total RNA Isolation .....	24
2.4	Reverse Transcriptase-Polymerase Chain Reaction (RT-PCR) .....	25
2.5	<i>In vivo</i> Splicing Phenotype Analysis for MYH6 miR-208.....	26
2.6	<i>In vivo</i> Splicing Phenotype Analysis for p120 minigene .....	27
2.7	<i>In vitro</i> Transcription Assay .....	27
2.8	<i>In vitro</i> pre-mRNA Splicing in cell-free Extracts .....	29
2.9	RT and qRT-PCR Analysis of miR-208 Expression .....	30

## III. RESULTS AND DISCUSSION

3.1	Construction of MYH Minigene Reporter System.....	32
3.2	Expression of MYH in Mouse Tissues.....	33
3.3	Optimization of Transient Transfection Conditions.....	34
3.4	<i>In vivo</i> Splicing Phenotypes of miRNA Stem-Loop Mutants .....	36
3.5	<i>In vitro</i> Splicing Assay for Stem-Loop Mutants .....	38
3.6	<i>In vivo</i> Splicing Phenotypes of MYH6 Intron27 5'splice site Mutants ...	42
3.7	<i>In vitro</i> Splicing Assay for 5'splice site Mutants .....	46
3.8	Quantitation of Mature hsa-miR-208 Expression.....	47

3.9	<i>In vivo</i> Splicing Phenotypes of Point Mutants.....	50
3.10	Quantitation of Mature hsa-miR-208 Expression in point Mutants .....	53
3.11	<i>In vivo</i> Splicing Phenotypes for U12-dependent miRNA Mutants .....	54
3.12	Quantitation of miR Expression in U12-dependent miRNA Mutants....	57
IV. CONCLUSIONS AND FUTURE DIRECTIONS		
4.1	Conclusions .....	59
4.2	Future Directions .....	62
BIBLIOGRAPHY.....		63
APPENDIX.....		72

## LIST OF FIGURES

1. Levels of gene expression control .....	3
2. Mechanism of RNA splicing .....	7
3. Major and minor spliceosomal pathways .....	8
4. miRNA biogenesis pathway .....	13
5. Representation of MYH6 gene and miR-208 .....	16
6. Regulation of cardiac contractility by miR-208 .....	17
7. Structure of the minigene reporter construct and location of primers .....	33
8. MYH6 expression in mouse organs.....	33
9. Optimization of transient transfection conditions .....	34
10. Optimization of primer sets located on MYH6 minigene construct.....	36
11. Structure of predicted miR-208 stem loop .....	37
12. <i>In vivo</i> splicing patterns of miRNA mutants .....	38
13. <i>In vitro</i> spliced phenotypes of stem loop mutants .....	39
14. Consensus 5' splice site of U2-type intron and intron 27 of MYH6 .....	42
15. <i>In vivo</i> splicing patterns of MYH6 intron 27 5' splice site mutants .....	43
16. Affect of 5' splice site mutation .....	45
17. <i>In vitro</i> spliced phenotypes of 5' splice site mutants.....	47
18. qRT-PCR analysis of miR-208 from stem-loop and 5' splice site mutants .....	48

19. qRT-PCR analysis of miR-208 from stem-loop and 5' splice site mutants .....	49
20. Pictogram representing 5' splice site of miRNA coding introns .....	50
21. Intron 27 of MYH6 showing 5' splice site point mutations .....	51
22. <i>In vivo</i> splicing patterns of 5' splice site point mutations .....	52
23. qRT-PCR expression of miR-208 in point mutants.....	54
24. Intron F of p120 minigene with miR-208 and mutations .....	55
25. <i>In vivo</i> splicing patterns of p120 mutants .....	56
26. qRT-PCR expression of miR-208 in p120 mutants.....	57

## LIST OF TABLES

1. snRNAs involved in splicing.....	6
-------------------------------------	---

## LIST OF ABBRAVIATIONS

ATP	Adenosine triphosphate
cDNA	complementary DNA
CHO	Chinese Hamster Ovary
CTP	Cytidine triphosphate
DMEM	Dulbecco's Modified Eagle's Medium
DNA	Deoxyribonucleic Acid
dNTP	deoxy Nucleotide triphosphate
EDTA	Ethylene Diamine Tetraacetic Acid
FBS	Fetal Bovine Serum
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GTP	Guanosine triphosphate
miRNA/ miR	microRNA
mRNA	messenger RNA
MYH	Myosin Heavy Chain
PAGE	Polyacrylamide Gel Electrophoresis
PCR	Polymerase Chain Reaction
Polybrene	hexadimethrine bromide
Pre-mRNA	Precursor messenger RNA

qRT-PCR	Quantitative Reverse Transcriptase PCR
RISC	RNA Induced Silencing Complex
RNA	Ribonucleic Acid
RT	Reverse Transcription
RT-PCR	Reverse Transcriptase- Polymerase Chain Reaction
snoRNA	small nucleolar RNA
snRNA	small nuclear RNA
TAE	Tris-acetate-EDTA
TBE	Tris-borate-EDTA
TRIS	Tris(hydroxymethyl) aminomethane
UTP	Uridine triphosphate
UTR	Untranslated Region
WT	Wild type

# **CHAPTER I**

## **INTRODUCTION**

### **1.1 Gene Expression**

The “central dogma” of developmental biology holds that growth and development is governed by regulated changes in gene expression. Gene expression involves transcription of DNA (genes) to RNA and then translation from RNA to proteins that directly affect cellular functions. Both transcription and translation processes are the basic fundamentals of molecular biology and both processes are strictly regulated at various steps (Hartwell et al, 2004). This regulation at different levels and stages in a cell leads to a differential expression and hence to different cell and tissue types.

Gene expression is influenced by a number of cellular processes, such as transcriptional regulation, RNA processing, compartmentalization, translational regulation etc (Figure 1). Each eukaryotic gene has its own promoter and its transcription is regulated by transcriptional activators or transcription factors. Eukaryotic genes also have one or more enhancers generally upstream of the promoter. The promoter is responsible for initiating low levels of transcription and determining the transcription start site, enhancers are responsible for increasing ("enhancing") transcription levels, and they are also responsible for regulating cell or tissue-specific transcription.

For example, human  $\beta$  globin gene cluster consists of 5 functionally active genes ( $\epsilon$ ,  $G\gamma$ ,  $A\gamma$ ,  $\delta$  and  $\beta$  genes) expressing at different stages of life, from embryonic to adulthood (Li et al, 1990). It has a ~12 kb enhancer region with binding sites for GATA-1, NF-E2, AP-1 and some other transcriptional activators. GATA-1 binds at embryonic stage (Takahashi et al, 1997), preferentially expressing the  $\epsilon$  gene. In fetal DNA,  $G\gamma$  and  $A\gamma$  are more strongly expressed and in adults, expression is switched to  $\beta$  gene, while the  $\delta$  gene is weakly expressed (Li et al, 2002).

Histone regulation and chromatin remodeling are also involved in regulation of transcription. Chromatin is condensed to inactive form- heterochromatin. Covalent modification such as methylations, acetylations, deacetylations etc of histones occurs, reducing transcription. Histone Acetyltransferases (HATs) acetylates lysine residues in histone proteins, leads to chromatin remodeling as well as making DNA accessible for transcription (Grunstein, 1997).

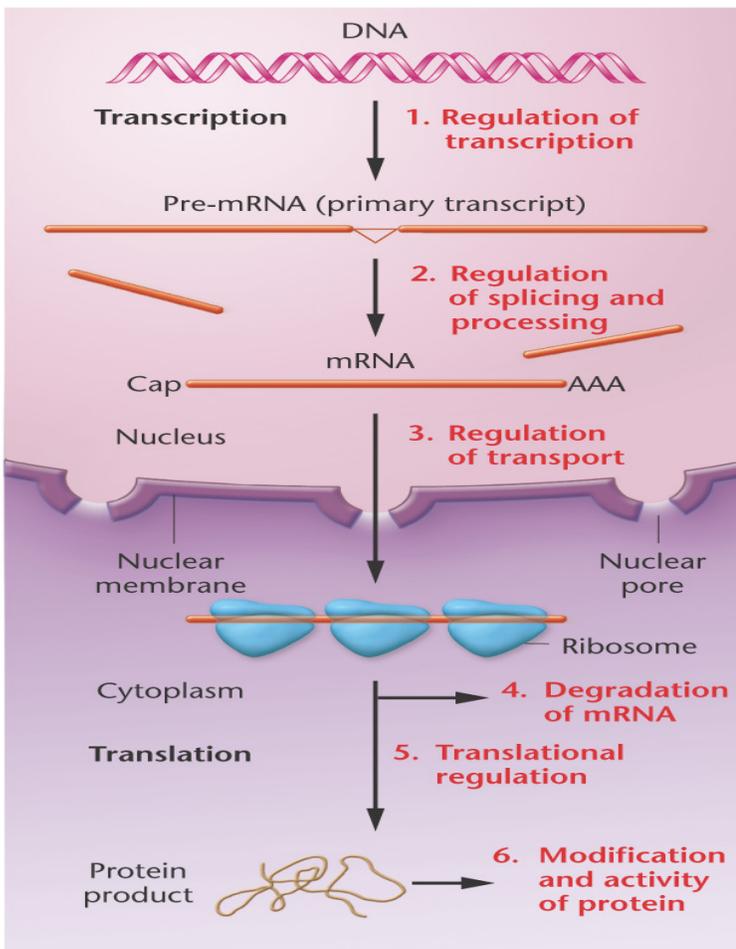
Most of the gene-expression changes which are involved in development are associated with changes in the level of mRNAs and their regulation (Figure 1). RNA polymerase II (RNA pol II) transcribes majority of protein-coding genes. The pre-mRNA synthesized by RNA pol II is processed before it can be translated. RNA processing involves 5' capping, 3' modification by addition of a 3' poly (A) tail, and intron splicing (Proudfoot et al, 2002; Shatkin and Manley, 2000; Staley and Guthrie, 1998). If pre-mRNA is capped and polyadenylated, it would be transported to the cytoplasm for translation otherwise it would be retained in the nucleus for degradation. Introns are spliced in the nucleus by formation of a series of dynamic complexes governing introns to be spliced (Moore et al, 1993). This process of splicing affects the function of the protein produced.

Some genes have exons that might be spliced, producing a different transcript and a functionally different protein for example; the calcitonin gene produces calcitonin hormone in parafollicular cells of thyroid (Copp and Cheney, 1962; Hirsch et al, 1963) and a neurotransmitter (CGRP) in neurons (Brian et al, 1985). This process is called alternative splicing and is now known to be involved in a number of diseases.

mRNAs from different genes have their approximate lifespan encoded in them. The information for lifespan is found in the 3' UTR. The sequence AUUUA, when found in the 3' UTR, is a signal for early degradation and hence a short lifetime (Chen et al, 2001). The

more times this sequence is present, the shorter the lifespan of the mRNA.

mRNA is also regulated at translational level. Translation in eukaryotes is divided into three stages (i) initiation (governed by initiation factors, eIFs), (ii)



Copyright © 2006 Pearson Prentice Hall, Inc.

**Figure 1: Levels of gene expression control:** Gene expression is regulated at various steps during transcription as well as translational processes.

elongation (governed by elongation factors, eEFs) and

(iii) termination (governed by release factors). Translational regulation is exerted mostly at the first stage i.e. initiation, where the AUG start codon (sometimes GUG) is identified. It is generally regulated by eIF2 or eIF4.

Recently, it has been shown that RNA pol II is involved in transcription of microRNAs, a novel class of small RNA molecules regulating translation (discussed in detail in section 2.3). These are small RNAs which modulate translation, mRNA degradation and perhaps other processes. Depending upon complementarity of miRNA with the 3'UTR of mRNA transcript, mRNA transcript is either degraded or translation is inhibited.

## **1.2 Nuclear Precursor mRNA (pre-mRNA) Splicing**

Pre-mRNA splicing involves removal of intervening introns and joining of exons forming a mature mRNA. This process appears to occur co-transcriptionally and is an essential step in RNA processing. When “interrupted” gene produces primary transcription product a heterogeneous nuclear RNA molecule (pre-mRNA), both exons and introns are represented. Introns, however, are removed from the primary transcript during the processing of pre- mRNA in specific splicing reactions. Splicing is usually constitutive, which means that all exons are joined together in the order in which they occur in the heterogeneous nuclear RNA (Sharp, 1985). In many genes, however, alternative splicing has also been observed, in which exons may be combined in some other way. It is a process by which certain exons are skipped and hence the mature mRNA formed is different. This leads to formation of different proteins. Alternative splicing is another mechanism by which a high degree of diversity and variance can be achieved. Alternative splicing is an

important aspect of gene regulation and about 75% of human genes encode pre-mRNAs that are alternatively spliced. Numerous diseases are associated with alternative splicing and splicing isoforms. Additionally, the proteins encoded by alternative isoforms of a particular gene have distinct biological functions (Black, 2003).

### **1.2.1 Intron: Types, Sequence Conservation and Splicing**

After the discovery of split genes in 1977, conserved sequences at both ends of introns were recognized. These sites are called 5' splice site and 3' splice site respectively. Another conserved sequence called the branch point region is located within the intron. It is usually 20–40 nucleotides upstream of the 3' splice site. The branch point is almost always an adenosine and can rarely be guanine. The sequences at intron ends were GT at the 5' splice site and AG at the 3' splice site and led to formulation of GT-AG rule (Breathnach et al, 1978). This rule holds true in most cases but a number of exceptions to this have been found. For example, GC is found at 5' splice site in case of intron 7 of Human laminin  $\alpha$ -4 (LAMA4) gene (Richards et al, 1997). After about 10 years, it was found that intron F of the gene encoding human P120, a proliferating cell nucleolar antigen, and intron 7 of the gene encoding human CMP, a cartilage matrix protein have AT and AC at the 5' and 3' intron ends respectively, instead of GT and AG (Jackson, 1991). Later on, intron 6 of the gene encoding Rep-3, a DNA repair protein, and intron 2 of the gene encoding Prospero, a *Drosophila melanogaster* homeodomain transcription factor were shown to have AT-AC ends (Hall and Padgett, 1994). It was also shown that AT-AC introns have highly conserved 5' splice site and presumptive branch site sequence elements that are not present

in the GT-AC introns, namely, ATATCCTY and CCTTRAY, respectively (Hall and Padgett, 1994).

The splice sites and branch point region interact with specific RNA and protein factors to engage the spliceosome. The spliceosome is formed by small nuclear RNAs (snRNAs) and a number of protein constituents called ribonuclear proteins (RNPs). snRNAs and RNPs are together called snRNPs (or snurps). Spliceosome is a dynamic complex structure which undergoes rearrangements and finally leading to catalysis (Konarska and Sharp, 1987).

<b>Major (U2-dependent) Spliceosome</b>	<b>Minor (U12-dependent) Spliceosome</b>
U1	U11
U2	U12
U4	U4atac
U5	U5
U6	U6atac

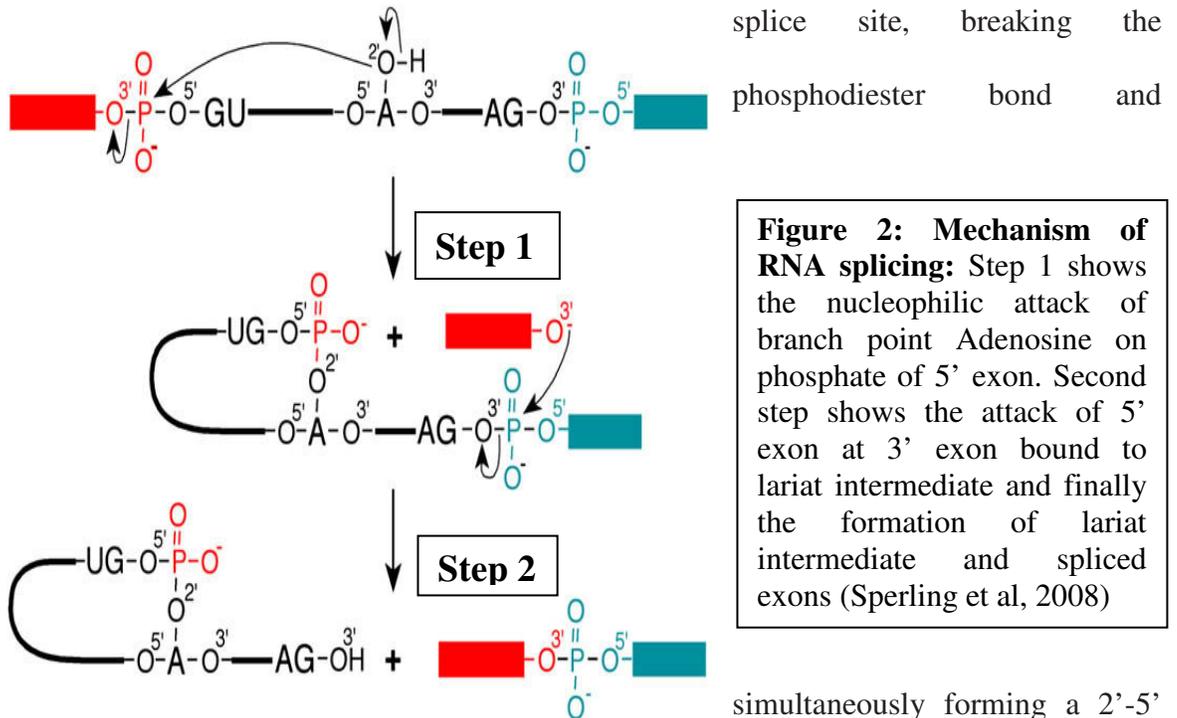
**Table 1: snRNAs involved in Splicing**

Depending on the conserved sequences in an intron, two spliceosomal systems are known to exist in a eukaryotic system. Both the systems recognize 5' splice site, branch point sequence and 3' splice site flanking the intron in a gene. However, the conserved consensus sequences recognized are different. But the chemistry involved in splicing through both pathways is same. Majority of introns are spliced by U2-dependent pathway

(major pathway) involving U1, U2, U4, U5 and U6 snRNAs (Table 1) (Tycowski et al, 2006). Generally GT-AC introns and GC-AG introns are spliced by major pathway however, some exceptions exist (Dietrich et al, 1997). AT-AC introns are spliced by the U12-dependent pathway (minor pathway) involving U11, U12, U4atac, U5 and U6atac snRNAs (Table 1) (Hall and Padgett, 1996; Tarn and Steitz, 1996 a, b).

### 1.2.2 Process of Splicing

Splicing occurs by formation of a large spliceosomal complex and involves two trans-esterification reactions. The splicing mechanism involves 2 steps: (i) cleavage of the 5' exon-intron phosphodiester bond and formation of a new 2'-5' bond resulting in an intron lariat and (ii) exon ligation. In the first step, the branch point adenosine attacks the 5'

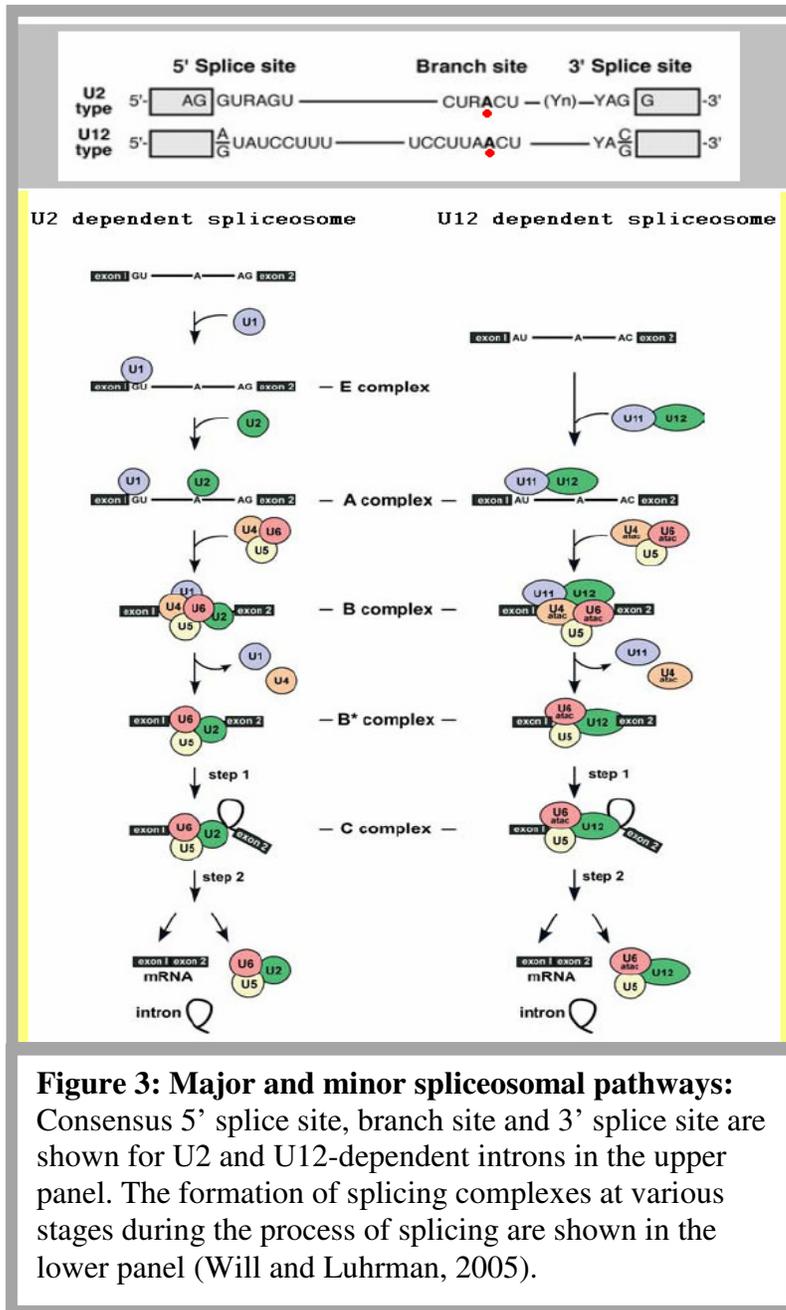


phosphodiester linkage between the branch point and the 5' terminal nucleotide of the intron. The products of the first reaction are free 5' exon and 3' exon bound to lariat intron.

After a conformational change within the spliceosome, the second step takes place. Here, the newly released 3' hydroxyl of the 5' exon, created by the first step, attacks the 3' splice junction, breaking that phosphodiester bond while forming a new phosphodiester bond between the 5' exon and 3' exon (Sperling et al, 2008). The products of the second step are

ligated exons and the free intron, released in the form of a lariat intermediate (Figure 2). These reactions involve rearrangement of snRNPs and formation of spliceosome.

Spliceosome formation is a complex process in which snRNAs and protein components of snRNPs are sequentially loaded on the intron to be spliced (Nilsen, TW, 1998; Will and Luhrmann, 2001, 2006). E complex is the first to be formed when U1 binds to the GU sequence at the 5' splice site (Zhuang



**Figure 3: Major and minor spliceosomal pathways:** Consensus 5' splice site, branch site and 3' splice site are shown for U2 and U12-dependent introns in the upper panel. The formation of splicing complexes at various stages during the process of splicing are shown in the lower panel (Will and Luhrman, 2005).

and Weiner, 1986), along with accessory proteins/enzymes such as U2AF (Figure 3). In next step, U2 binds to the branch site; this is called the A Complex (Wu and Manley, 1989). A stable association of U2 snRNA with branch site requires ATP to form the pre-spliceosome. B complex is formed following A complex, by the binding of U5/U4/U6 tri-snRNP (Konarska and Sharp, 1987). The U5 snRNA binds exons at the 5' splice site and U6 snRNA binds to U2 with release of U1 snRNA. Then, U5 shifts from exon to intron and U6 binds at the 5' splice site by RNA: RNA interactions (Madhani and Guthrie, 1994). For the C complex formation, U6 and U2 catalyze trans-esterification reaction and U4 is released from the complex (Lewis and Seraphin, 1993). In next step, with U2, U5 and U6 still bound, exons are spliced (Nilsen, 1998; Staley and Guthrie, 1998). The final products after splicing are spliced exons and lariat intron. Snurps are recycled again used for the next round of splicing (Moore et al, 1993).

snRNAs involved in U12-dependent pathway are functional analogs of their U2 dependent counterparts. U11 is a functional analogue of U1 snRNA, U12 of U2, U4atac of U4 and U6atac of U6. U5 snRNA is common to both the pathways. There are a few differences in splicing by minor spliceosomal pathway (Figure 3). In case of minor pathway, U11/U12 di-snRNP is formed and binds to 5' splice site and branchpoint adenosine resulting in formation of A complex. U4atac/U5/U6atac tri-snRNP then binds the intron forming B complex (Tarn and Steitz, 1996(b)). Transesterification reactions for U12-dependent pathway are similar to U2-dependent pathway (Nilsen, 1996).

### 1.2.3 Methods to Study pre-mRNA Splicing

Pre-mRNA splicing is generally studied by two approaches, *in vivo* and *in vitro* methods. These studies have demonstrated the roles of sequence elements in the pre-mRNA and the snRNAs in various interactions and the stepwise assembly of the spliceosome. The *in vivo* approach involves transient transfection of wild-type or mutant minigenes and then analyzing the spliced variants by RT-PCR method. For *in vivo* studies, a suppressor system is developed in which few mutations are introduced in minigene and compensatory base changes in snRNAs are made, allowing minigene to splice as good as wild type. It is known that the 5' end of U1 snRNA and the intronic 5' splice site interact via base pairing (Zhuang and Weiner, 1986; Siliciano and Guthrie, 1988). For major pathway, a compensatory base change in U1 snRNA can suppress 5' splice site mutation (Zhuang and Weiner, 1986), allowing splicing to occur as wild type. In case of AT-AC pre-mRNA splicing (minor splicing), suppressor assay system has also been established. In an AT-AC intron, it has been shown that CC at 5<sup>th</sup> and 6<sup>th</sup> position in the conserved 5' splice site sequence, ATATCCTY, are absolutely necessary for splicing to occur by the U12-dependent pathway (Hall and Padgett, 1996; Kolossova and Padgett, 1997; Incorvaia and Padgett, 1998). Suppressor AT-AC snRNAs can be cotransfected with AT-AC intron mutants to study base pairing interactions between intron and snRNAs.

An *in vitro* splicing method involves splicing of a pre-mRNA transcript in presence of nuclear extract. In an *in vitro* transcribed mRNA transcript is called substrate which upon *in vitro* splicing reaction leads to formation of free lariat intermediate, spliced exons, free 5' exon, lariat intermediate with 3' exon and all these products can be detected on the gel. The *in vitro* systems for AT-AC pre-mRNA splicing make possible to study the

biochemical mechanisms of the reaction. In vitro splicing of has been established for both U2-dependent pathway (Moore et al, 1993) and U12-dependent pathway (Tarn and Steitz, 1996 a). Model minigene system for U2-dependent pathway is human adenovirus-2 gene and for U12-dependent pathway is intron F of p120 gene. Both U2-dependent and U12-dependent splicing occurs in two steps involving trans-esterification reactions and the results for splicing are similar in both in vitro splicing systems.

Both in vivo and in vitro methods have certain advantages and disadvantages. Almost every construct transfected in cells splices upon transfection in contrast to in vitro splicing methods where the splicing ability is limited. Utilizing the in vivo system, different cell lines can be tested. However, in vivo splicing is prone to indirect effects since intact cells are used. Using the in vitro approach, we can differentiate whether splicing occurs by major (U2-dependent) or minor pathway (U12-dependent) by blocking one or more snRNAs. This inactivation of snRNAs is achieved using 2'O-methyl oligos. For blocking U2-dependent splicing, inactivation of U2 snRNAs in the nuclear extract is required to detect splicing. However, inactivation of U2 snRNA in addition to U12 snRNA is required to block the U12-dependent splicing (Tarn and Steitz, 1996 a). It is believed that since, U2 is about a hundred times more abundant than U12, U2 competes with U12 for binding to the AT-AC branch site (Montzka and Steitz, 1988). Another advantage of the in vitro system is enabling to study the kinetics of splicing reaction since an in vitro reaction can be stopped at different time points and studied for complex formation. Also different snRNAs can be blocked (using 2'O-methyl oligo) and intermediate products can be studied, giving insight on the process.

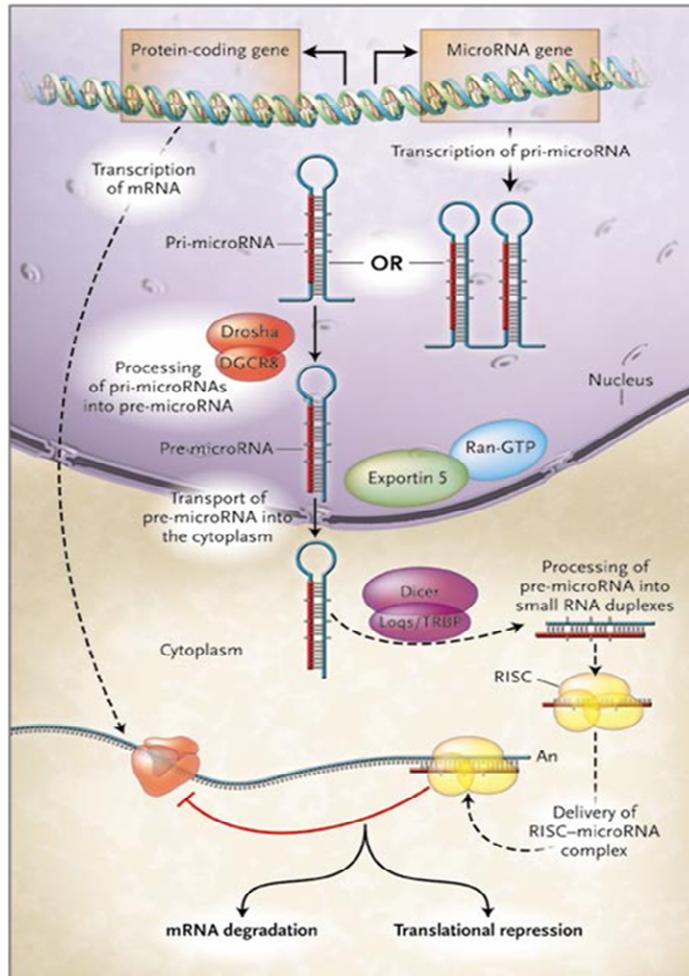
### **1.3 MicroRNAs**

MicroRNAs (miRNAs) are short, endogenous, 20-22 nucleotide long, non-coding RNA molecules that can regulate gene expression by base pairing with target mRNAs (Lee et al, 1993; Lee and Ambros, 2001; Lagos-Quintana et al, 2001; Kim, 2005; Kim and Nam, 2006). The founding microRNA *lin-4* was discovered more than 25 years ago in the nematode *Caenorhabditis elegans* (Lee et al, 1993). Large numbers of miRNAs have been identified since then, making it one of the largest growing fields in the RNA world. To date, more than 10884 mature miRNAs have been annotated and 721 of these are human microRNAs. They are known to exist in 115 different species, from viruses, fungi to plants and vertebrates (mirBase, 14th release; Griffiths-Jones, 2008). MiRNAs are computationally found across the genome and then they are confirmed experimentally by cloning (Bentwich, 2005). Targets for some of miRNAs are also found computationally first and later on confirmed experimentally (Bentwich, 2005). Many miRNAs are conserved across the species indicating their evolutionary significance. MiRNAs are known to have important functions in cellular biology for example; miR 134 is involved in embryonic stem cell differentiation (Tay et al, 2008), miR 17-92 cluster in adipogenesis (Wang et al, 2008) and in cell proliferation and cancer as well (Hayashita et al, 2005). miRNA involvement is also implicated in a number of diseases such as, miR 9 and miR 223 for ovarian cancer (Laios et al, 2008), miR 375 for Diabetes (Poy et al, 2004, 2009).

#### **1.3.1 miRNA Biogenesis**

MiRNA biogenesis refers to the processing microRNA from primary transcript to mature form (Figure 4). Majority of miRNAs are transcribed by RNA polymerase II

(Cullen et al, 2004; Lee et al, 2004; Lee et al, 2002) and some through RNA polymerase III (Borchert et al, 2006). The transcript formed is called as primary miRNA and has a stem-loop structure of about 125 nucleotides. These are processed in a number of steps in both,



nucleus and cytoplasm to yield mature miRNA. In nucleus, pri-miRNA is excised by Drosha, an RNase III-type enzyme (Lee et al, 2003). Drosha functions as a large protein complex called Microprocessor complex that contains DGCR8/Pasha as well as several splicing factors. Cleavage of pri-miRNAs by the microprocessor complex (Drosha) results in the production of shorter, stem-loop-shaped RNAs called

**Figure 4: miRNA biogenesis pathway:** MiRNA genes are transcribed by RNA polymerase II to generate the pri-miRNAs. The initial cropping step by the Drosha-DGCR8 complex, results in formation of pre-miRNA (~70-nt), which is exported by the Exp5-Ran complex. Upon export, Dicer participates in the second step of processing to produce miRNA duplexes. The duplex is separated and one strand is selected as mature miRNAs (<http://www.euchromatin.com/ChenCZ01F1.jpg>)

precursor miR (pre-miR). These RNAs are ~70 nucleotides long and are exported to the cytoplasm by Exportin 5/Ran-GTP complex (Bohnsack et al, 2004; Yi et al, 2003). In the

cytoplasm, pre-miRs are cleaved by the cytoplasmic RNase III-type enzyme, Dicer (Ketting et al, 2001). The cleavage reaction yields short RNA duplexes and, subsequently one strand of each duplex is incorporated into the RNA-induced silencing complex (RISC). RISC complex delivers miRNA to its target mRNA through base pairing with the 3' untranslated regions (UTR). If the miRNA is fully complementary to the 3'UTR of target mRNA, it leads to mRNA degradation and if it has incomplete complementarity with target, it induces translational repression.

### **1.3.2 Genomic Locations of miRNA**

Based on genomic location relative to protein-coding gene locus, miRNA can be divided into intergenic or intronic miRNAs. Intergenic miRNAs are found in region between the two genes. They generally are independent transcription units, and have their own promoters, transcript sequences and terminator sequences; such as *lin-4* and *let-7* miRNAs from *C. elegans* (Bartel, 2004; Kim, 2005). On the other hand, intronic miRNAs are located in the introns of a protein-coding gene. They do not have their own promoters but are transcribed from the promoter of the protein gene (Kim, 2005). Intronic miRNA is similar to the intergenic miRNA, structurally and functionally; difference being in location and requirement for RNA polymerase II for transcription for most intronic miRNAs (Cai et al, 2004). However, RNA polymerase III has also been shown to transcribe miRNAs from repetitive sequences where the human C19MC 5' Alu function as miRNA promoters (Borchert et al, 2006).

### 1.3.3 Intronic miRNAs

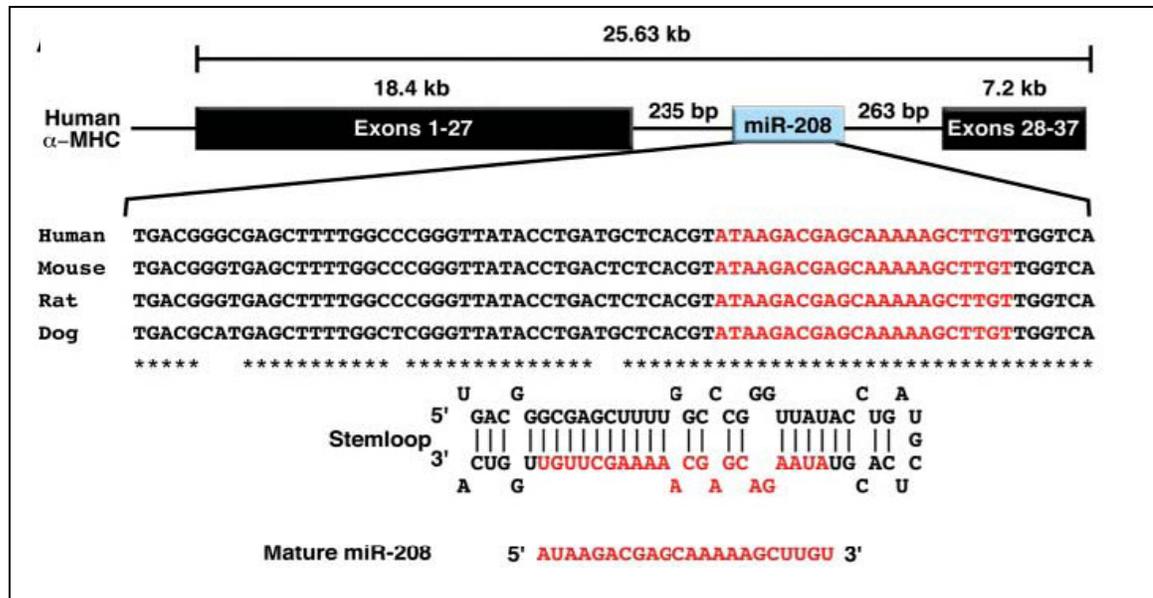
Much is known about miRNA biogenesis and the post-transcriptional effects it has on protein coding genes, but very less is known about the transcriptional regulation of intronic miRNAs. Recently with advances in the field, it was found that about 80% of miRNAs are encoded in the intronic regions of protein-encoding or non-coding genes in mammals (Rodriguez et al, 2004; Kim and Kim, 2007) and 75% in *Xenopus tropicalis* (Tang and Maxwell, 2008). This also indicates that the miRNA harbored in introns might be evolutionarily conserved in vertebrates. It was earlier thought that all miRNAs have their own promoter regions but the expression of intronic miRNAs largely coincides with that of their host genes (Baskerville and Bartel, 2005). miRNAs are transcribed from their host gene promoters and processed post-transcriptionally from the host pre-mRNAs (Sikand et al, 2009). According to computational analysis miRNAs may regulate about 60% of protein encoding genes, suggesting some important, yet undiscovered regulatory mechanisms.

The fact that mature mRNAs and intronic miRNAs are generated from same transcript may have some unknown significance. It was shown that the miRNA-containing intron is spliced out slower than the other introns (Kim and Kim, 2007) and a model was proposed in which Drosha cleaves the pre-miRNA before the intron is spliced out and the two RNA fragments, produced by the pre-miRNA cropping from the pre-mRNA, and might subsequently be *trans*-spliced. Recently, Kataoka et al (2009), using in vitro splicing assay system have shown that pre-miRNA cropping could occur prior to mRNA splicing. Moreover, the presence of miRNA in the intron slowed splicing while increasing pre-miRNA production. They have also shown a Y-shaped intermediate and a Y shaped intron

indicating *trans*-splicing. Recently, the association of Drosha (microprocessor) and spliceosomal snRNPs has been shown using northern blotting analysis (Kataoka et al, 2009). This indicates there is a cross-talk between the spliceosomal machinery and the micro processor which leads to processing of miRNA from the intron as well as splicing of the pre-mRNA.

#### 1.4 Model System: Myosin Heavy Chain and miR 208

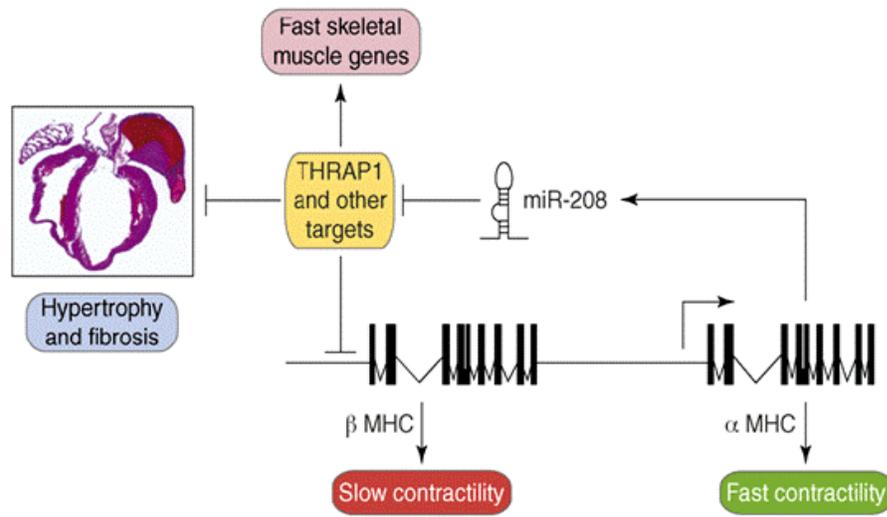
Myosin heavy chain (MHC or MYH) genes are expressed in the heart and largely regulate cardiac contractility. MHC gene has two isoforms,  $\alpha$  and  $\beta$ , which are regulated developmentally.  $\beta$ -MHC (MYH7), is the dominant isoform expressed in the embryonic heart and slow skeletal muscle whereas  $\alpha$ -MHC (MYH 6) is expressed postnatally in the heart (Weiss and Leinwand, 1996). Thyroid hormone (T3) signaling stimulates  $\alpha$ -MHC and



**Figure 5: Representation of MYH6 gene and miR-208:** Shown is the location of miR-208 in intron 27 of MYH6 gene which is a 24.63 kb long gene. Asterisks show the sequence conservation across humans, mouse, rat and dog. Also depicted is the proposed stem-loop structure of miR-208 and mature miR-208 (van Rooij et al, 2007)

inhibits  $\beta$ -MHC transcription after birth (Morkin, 2000). However, in state of heart disease there is a switch in MHC genes from  $\alpha$  (fast contracting myosin) to  $\beta$ -MHC (slow, embryonic myosin).

Human  $\alpha$ -MHC (MYH 6) gene is 25.63 kb long and has 37 exons. Intron 27 of the gene contains a highly conserved miRNA, miR-208a (Figure 5). MYH7 gene is a 23.8 Kb long gene having 41 exons and intron 31 harboring miR-208b. miR-208a (referred as miR-208) and miR-208b are 90% identical sequentially. The sequence alignment of miR-208 a stem-loop, proposed stem-loop structure of miR-208 and mature miR-208 sequence are shown in Figure 5. miR-208 stem-loop sequence is 90% identical across four species, namely human, mouse, rat and dog, and the mature sequence is completely conserved. This



**Figure 6: Regulation of cardiac contractility by miR-208:** miR-208 expressed by MYH6 negatively regulates THRAP1 expression. miR-208 absence results in upregulation of skeletal muscle genes and hence cardiac hypertrophy (vanRooij et al, 2008).

indicates that miR-208 plays some important functions in heart. miR-208 plays an important role in the regulation of  $\beta$  to  $\alpha$  switch, cardiac hypertrophy and fibrosis. Mice with homozygous deletion of mir-208 are viable but do not upregulate MYH7 expression in

response to cardiac stress; however other stress responsive genes were induced. This indicated the role of this tiny miRNA in specifically controlling MYH7 expression.

miR-208 has a half-life of 14 days and therefore can exert its effects even in the absence of MYH6 mRNA expression. miR-208 knockout mice at 2 months of age showed expression of skeletal muscle contractile protein genes which otherwise are not expressed in the heart. This indicated that miR-208 is expressed with MYH6 to repress the expression of skeletal muscle genes in the heart (vanRooyj et al, 2007).

MYH7 expression is repressed by thyroid hormone signaling and is upregulated in hypothyroid state. Hypothyroidism can also be induced by treating mice with thyroid hormone inhibitor, propylthiouracil. miR-208 knockout mice treated with propylthiouracil were found to be resistant to MYH7 expression upregulation. Thyroid hormone receptor (TR) coregulator TR-associated protein 1 (THRAP1) is the strongest predicted target of miR-208 which can exert both positive and negative effects on transcription (vanRooyj et al, 2008). The TR represses MYH7 expression in the adult heart through a negative thyroid hormone responsive element. In the absence of miR-208, THRAP1 expression increases and hence would repress the MYH7 expression (Figure 6).

Knowing these facts about miR-208, we were interested in knowing if the miRNA processing from the intron and splicing of pre-mRNA are related. Kataoka et al (2009) have shown that the processing of miR-330 from intron of EML2 gene occurs prior to splicing in vitro and both processes are coupled. If this is true universally, miR-208 is also an intronic miRNA; hence it's processing from the must be coupled to the process of splicing. Another question to be answered is miRNA processing dependent on splice sites. If the processes of miRNA processing and pre-mRNA splicing are coupled and trans-

splicing occurs (Kataoka et al, 2009), miR processing could be related to 5' splice site, 3' site and conserved branch site sequences. Trans-splicing is a common phenomenon in nematodes (Nilsen, 1993) and trypanosomes (Sutton and Boothroyd, 1986). Trans-splicing has been established in vitro system as well and essentially involves the transfer of a 5' exon from a pre-mRNA sequence containing an exon, 5' splice site and adjacent intron to a pre-mRNA transcript with exon, 3' splice site and adjacent intron (Konaraska et al, 1985). For introns harboring miRNAs trans-splicing has been predicted, and after the processing of miRNA from intron, trans-splicing takes place, hence the splice sites of the intron must play an important role in the process. If this is true and splice sites play an important role, how flexible or how permissive could the splice sites be to allow miRNA processing and pre-mRNA splicing to occur.

Nearly 97% of the human genome is non-coding DNA and the intron occupies most of it. Numerous intronic sequences have been recently found to encode microRNAs (miRNAs), responsible for miRNA-mediated gene silencing (Ying and Lin, 2009). Several kinds of intronic miRNAs have been identified in *C. elegans*, mouse, and human cells. However, their functions and applications are not clear yet. Till date, intronic miRNAs are known to exist only in U2-dependent introns only. This could possibly be due to the requirement of RNA polymerase II and spliceosomal components for their biogenesis (Ying et al, 2010). An interesting question arises, why is a miRNA not present in a U12-dependent intron. Is minor spliceosome not compatible with miRNA processing. If a miRNA is incorporated in a U12-dependent intron, will it be processed and pre-mRNA splicing will also occur. Recently, it was shown that revealed that part of the pre-miRNA cofractionated with the spliceosome in glycerol gradient sedimentation experiments

(Kataoka et al, 2009. U2, U5, and U6 but not U1 or U4 snRNAs coprecipitated with Drosha suggesting that major spliceosome associates with Drosha (microprocessor).

## CHAPTER II

### MATERIALS AND METHODS

#### 2.1 Minigene Construction

A genomic fragment of  $\alpha$ -MHC gene consisting exons 26 through 29 was amplified by PCR using primers MYH Xho F 5'-GCCTCGAGGAGAGTTGGCCCG GCAG-3' (forward) and MYH BamH1 R 5'-ACGGGATCCCCTGAAGGTTCTTGTTCC-3' (reverse). The amplified fragment was ligated into pCDNA 3.1(-) plasmid vector (Invitrogen, CA) under Xho1 and Bam H1 restriction enzyme sites, clones were selected and verified by sequencing. The resultant plasmid was named MYH-208 WT (WT). Splice site mutants and miR mutants were constructed by overlapping PCR using following sets of primers and then ligating in pCDNA 3.1 (-) vector. 5'SS mutant was constructed with primer 2085'SS F 5'-GAGGCCAACCGGGCTCCAG-3' (forward) and 208 5'SS R 5'-CTGGAGCCCGGTTGGCCTC-3' (reverse); U12 5'SS mutant with 208 U12 5'SS F primer 5'-GAGGCCAAGTATCCTTCAGATACCC-3' (forward) and 208 U125'SS R 5'-GGGTATCTGAAGGATACTTGGCCTC-3' (reverse); 208 U12BS mutant with primer 208 U12BS F 5'-CACACACATCCTTAACCCTCACC-3' (forward) and 208 U12BS R 5'-GGTGAGGGTTAAGGATGTGTGTG-3' (reverse).

Primers used for lower helix mutant (LH) were 208 LH F 5'-CTGTGACGCCGCTCGAAAAGGCCCGGG-3' (forward) and 208 LH R 5'-CCCGGGCCTTTTCGAGCGGCGTCACAG-3' (reverse). Primers used for upper helix mutant (UH) were 208 UH F 5'-GGCCCGGGAATATCCTGATGCT-3' (forward) and 208 UH R 5'-AGCATCAGGATATTCGCCGGCC-3' (reverse). PCR-based site-directed mutagenesis was performed using Change-IT multiple mutation site directed mutagenesis kit (USB) to generate loop mutant (Loop) in WT background using 208 Loop primer 5'-CCCGGGTTATACCACTACGAGTCGTATAAGACGAG-3' and MYH CC5/6GG mutant was created in 208 U125'SS mutant background using 208 CC/GG primer 5'-GAGGCCAAGTATGGTTCAG-3'.

For in vitro transcription, all the above mentioned mutants were PCR amplified using forward primer with T7 promoter sequence, MYH T7BamH1 F 5'-ATATGGATCCTAATACGACTCACTATAGGGGGAGCAGTACGAGGAGGAGA-3' and MYH Xba1 reverse primer 5'-ACGGTCTAGACGTCCACCATCAAGTCCTCT-3' and cloned between BamH1 and Xba1 sites of pUC19 vector.

Point mutants were made using Change-IT multiple mutation site directed mutagenesis kit (USB) using the mentioned primers: 5'-CCAAGTAAGCTCC-3' for G3A; 5'-CCAAGTCAGCTCC-3' for G3C; 5'- CAAGTGCGCTCCA-3' for A4C; 5'- CAAGTGGGCTCCA-3' for A4G; 5'-AAGTGA ACTCCAG-3' for G5A; 5'- AAGTGACCTCCAG-3' for G5C; 5'- AGTGAGGTCCAGA-3' for C6G; 5'- AGTGAGTTCCAGA-3' for C6T; 5'-GTGAGCACCAGAT-3' for T7A; 5'-GTGAGCGCCAGAT-3' for T7G.

p120 minigene construct (p120 WT) was a gift from Richard Padgett, Learner Research Institute, Cleveland Clinic Foundation. It has a well characterized U-12 dependent intron (intron F). p120 CC5/6GG mutant was made by mutating CC5/6GG in intron F of p120 WT. For p120-miR208 mutant, site-directed mutagenesis was performed using Change-IT multiple mutation site directed mutagenesis kit (USB) to introduce PmeI and NotI sites in the p120 WT construct in pCB6 vector using 5'-GAGGAAGGCGGCCGCGAGAGG G-3' to introduce NotI site and 5'-GAGAGGGGTTTAAACGCCTTAG-3' primer to introduce PmeI site. The DNA fragment containing pre-miR330 was amplified by PCR from the WT plasmid using 5'-GAATGCGGCCGCTGAACTCTGGCTGGCCCTGA-3' (forward) and 5'-GCATGTTTAAACCACACAGGCTGATCGACGGT-3' (reverse) and cloned into PmeI and NotI modified p120 WT. Site directed mutagenesis was performed using Change-IT multiple mutation site directed mutagenesis kit (USB) and called p120 5'SS mutant with primer 5'-GATGGAGCAGGATGAGTTTGCAGGGCAGAG-3' and second mutant called p120 CC5/6GG using 5'-TGGAGCAGGATATGGTTGCAGGGCAGAG-3'. U11 and U6atac suppressors cloned in pALTER expression vector was a gift from Richard Padgett, Learner Research Institute, Cleveland Clinic Foundation.

All plasmids were transformed in DH5 $\alpha$  competent cells, plated on Ampicillin (Amp 100 mg/ml) containing Agar plates. Colonies were picked and grown overnight in 5ml Luria Broth (LB, Amresco, OH) containing 5  $\mu$ l Amp (100 mg/ml) at 37 °C. Mini-preps were done using Zyppy miniprep kit (Zymo Research, CA). Mutations were confirmed by sequence analysis and then maxiprep was done using Plasmid Maxi Kit (Qiagen, AL).

## **2.2 CHO Cell Maintenance and Transfection**

CHO cells were cultured in 10 ml complete CHO media containing DMEM-1x with 4.5 g/l glucose and L-glutamine (CCF, OH), P/S (CCF, OH), 5% FBS (Atlanta Biologicals, GA), 1 mM L-proline (Sigma, MO) and 10 mM HEPES (SIGMA, MO). All cells were kept in a humidified incubator at an atmosphere of 5% CO<sub>2</sub> at 37 °C.

For transfection, a 100% confluent 100 mm plate was split at 1:8 ratio in 100 mm plates using 10 ml complete CHO media. The cells were ~70 % confluent after 24 hrs and morphologically healthy for transfection. Cells were transfected using 5 ml complete CHO media containing 12 µl polybrene (10 mg/ml Sigma, MO) and 1.2 µg plasmid DNA per plate. 9 µg pUC19 was used as carrier and to equalize amount of DNA being added. In wells where both U11 suppressor plasmid and U6atac suppressor mutants were transfected together, pUC19 vector was omitted. After 8 hrs, media from the wells was aspirated and cells were treated with 5 ml of 30 % DMSO (in complete CHO media) and incubated at room temperature for 5 minutes. Plates were washed with 5 ml complete CHO media and finally replenished with 10 ml CHO media. Cells were cultured at 37 °C in a humidified 5 % CO<sub>2</sub> incubator. The cells were harvested after 36 hours of incubation post-DMSO shock.

## **2.3 Total RNA Isolation**

Total RNA isolation from cells growing in monolayer was done by TRIzol Method, a modification of the method developed by Chomczynski, P. and Sacchi, N. (1987). Briefly, the cells were lysed directly in culture dish by adding 1 ml TRIzol (Invitrogen, CA) reagent to each well, transferred to an eppendorf and added 0.2 ml Chloroform (Amresco, OH). Samples were centrifuged and aqueous phase was collected.

RNA present in aqueous phase was precipitated and washed using Isopropanol (Amresco, OH) and 70 % Ethanol (Amresco, OH) respectively. Pellet obtained finally was resuspended in DEPC treated water and stored in -80 °C till further use. The isolated RNA was quantitated spectrophotometrically and absorbance was taken at 260 nm and 280 nm. Integrity of RNA was checked on denaturing 1% MOPS-Agarose gel.

#### **2.4 Reverse Transcriptase- Polymerase Chain Reaction (RT-PCR)**

Reverse transcription was performed using GeneAmp ThermoStable rTth Reverse Transcriptase RNA PCR Kit (Applied Biosystems, CA) using 0.5 µg total RNA per reaction. A master mix was made as per manual but scaling down reaction volume to 10 µl.

For MYH-mir208 stem-loop and splice site mutants, RT was done using vector specific RT primer 5'-CAACAGATGGCTGGCAACTA-3' at 70 °C for 15 min in a PTC-100 (Bio-Rad Laboratories, CA) thermal cycler. The reaction was stopped by keeping tubes on ice for at least 5 min. PCR was performed using master mix specified in the manual scaling down reaction to 40 µl using Exon Junction F primer 5'-GAGGGCAAGGCGAAGAAC-3'. Amplification was done for 40 cycles consisting of denaturation at 95 °C for 1 min, annealing/extension at 55 °C for 1.05 min and a final extension at 68 °C for 7 minutes. RT for point mutants was performed using RT primer according to user manual (Promega, WI).

For p120 minigene, RT was done using p120 mini-gene specific reverse primer E-7-8-2 primer 5'-CTTCTAAGAACTCCACCAGCTCAGA-3' at 70 °C for 15 min in a PTC-100 (Bio-rad Laboratories, CA) thermal cycler. The reaction was stopped by

incubating tubes on ice for at least 5 min. Then PCR was performed using master mix specified in the manual scaling down reaction to 40  $\mu$ l using E-5-6 primer 5'-GGCCCGGGAAGCTGCTGCTGGGATC-3'. Amplification was done for 40 cycles with denaturation at 95 °C for 1 min, annealing/extension at 60 °C for 1.05 min and a final extension at 68 °C for 7 minutes.

### **2.5 *In vivo* Splicing Phenotype Analysis for MYH6 miR-208**

Nested PCR was performed using GoTaq PCR core system II (Promega, WI) setting up a reaction of 25  $\mu$ l according to manual provided. Primers used for amplification were F2 5'- GGAGCAGTACGAGGAGGAGA-3' (forward) and R2 5'-CGTCCACCATCAAGTCCTCT-3' (reverse). PCR was done using program "55" involving denaturation at 95 °C for 3 min followed by 25 cycles of denaturation at 95 °C for 1 min and annealing/extension at 60 °C for 1.05 min final extension at 68 °C for 7 min. The products were analyzed on 2 % Agarose-1X TBE gels.

Nested PCR was performed using GoTaq PCR core system II (Promega, WI) was also preformed to check various spliced products with upstream and downstream exons. Primers were used in different combinations as indicated. Primers used are indicated as F for forward and R for reverse, F1: 5'-GGAAAAGGAGGCGCTAATCT-3'; F2: 5'-GGAGCAGTACGAGGAGGAGA-3'; F3: 5'-GGCTGTTAATGCCAAGTGCT-3'; R1: 5'-GCGTCCGTCTCATACTTGGT-3'; R2: 5'-CGTCCACCATCAAGTCCTCT-3'; R3: 5'-TCGTAGGCGTTCTTGAGCTT-3'. PCR amplification was done using PTC-100 (Bio-Rad, CA) thermal cycler. PCR program used was denaturation at 95 °C for 2 min, followed by 36 cycles of denaturation at 95 °C for 1 min, annealing at 55 °C for 1 min,

extension at 72 °C for 1 minute and a final extension at 72 °C for 5 min. PCR products were checked on a 2 % Agarose gel.

## **2.6 *In vivo* Splicing Phenotype Analysis for p120 Minigene**

Nested PCR was performed using Hi Fidelity PCR Kit (Roche, IN) setting up a reaction of 25µl according to manual provided. Primers used for amplification were E6 long 5'- TTGTGCTGCCCCCTGCTGGGGAGATG-3' (forward) and E7 long 5'- TGAGCCCCAAAATCAGCAGAATTCC-3' (reverse). PCR was done using amplification program involving denaturation at 95 °C for 3 min followed by 25 cycles of denaturation at 95 °C for 1 min and annealing/extension at 60 °C for 1.05 min, final extension at 68 °C for 7 min. The products were analyzed on 3 % (2 % Agarose: 1 % NuSieve) Agarose-1X TBE gels.

## **2.7 *In vitro* Transcription Assay**

Reaction was set up at room temperature using maxiscript kit (Ambion, TX). A 50 µl reaction was set up and reagents were added in following order: 12.5 µl Nuclease-free water, 5 µl 10X Transcription Buffer (with DTT), 2.5 µl 10 mM ATP, 2.5 µl 1 mM UTP, 2.5 µl 10 mM CTP, 2.5 µl 1 mM GTP, 2.5 µl 10 mM m7G(5')ppp(5')G cap analog (Ambion, TX), 10 µl (2 µg) linearized DNA template, 5 µl [ $\alpha$ -32P] UTP 800 Ci/mmol (10 mCi/mL- MP Biomedicals, OH) and finally 5 µl T7 RNA Polymerase (10 U/µl). All reactions were incubated at 37 °C for 2 hrs. Then, 2 µl RNase-free DNase I (2 U/µl) was added to the reaction mixture to remove the DNA template, mixed thoroughly, and incubated for 20 min at 37 °C. Gel loading Buffer II (Ambion, TX) was added to

samples. Samples were loaded on pre-electrophoresed 8 M Urea 5% polyacrylamide gel. Covered gel with saran wrap and exposed it to phosphor imager screen for about 1 min and scanned and extracted the band. To elute RNA from gel, added 400  $\mu$ l G50 buffer (1.7 g Sodium Acetate, 0.5 ml 2 M Tris pH 7.5) to the band and eluted overnight on rotator at 4 °C. Next day, collected supernatant in a fresh tube, 900  $\mu$ l ethanol and 40  $\mu$ g glycogen (Roche, IN) were added and precipitated RNA at -20 °C overnight. Following day, samples were spun at 13000 rpm 20 min and pellet obtained was resuspended in nuclease-free H<sub>2</sub>O and store at -20 °C in a  $\beta$ -blocker.

Century Markers (Ambion, TX) were transcribed as per manual guidelines. Briefly, reaction was set using 6  $\mu$ L Nuclease-free water, 2  $\mu$ L 10X Transcription Buffer (with DTT), 1  $\mu$ L 10 mM ATP, 1  $\mu$ L 10 mM CTP, 5  $\mu$ L 10 mM GTP, 1  $\mu$ L 1 mM UTP, 1  $\mu$ L 0.5  $\mu$ g RNA Century Marker Template Set, 2  $\mu$ L [ $\alpha$ -32P] UTP 800 Ci/mmol (10 mCi/mL in aqueous solution) and 1  $\mu$ L T7 RNA Polymerase (10 U/ $\mu$ L). Reaction mix was incubated at 37°C for 1 hr and 1  $\mu$ L RNase-free DNase I (2 U/ $\mu$ L) was added to the reaction to remove the DNA templates, and incubated for an additional 15 min at 37°C. The terminated transcription reaction was passed through a G-25 or G-50 chromatography column to remove unincorporated nucleotides. Counts per minute (cpm) were recorded and sample was diluted to 2000 counts/ $\mu$ L.

Decade Markers (Ambion, TX) were transcribed as per manual guidelines. Briefly, reaction mix was made in a nuclease-free tube using 1  $\mu$ L Decade Marker RNA (100 ng), 6  $\mu$ L Nuclease-free water, 1  $\mu$ L 10X Kinase Reaction Buffer, 1  $\mu$ L [ $\gamma$ -32P] ATP and 1  $\mu$ L T4 Polynucleotide Kinase. Incubated at 37°C for 1 hr. Then 8  $\mu$ L Nuclease-free water and 2  $\mu$ L 10X Cleavage Reagent were added. Reaction mix was

further incubated at room temperature for 5 min. Then 20  $\mu$ L Gel Loading Buffer II was added to stop the reaction. The mixture was heated at 95°C for 5 min before loading on the gel.

## **2.8 *In vitro* pre-mRNA Splicing in Cell-free Extracts**

All reactions were set up on ice. Mix1 was made fresh every time and contained 25  $\mu$ l ATP/Phosphocreatine/Cordycepin 5' phosphate mix (60  $\mu$ l 100mM ATP (Roche, IN), 50  $\mu$ l 0.5 mM Cordycepin 5' phosphate (Sigma, MO), 100 mM Cordycepin 5'triphosphate- 15  $\mu$ l and H<sub>2</sub>O- 375 $\mu$ l), 6  $\mu$ l 0.5M Phosphocreatine and 49  $\mu$ l H<sub>2</sub>O. To set up a 25  $\mu$ l reaction, added 8.0  $\mu$ l Mix 1, 1.0  $\mu$ l RNase inhibitor (SUPERase In- 20 U/ $\mu$ l- Ambion, TX), 1  $\mu$ l 25mM MgCl<sub>2</sub>, 4  $\mu$ l nuclease-free H<sub>2</sub>O, 10  $\mu$ l nuclear extract and finally 1  $\mu$ l substrate mRNA (P-32 labeled) in an eppendorf. If anti-U2 or anti-U12 2'O-methyl oligos (Dharmacon, CO) were used, the reactions were incubated at 30 °C for 15 minutes and then labeled-pre mRNA was added. To block U2 dependent splicing, 8  $\mu$ M anti-U2b oligo was used whose sequence is complementary to U2 nucleotides 27–49 (Dietrich et al, 1997). To block U12 dependent splicing 8 $\mu$ M anti-U2b oligo along with 5  $\mu$ M anti-U12 oligo (Dharmacon, CO) was used whose sequence was complementary to U12 nucleotides 11–28 (Dietrich et al, 1997). All reactions were incubated at 30 °C for 90 minutes in case of adeno and adeno-miR constructs and 3 hours for reactions involving p120 and MYH. Then centrifuged and added 200  $\mu$ l G50 buffer (1.7g Sodium Acetate, 0.5 ml 2M Tris pH 7.5) followed by 5  $\mu$ l Proteinase K (Roche, IN) and incubated the reaction at 42 °C for 20 minutes. Then added 200  $\mu$ l of phenol/ chloroform (Ambion, TX), vortexed 3-4 times and centrifuged at 13000 rpm for 5 minutes at room

temperature. Carefully collected the supernatant in another eppendorf and added 200  $\mu$ l chloroform (Amresco, OH), vortexed 3-4 times and centrifuged at 13000 rpm for 5 minutes at room temperature. Again collected supernatant in a new tube and added 40  $\mu$ g glycogen (Roche, IN) followed by 500  $\mu$ l ethanol. Samples were kept at -20 °C overnight for precipitation. Next day, centrifuged samples at 13000 rpm for 20 minutes at 4 °C and pellet obtained was resuspended in 10  $\mu$ l gel loading buffer II (Ambion, TX). Samples were incubated at 95 °C for 2 minutes, immediately chilled on ice and loaded on heated Urea-PAGE gel. Gel was run at 15-18W constant power. Gel was dried and exposed to phosphor imager screen at least overnight.

## **2.9 Reverse Transcription and Quantitative Real-Time (qRT) PCR Analysis of Mature miRNA Expression**

First strand cDNA was synthesized from 100ng of total RNA using primers specific for mature miR-208 and small nucleolar RNA 202 (sno202). The miRNA-specific primers were obtained from TaqMan MicroRNA Assays (Applied Biosystems Foster City, CA). Reagents for cDNA synthesis were obtained from TaqMan MicroRNA Reverse Transcription kit (Applied Biosystems, CA). For each sample, a 15 $\mu$ l reverse transcription (RT) reaction was set up containing 10ng of total RNA, 1X RT buffer, 1mM of dNTP mix, 50 units of MultiScribe reverse transcriptase, 3.8 units of RNase inhibitor and 3 $\mu$ l of miRNA-specific RT primer. The reactions were incubated in a thermal cycler (Bio-Rad PTC-100) at 16°C for 30 minutes, 42°C for 30 minutes, and 85°C for 5 minutes and then held at 4°C. The 'reverse transcriptase minus' controls were also synthesized under the same conditions. In order to quantify the mature miR-208 and sno202 in each

sample, the cDNAs were amplified using TaqMan MicroRNA Assays together with the TaqMan 2X Universal PCR Master Mix (Applied Biosystems, CA). Briefly, a 20 $\mu$ l reaction was set up containing 1.33 $\mu$ l product from RT reaction, 1 $\mu$ l of 20X TaqMan microRNA assay mix (mixture of miRNA-specific forward and reverse primers, and miRNA-specific TaqMan MGB probe labeled with FAM fluorescent dye) and 10 $\mu$ l of TaqMan 2X Universal PCR Master Mix. These reactions were dispensed into a 96-well optical plate and the plate was positioned in 7500 Real-time PCR System (Applied Biosystems) under the following conditions: 95°C for 10 minutes followed by 40 cycles of 95°C for 15 seconds and 60°C for 1 minute. Three replicates were performed per RT reaction together with the ‘reverse transcriptase minus’ and ‘no template’ controls. Duplicate PCRs were performed for all the RNA sample. The mean  $C_t$  was determined from the replicates. Sno202 expression was used as an invariant control. The relative expression of each miRNA was calculated as  $2^{-\Delta C_t}$  where  $\Delta C_t = C_t$  value of each miRNA in a sample –  $C_t$  value of sno202 in that sample. All experiments were repeated with three and all replicates provided similar results.

### **Statistical Analysis**

The statistical analysis of expression levels of miR-208 and was performed using the SPSS package. The index of expression of each miRNA was  $2^{-\Delta C_t}$  after normalization to sno202 expression levels. A student t-test was applied to the samples. Expression levels are considered significant if  $p < 0.5$ .

## **CHAPTER III**

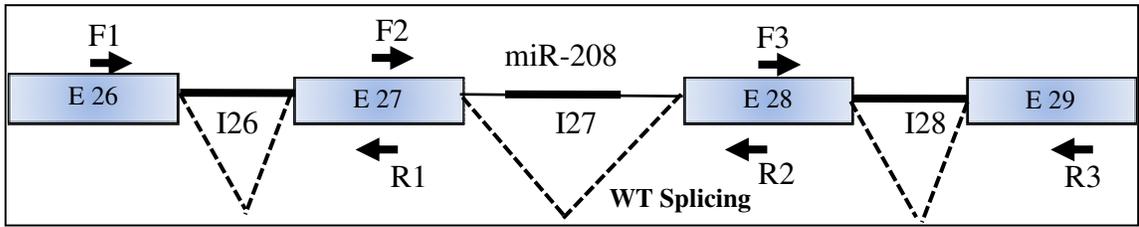
### **RESULTS and DISCUSSION**

#### **3.1 Construction of MYH Minigene Reporter System**

To study the events of pre-mRNA splicing and miRNA processing, a reporter system that could experimentally enable us to study both processes was required. To achieve this aim, a minigene reporter was constructed from Myosin Heavy Chain 6 gene (MYH6). A 1764 base long fragment (Appendix 2) from exons 26 through 29 was PCR amplified from human genomic DNA and cloned between the Xho1 and BamH1 sites of pCDNA3.1 (-) under CMV promoter (Appendix 1). Hence, the reporter system consists of exons 26 to 29, including introns 26, 27 and 28. Intron 27 of MYH6 gene encodes the miRNA, miR-208. miR-208 has been implicated in regulating cardiac hypertrophy and myocardial infarction (vanRooy et al, 2007; Callis et al, 2009).

A schematic of minigene reporter construct is represented in Figure 7. E 26 to E29 correspond to the exons 26, 27, 28 and 29 from the MYH6 gene and I 26 to I 28 represent the introns 26, 27, 28 respectively. Forward primers designated as F1, F2 and F3 are located in exons 26, 27 and 28 respectively and R1, R2 and R3 represents the

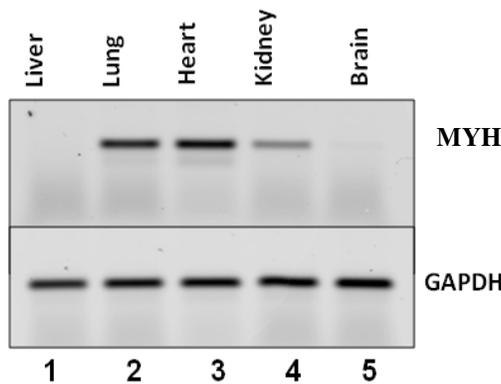
reverse primers located in exons 27, 28 and 29 respectively (further details Chapter II and Appendix).



**Figure 7: Structure of the minigene reporter construct and location of primers:** E represent the exons 26-29 and I represent the introns 26-28. Forward primers are denoted by F and R represents reverse primers. miR-208 is encoded in the intron 27 of the minigene construct. Dashed lines show wild-type splicing junctions.

### 3.2 Expression of MYH in Mouse Tissues

miR-208 is encoded by intron 27 of MYH6 gene and is expressed specifically in the heart and tracey in lungs (van Rooij et al, 2007; Callis et al, 2009). It was important

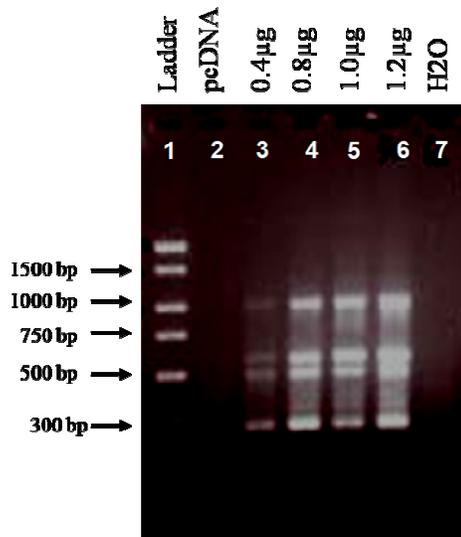


**Figure 8: MYH6 expression in mouse organs:** RNA was isolated from mouse tissues and RT-PCR analysis was done using specific primers. Upper panel shows results for MYH6 gene expression in mouse tissues and lower panel shows expression of GAPDH (internal control).

to know if MYH6 is expressed in other tissues or is it tissue-specific as miR-208. We performed RT-PCR on RNA isolated from mouse organs, heart, kidney, liver, brain and lungs. GAPDH was used as an internal control. Our data shows that MYH6 expresses in lungs, heart and kidney (Figure 8) if PCR amplification is done for 30 cycles. Expression of MYH6 gene was observed in brain as

well if the PCR cycles are increased from 30 to 40. However, MYH6 expression was not detected in liver even at 40 cycles. Since the miR-208 expresses only in mouse heart and traceably in lungs (van Rooij et al, 2007), its expression is tissue specific. Its expression could also be regulated post-transcriptionally, however, any factors involved in miR-208 processing from intron 27 of MYH6 are not yet known. This experiment also suggested us to choose a cell system not derived from heart, kidney, lungs or brain so that there is no interference of endogenous MYH6.

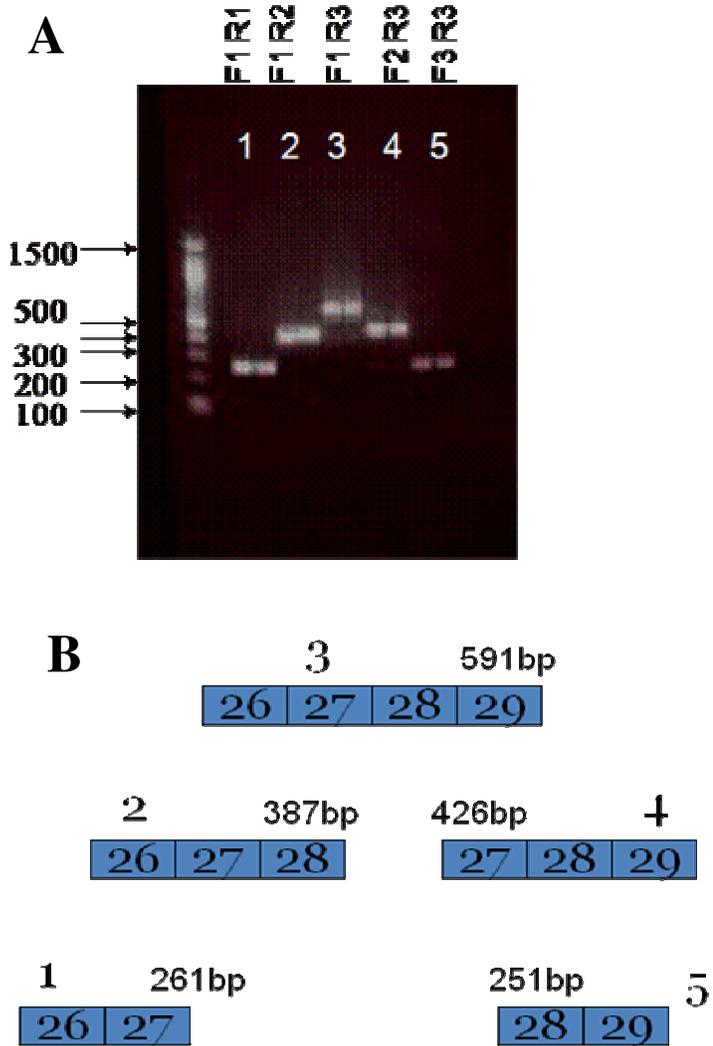
### 3.3 Optimization of Transient Transfection Conditions



**Figure 9: Optimization of transient transfection conditions:** Minigene construct was transfected at different concentrations as indicated, RNA was isolated 36 hrs post transfection and RT using vector backbone primer and PCR using F2 and R2 primers which flank miR-208 containing exons was performed.

To study the events of splicing and miRNA processing it was necessary to recapitulate the MYH6 splicing and miR-208 processing in an *in vivo* assayable system. To achieve this objective, MYH6 minigene reporter system was constructed (as described in Chapter II). MYH6 WT was transfected in HEK293T cells using lipofectamine at concentrations of 0.4 µg, 0.8 µg, 1.0 µg and 1.2 µg per well of a 6-well plate along with appropriate controls. Cells were harvested after 36 hours and RNA was isolated using

Trizol. To specifically amplify exogenously transfected MYH6 minigene construct,



**Figure 10: Optimization of primers sets located on MYH6 minigene construct:** PCR was performed with all primers designed and in Panel A spliced phenotypes and primer sets used for amplification are mentioned on top. Panel B depicts schematics of spliced products obtained using all combination sets of primers in panel A along with their sizes. Each box denotes an exon. For example, lane 1 in Panel A corresponds to spliced phenotype obtained using primers F1 and R1. Schematic of spliced product amplified shown in lane 1 is represented in panel B, product is spliced exons 26 and 27 and is 261 bp long.

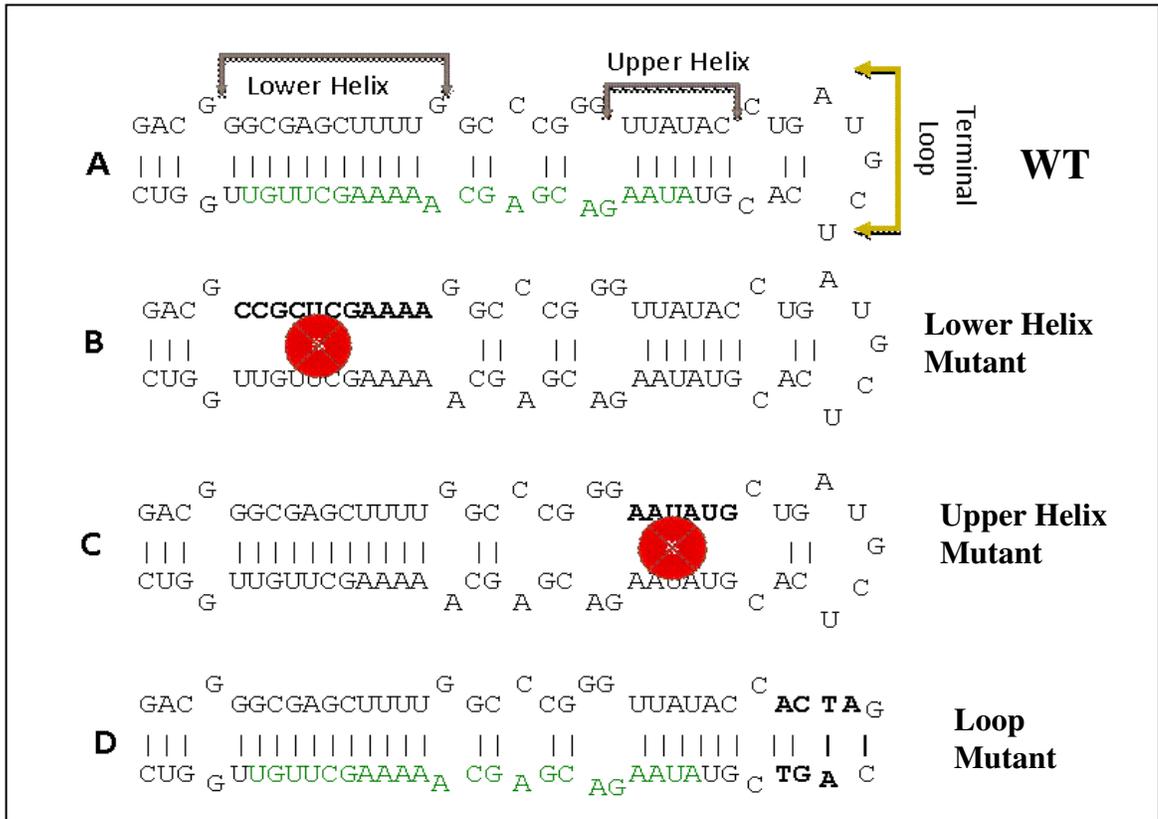
reverse transcription was done using a vector backbone specific primer (RT primer) and then PCR was done using F2 and R2 primers located on exons 27 and 28 respectively (Figure 7). The PCR product was loaded on a 2% Agarose-1x TBE gel and compared the splice variants (desired ~263 bp) with the 100 bp molecular weight ladder (Figure 9). Some other non-specific products were observed at ~600 bp and ~1000 bp. When the cDNA was cleaned up and purified to remove RT primer, none of the non-specific products was observed. The best spliced phenotypes were observed at a concentration of

1.2  $\mu\text{g}$  transfected in HEK293T cells and this optimized concentration was used for further transfections.

It was necessary to check if the *in vivo* splicing takes place from upstream and downstream exons as well as the efficiency of MYH minigene reporter construct. RT-PCR was performed using the vector specific primer and cDNA was cleaned using DNA clean and concentrator kit. PCR amplification was done using appropriate sets of primers as represented in Figure 10, panel A. If splicing occurs from the correct splice site junctions, products of specific lengths shown in Figure 10, panel B are expected. Since F1 primer is located on exon 26 and R1 primer is on exon 27, intron 26 will be spliced out and the expected product size is 261 base pairs long (Figure 10, Panels A and B, lane 1). Similarly, F1 and R2 primers would yield a 418 bp long product by splicing of exons 26, 27 and 28 (Figure 10, Panel A and B, lane 2). Primers F1 and R3 are located in exons 26 and 29 respectively and would yield largest product of 590 bp by splicing of exons 26 to 29 (Figure 10, Panel A and B, lane 3). F2 and R3 primers would yield a product of 435 bp, by splicing exons 27 to 29 (Figure 10, Panel A and B, lane 4). Primers F3 and R3 are located on exons 28 and 29 respectively and yield a spliced product of 252 bp long containing exons 28 and 29 (Figure 10, Panel A and B, lane 5). PCR products were isolated from the gel and sequenced to identify the spliced junction sequences. The sequencing results obtained showed that the PCR products spliced at correct splicing junctions.

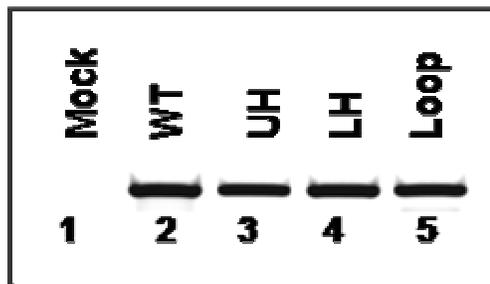
### 3.4 *In vivo* Splicing Phenotypes of miRNA Stem-Loop Mutants

It has been proposed that RNA-RNA interactions or RNA-protein interactions are involved in structure recognition and processing precursor miRNA to mature sequence. Lower helix, upper helix and the terminal loop are the required structural regions for the miRNA processing to occur (Kim and Kim, 2007). After optimization of the minigene system, we wanted to answer specific question- whether miRNA sequence affects pre-mRNA splicing or not. To answer this question we introduced mutations in the stem-loop structure of hsa-miR208 (Figure 11 A).



**Figure 11: Structure of predicted hsa-miR-208 stem loop:** (A) correspond to predicted stem-loop structure and sequence/regions of precursor miR-208. Sequence of mature miR-208 is illustrated in green letters. (B)- (D) represent lower helix, upper helix and loop mutants of hsa-miR-208. Regions marked with crosses in the mutants indicate abolished base pairing interactions whereas extended helix is predicted to form in loop mutant. Mutated bases are shown in bold letters.

We introduced non-complementary mutations in the helical regions to disrupt the base pairing interactions between 5' and 3' end of predicted miRNA stem-loop structure (Figure 11, B and C). These mutants were called the lower helix (LH) or the upper helix (UH) mutants. For the loop mutant, base pairing interactions were introduced at the terminal loop region and hence formation of an extended helix is predicted (Figure 11 D).



**Figure 12: *In vivo* splicing patterns of miRNA mutants:** UH, LH and Loop mutants were transfected with appropriate controls. PCR was performed using F2 and R2 primers. PCR products were loaded on 2 % Agarose-1X TBE gels.

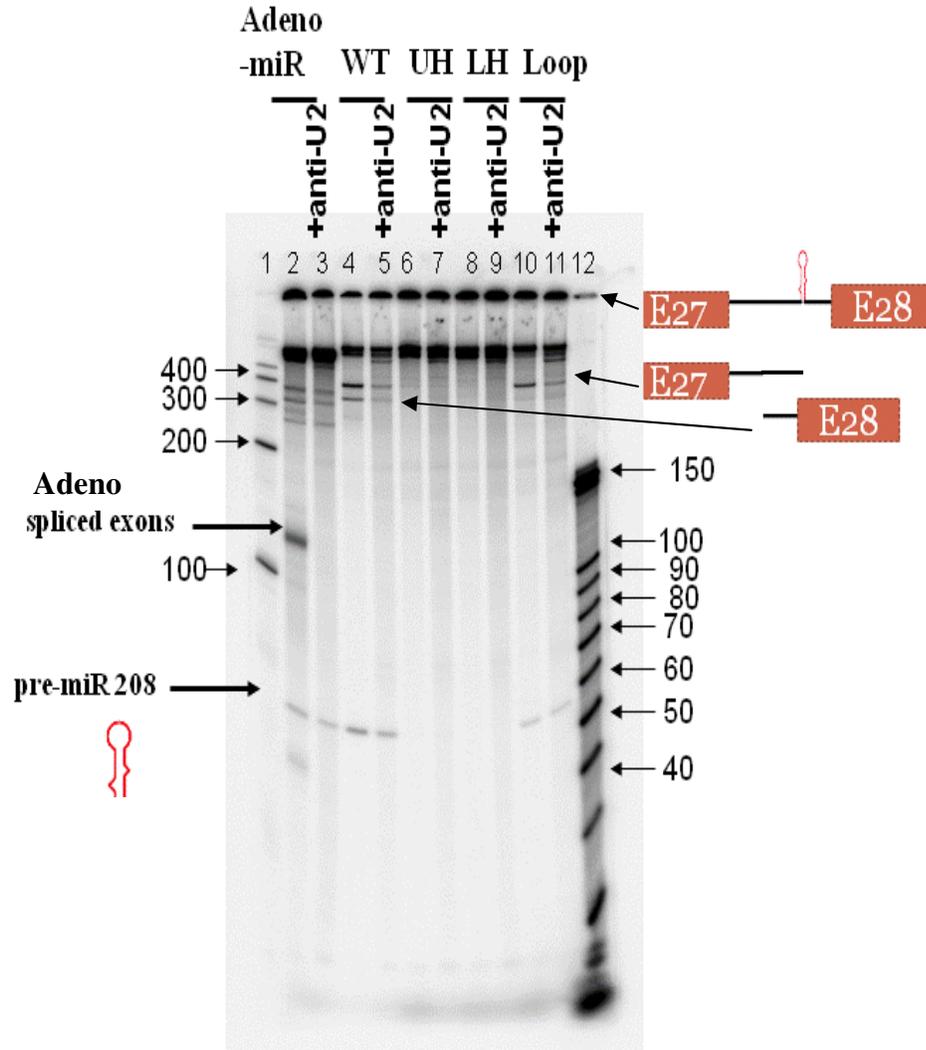
All splice site mutants along with wild type minigene construct and appropriate controls were transfected in CHO cells using polybrene and upon RNA isolation analyzed for spliced phenotypes. Wild type (WT) spliced product is 263 bp long if splicing occurs at correct splice junctions (Figure 12, lane 2). Upper helix (UH), lower helix (LH) and Loop mutants did not have any effect on splicing and yield product similar to wild-type (Figure 12, lanes 3, 4, 5). To verify if these mutants spliced from the wild type splice junctions, PCR products were eluted from the gel and sequenced. The analysis of sequencing data from the three miR-208 stem-loop mutants obtained showed that splice junctions are same as that of the wild-type. This data from the miRNA stem-loop mutants indicates that miRNA mutations do not affect splicing *in vivo*.

### 3.5 *In vitro* Splicing Assay for Stem-Loop Mutants

There are advantages of using the *in vitro* system over the *in vivo* system. Firstly, using an *in vitro* system all intermediate stages and products of splicing and miRNA processing can be resolved on the same gel, on the other hand, using *in vivo* system only the final spliced product can be identified. Secondly, one or more snRNAs can be blocked at different time points to establish the role of snRNAs in splicing (and miRNA processing). Splicing reaction kinetics can be studied and different complexes formed during the process of splicing can be identified.

A 788 bp region containing exon 27 and exon 28 including the miR-208 harboring intron 27 was PCR amplified from miRNA stem-loop mutants as well as the wild-type and cloned in pUC19 vector under T7 promoter. Plasmid DNA was linearized with restriction enzyme, XbaI and transcribed *in vitro* as described in Chapter 2. *In vitro* splicing assay was performed in HeLa cell nuclear extracts as described earlier and *in vitro* spliced phenotypes of MYH6 wild-type and miRNA stem-loop mutants were analyzed on 8M Urea-8% PAGE gel. To compare the sizes of the products generated in the assay, century (100 bp RNA ladder) and decade markers (10 bp RNA ladder) were used (Figure 13, lanes 1 and 12 respectively) since products ranging from 70 bp (pre-miR-208) to 788 bp (pre-mRNA transcript) were expected. Adeno-viral mRNA with hsa-miR-208 incorporated in intron was used as control to optimize *in vitro* splicing reaction conditions. Both adeno-viral and MYH6 minigene constructs splice using major spliceosomal pathway. Adenoviral construct containing miR-208 in intron was spliced *in vitro* and spliced exons and precursor miR-208 (pre-miR-208) were observed (Figure 13, lane 2), indicating that the nuclear extract is sufficient for *in vitro* splicing to occur.

Spliced exons were not observed in adenoviral construct incubated with anti-U2 oligo (Figure 13, lane 3). This showed that the anti-U2 oligo could block U2-dependent



**Figure 13: *In vitro* spliced phenotypes of stem loop mutants:** Lane 1 shows century marker, lane 12 shows decade marker, lanes 2 and 3 show *in vitro* spliced adeno-miR construct. Lanes 4-5, 6-7, 8-9 and 10-11 show *in vitro* spliced MYH6 wild-type, UH, LH and Loop mutants respectively. Anti-U2 indicates samples where anti-U2 oligo was added to block U2-dependent splicing. Drosha cleaved products namely, miR-208 stem-loop, upstream exon to 5' end of miR and from 3' end of miR to 3'exon were observed as indicated.

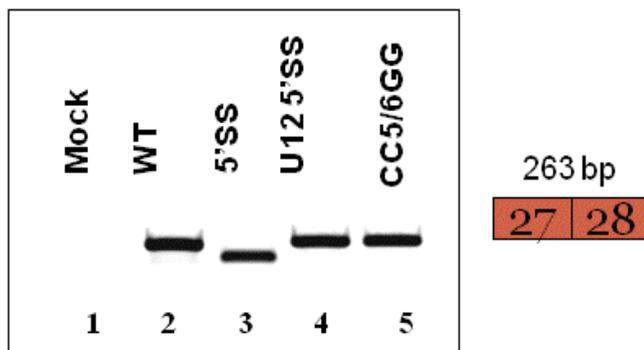
splicing (as its sequence is complementary to U2 nucleotides 27–49 (Dietrich et al, 1997). Interestingly, blocking splicing did not abolish miRNA processing upto the

precursor stage, though the level of pre-miR-208 observed was reduced. Wild-type MYH construct appears to be splicing deficient, however precursor microRNA was observed (Figure 13, lane 4). Interestingly, two pre-mRNA by-products were also observed between 300 and 400 bp. These products appear to be Drosha cleaved products, i.e. the region from starting of mRNA to 5' end of pre-miR-208 and 3' end of pre-miR-208 to the end of mRNA (Figure 13) and the length of these products should be 387 bp and 330 bp (by analyzing the genomic sequence). However, the identity of these products has not been experimentally proven yet. Wild type MYH construct along with anti-U2 oligo showed reduced levels of the Drosha cleaved pre-mRNA products as well as the pre-miR-208 (lane 5). Interestingly, pre-miR-208 along with Drosha cleaved pre-mRNA was also not detected in upper helix (UH) and the lower helix (LH) mutants, in the presence or absence of anti-U2 oligo (Figure 13, lanes 6-9). *In vitro* spliced products of loop mutant, with and without anti-U2 oligo are shown in lanes 10 and 11. This mutant is also splicing deficient as the wild-type and hence spliced products were not observed, however Drosha cleaved pre-mRNA products and pre-miRNA in low level were seen.

Recently, mRNA splicing and pre-miRNA processing have been detected in an *in vitro* system with HeLa cell nuclear extract using the EML2 gene harboring miR-330 (Kataoka et al, 2009). The EML2 pre-mRNA was a good substrate for splicing as well as miRNA processing in contrast to the MYH6 pre-mRNA. We also performed *in vitro* splicing assay with a 900 bp long (including full exons 27 and 28) MYH6 pre-mRNA substrate but observed similar results. We suspect that MYH6 gene might have some key cis-regulating elements, upstream of exon 27 or downstream of exon 28, necessary for MYH6 pre-mRNA to splice. Kataoka et al (2009) showed that pre-mRNA processing



Since, 5' splice site binds to U1 and U6 snRNAs by Watson- Crick base pairing and these RNA-RNA interactions are phylogenetically conserved and essential for splicing to occur, we mutated the 5' splice site of the intron 27 of the minigene reporter construct to eliminate the wild type site (Figure 14, panel B). The 5' splice site was mutated to consensus U12-type 5' splice site (GUAUCCUU) and this mutant was called U125'SS mut. Firstly, we wanted to know if this mutant would splice, since branch site and 3' splice site were wild-type (i.e. U2-dependent) and secondly which spliceosomal machinery (major or minor) would be used. Mutations were also introduced in the 5' splice site such that it is neither U2-type nor U12-type and the mutant was called 5'SS mut (CCGGGCUC). This mutant was constructed to completely eliminate the binding of U1 and U6 snRNAs to the splice site junctions and identify if the mutant would splice. The third mutation introduced was CC5/6GG mutation in the U125'SS mut to generate the U125'SS CC/GG mutant (referred as CC/GG mutant). Using this mutant, we wanted to know whether major or minor spliceosome, is splicing the MYH6 U125'SS mut.



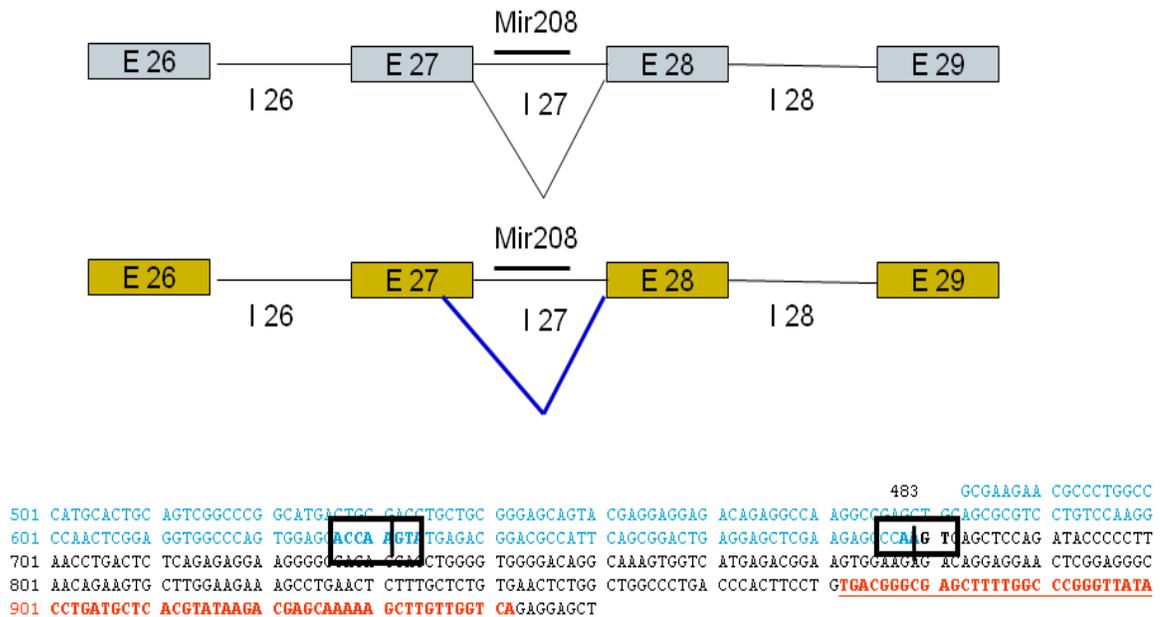
**Figure 15: *In vivo* splicing patterns of MYH6 intron 27 5' splice site mutants:** 5' splice site mutants were transfected in mammalian cells, RT-PCR was performed and analyzed for splicing. U12-type 5'SS mutant and U12 5'SS CC5/6GG mutant spliced products are similar to wild-type. However 5'SS mutant activates a cryptic splicing event.

Since, it is known that wild-type splicing gets abolished and cryptic splicing is activated if CC5/6GG is mutated in U12-dependent intron F of p120 gene. Using this construct *in*

*in vivo* suppression system, developed for U12-dependent (p120 gene) intron, could be utilized (Hall and Padgett, 1996; Kolossova and Padgett, 1997; Incorvaia and Padgett, 1998). All splice site mutants were transfected in CHO cells using polybrene and upon RNA isolation analyzed for spliced phenotypes. As expected, bands were not observed seen in case of mock (negative) control. Wild type (WT) spliced product is 263 bp long (Figure 15, lane 2). U125'SS mutant (lane 4) which had consensus U12-type 5' splice site showed phenotype similar to the wild type, suggesting either splicing takes place in mutant utilizing the minor spliceosomal pathway or there is no effect of mutation on the splicing. To answer this question, we utilized the CC5/6GG mutant (lane 5). RNA-RNA interactions between the snRNAs and pre-mRNA are the basis of splicing and any mutation in one of the interacting partners leads to inhibition of these interactions and hence affects splicing. It has been shown that CC at positions 5 and 6 in a U12-dependent intron are necessary for splicing to occur. If CC at 5 and 6 positions in a U12-dependent intron are mutated, splicing does not take place. However, MYH6 U12 5'SS CC5/6GG mutant spliced phenotype is similar to wild-type (Figure 15, lane 5). We co-transfected suppressor U11 and U6atac snRNAs along with U125' SS CC5/6GG mutant since, suppressor snRNAs are known to restore splicing. U11 snRNA with a mutated GG6/7CC and U6atac snRNA with mutated GG14/15CC are called suppressor snRNAs as they restore the base pairing interactions with CC5/6GG in the U12 5'SS CC/GG mutant. However, the results obtained were quite ambiguous and could not be reproduced. Upon careful analysis of the wild-type genomic sequence and U12 5'SS CC/GG, we found that both wild-type and CC/GG mutant have guanine (G) at the 5<sup>th</sup> position in intron 27. This observation suggested that G at 5<sup>th</sup> position could be important in splicing intron 27. This

observation was further supported by splicing phenotypes observed from *in vivo* spliced point mutants, G5A and G5C (Figure 22, lanes 8 and 9) where splicing was completely abolished (section 3.9).

The 5'SS mut which has 5' splice site mutations (GTA 124/CCG) such that the splice site is completely mutated for splicing, cryptic splicing was activated. Spliced phenotype of 5'SS mut is smaller than the wild type (Figure 15, lane 3).



**Figure 16: Affect of 5' splice site mutation:** The Figure illustrates the comparison of WT (upper panel) and cryptic (lower panel) splicing of intron 27 of MTH6. Blue line indicates that cryptic splice site is located in the exon 27. Sequence similarity of both the sites recognized is shown in the lower panel. In the lower panel, region shown in blue represents exon 27 and region in black illustrates intron 27.

The band was extracted from the gel, PCR product was eluted and its sequence was analyzed. Since in this mutant, the intronic 5' splice site GT was mutated to CC hence the splice site no longer was U2 dependent, a cryptic 5' splice site was utilized for the splicing of the intron. Spliced sequence analysis showed that there is a 50 nucleotide shift in 5' splice site upstream of the normal splice site that produced smaller splice product

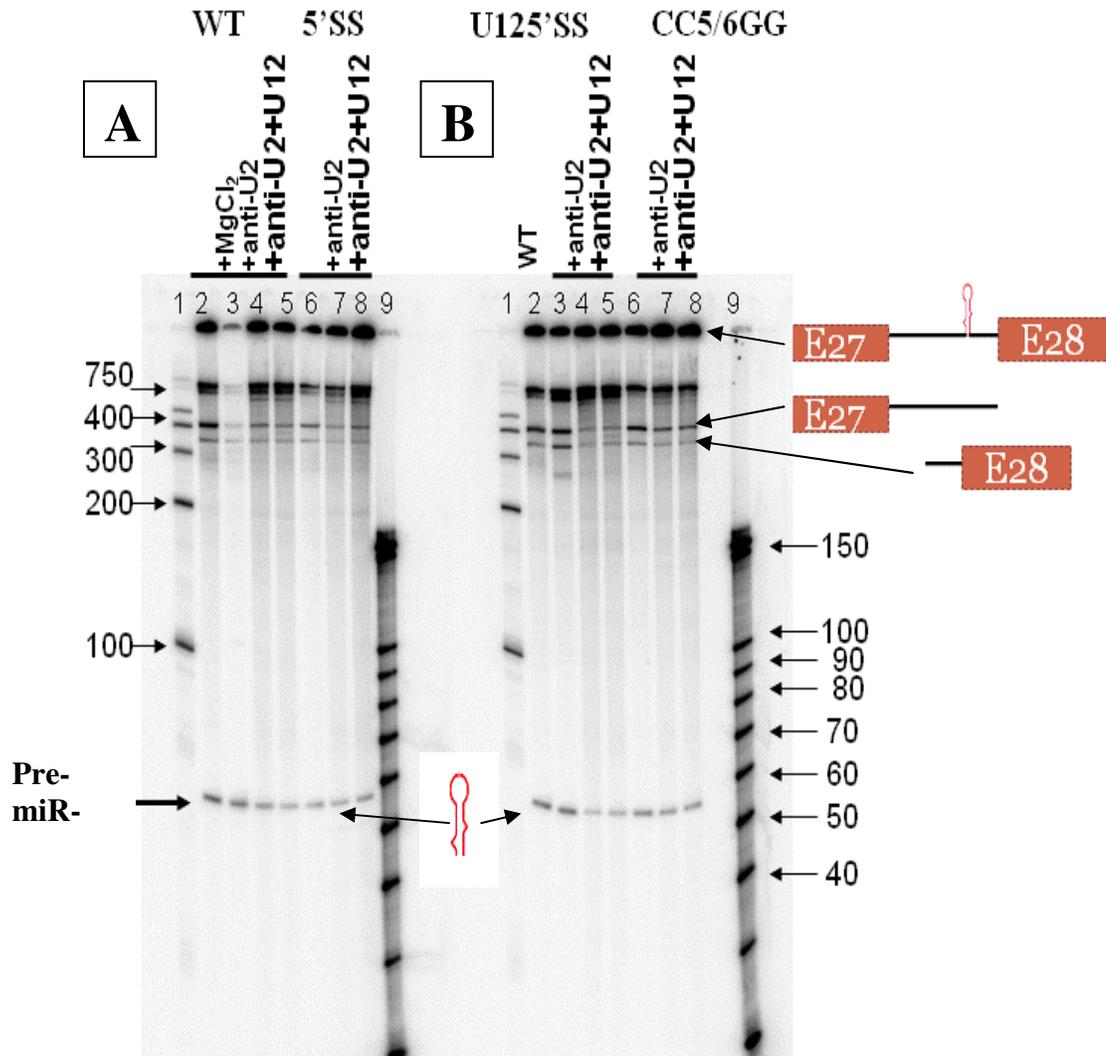
(Figure 16). The upstream cryptic 5' splice site is identical to the WT 5' splice site which is CCAA/GT.

### **3.7 *In vitro* Splicing Assay for Splice Site Mutants**

To study the effect of 5' splice site mutations on miRNA processing, we utilized the *in vitro* splicing assay system. The splice site mutants were PCR amplified from exon 27 to exon 28 including intron 27 containing miRNA and cloned in pUC 19 vector under T7 promoter. Plasmid DNA was linearized with Xba1 and the linear plasmid DNA was transcribed *in vitro* in presence of  $\alpha$ -P<sup>32</sup> radio-labeled UTP. Then *in vitro* splicing assay was performed in HeLa cell nuclear extracts as described in Chapter 2. Similar to miRNA stem-loop mutants, splice site mutants also failed to splice *in vitro* and, pre-miR208 along with Drosha cleaved mRNA products were detected (Figure 17). Anti U2 oligo to block U2-dependent splicing and to block U12 dependent splicing both anti-U2 oligo and anti-U12 oligo were added. As mentioned in section 2.8, sequence of anti-U2 oligo is complementary to U2 snRNA nucleotides 27–49 (Dietrich et al, 1997) and sequence of anti-U12 oligo is complementary to U12 snRNA nucleotides 11–28 (Dietrich et al, 1997). We wanted to confirm if MYH6 U12 5'SS mutant construct splices using major or minor spliceosomal machinery for splicing. But since the MYH6 substrates (wild-type as well as the mutants) fail to splice *in vitro* system, we could not conclude if U125'SS mutant splices using major or minor spliceosomal machinery. However, precursor miR-208 processed from wild type as well as 5'SS, U125'SS and U12 5'SS CC/GG mutants (Figure 17). It appears that miRNA processing can occur even if splice products are not

observed. MYH6 pre-mRNA is a good substrate for miRNA processing but a poor one for splicing.

### 3.8 Quantitation of Mature miR-208 Expression in Splice Site and Stem-Loop Mutants



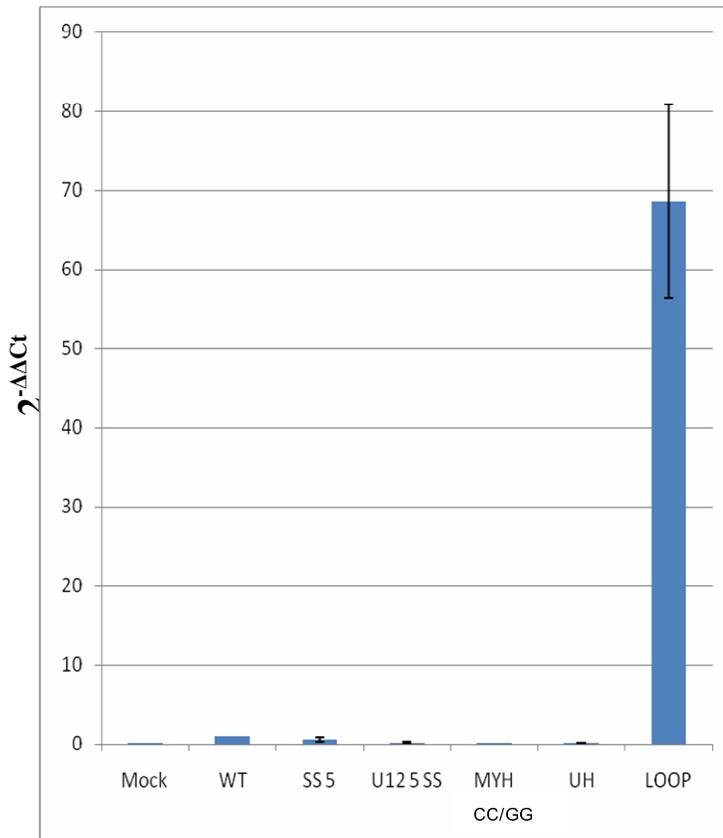
**Figure 17: *In vitro* spliced phenotypes of 5' splice site mutants:** Lane 1 and lane 9 shows century marker and decade markers, lanes 2-5 show *in vitro* spliced wild-type MYH construct. Lanes 6-8 show 5'SS mutant. Panel B lanes 3-5 show U125'SS and lanes 6-8 show U125'SS CC/GG mutant. Anti-U2 indicates lanes to which anti-U2 oligo was added and U12 indicates addition of anti-U12 oligo to block U2-dependent and U12-dependent splicing respectively.

miRNA processing is a compartmentalized process and first step takes place in the nucleus where precursor miRNA is made and then second step occurs in cytoplasm where precursor is processed to mature miRNA (Lee et al, 2002; Kim and Kim, 2007).

*In vitro* splicing assay gave us some insight about the precursor miR-208 since it takes place in the nuclear extract. We detected only the spliced product by RT-PCR, mature miRNA could be detected *in vivo* by quantitative RT-PCR (real-time PCR or qRT-PCR) which is a sensitive and efficient method to detect small RNA species.

Mature miR-208 expression from stem-loop and the splice site mutants was detected using qRT-PCR. The relative expression of miR-208 is plotted as  $2^{-\Delta\Delta Ct}$  after normalizing to snoRNA202 and then to WT. miR-208 expression in case of loop mutant

is expressed approximately 30



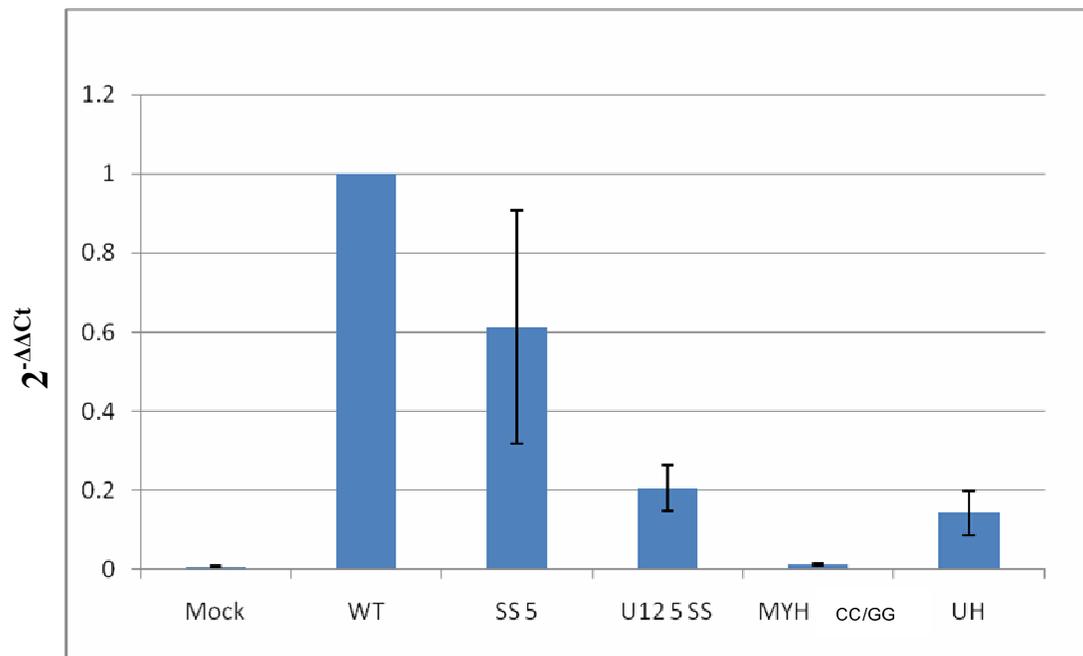
**Figure 18: Quantitative RT-PCR expression of mature miR-208 from MYH6 stem-loop and 5' splice site mutants:** The Figure shows relative expression of miR-208 expressed from various mutants. Relative expression is plotted as  $2^{-\Delta\Delta Ct}$  after normalizing to sno202 expression levels and then to WT mir-208 levels.

fold as compared to the wild type (Figure 18). miR-208

expression at such a high level from loop mutant could be due to any changes in the

secondary structure of miRNA or change in sequence at the terminal loop region could have caused such an increase. It is thought that the terminal loop region plays a very important role in miRNA processing (Kim and Kim, 2007). Figure 19 represents same graph as Figure 18 without comparing loop mutant as it becomes difficult to compare expression profile of miR-208 from other mutants with high level of expression from the loop mutant. The relative expression of miR-208 in all the mutants namely 5'SS, U12 5'SS, U12 5'SS CC5/6GG (referred as MYH CC/GG in Figure 18 and 19) and UH mutant is reduced as compared to the wild type. All these mutants besides 5'SS mut show similar splicing phenotypes as the wild-type (Figures 13 and 15).

This indicated mutations in the 5' splice site of the intron affect miRNA processing and also mutating the stem-loop structure of miRNA affect miRNA-

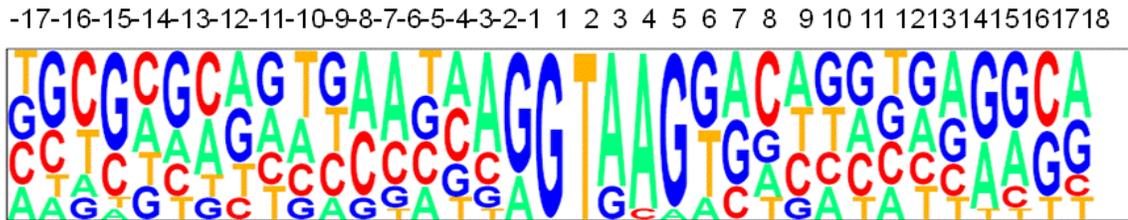


**Figure 19: Quantitative RT-PCR expression of miR-208 from MYH6 stem-loop and 5' splice site mutants:** The figure shows relative expression of miR-208 expressed from various mutant constructs. Relative expression is plotted as  $2^{-\Delta\Delta C_t}$  after normalizing to snoRNA202 expression levels and then to WT.

processing as well. Interestingly, U12 5'SS CC5/6GG almost failed to express miR-208 showing that CC5/6 are important for miR-208 processing. These results suggest there is some interplay between the splice site and miRNA processing.

### 3.9 *In vivo* Splicing Phenotypes of Point Mutants

Pictogram is a pictorial representation of analysis of nucleotides at a particular position in a DNA sequence. A pictogram representing the 5' splice site sequence content of miRNA harboring introns was designed (Figure 20). Size of nucleotide at a position is

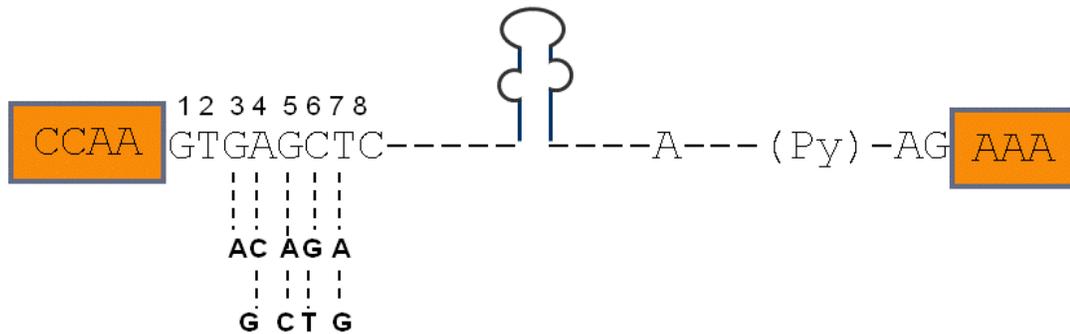


**Figure 20: Pictogram representing 5' splice site of 15 miRNA coding introns:** -1 to -17 represent the sequence at exonic splice site junction and 1 to 17 represent the bases in the intron and -1 to -17 represent the 5' exon. Bigger size of base at a position corresponds to higher conservation.

proportional to its conservation for example; A is more conserved at position 3 than G. Based on this splicing sequence content, point mutations in intron 27 of minigene were made. The 5' splice site of intron 27 of the myosin heavy chain was mutated at positions 3 through 7 (Figure 21). If the 5' splice site sequence of MYH intron 27 is compared with that of pictogram, considerable similarities and dissimilarities are seen. G and T at positions 1 and 2 respectively are completely conserved, indicating their evolutionary importance and could possibly have a role in pre-mRNA splicing and miRNA processing. Purines are largely present at positions 3-7. AAG at positions 3-5 are conserved as well,

however GCA can be present in some cases (Figure 20). Positions 6 and 7 are quite degenerate with lack of C and T at 6<sup>th</sup> and 7<sup>th</sup> position respectively.

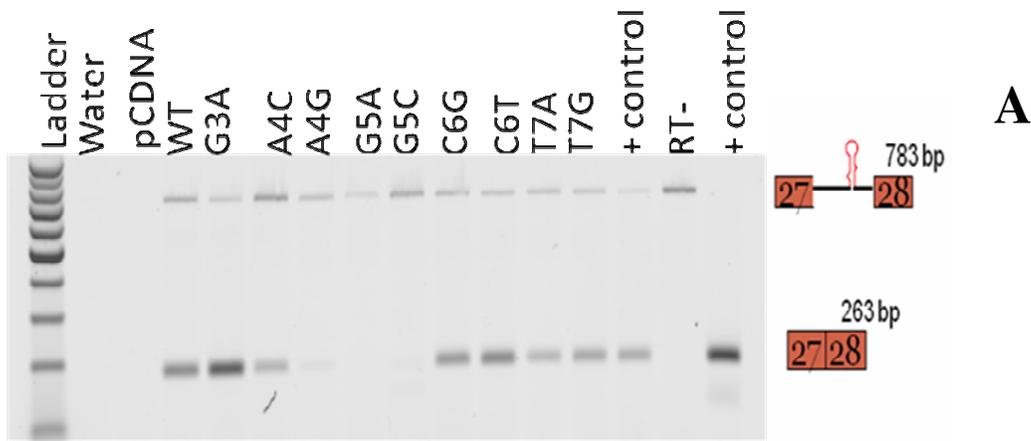
As mentioned before, the 5' splice site of MYH intron 27 was mutated based on sequence conservation of miRNA coding introns. Mutant G3A had a mutation from G to A at position 3 since A is more dominant than G at the position (Figure 21). At position 4, base A is present in majority of introns studied, although C is found as well; intron 27 also has A at 4 position so we mutated A at 4 position to C and called it A4C mutant and in A4G mutant mutated A at 4 position to G keeping in mind that G is not present in miRNA coding introns. Similarly G5A and G5C mutants were made for 5<sup>th</sup> position where A is rarely present and C is not present as per pictogram. Pyrimidine C is not present at 6<sup>th</sup> position in the pictogram, however it is present in MYH intron 27 so C6G and C6T mutants were made for position 6. For the 7<sup>th</sup> position T7A and T7G mutants were made by comparing intron 27 with the pictogram.



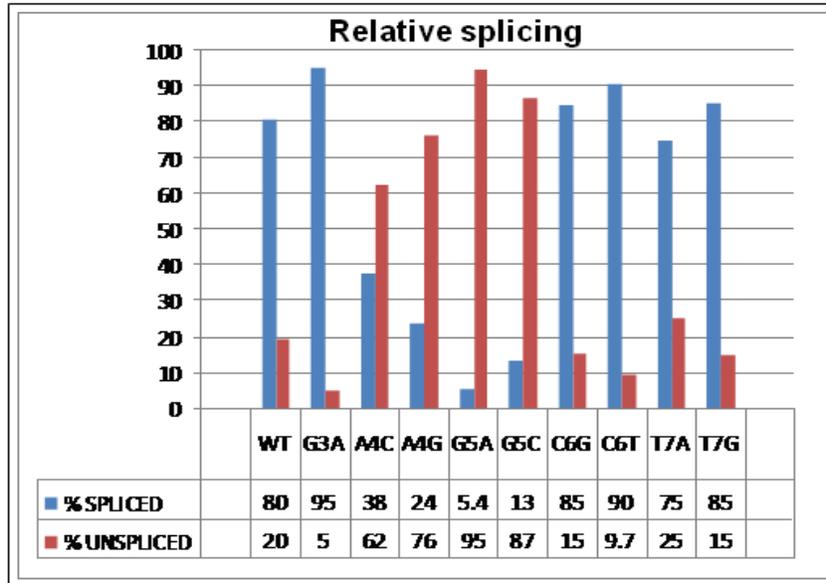
**Figure 21: Intron 27 of MYH showing 5' splice site point mutations:** Upstream and downstream exons are shown in orange boxes and miR-208 is shown in the intron. Numbers 1-8 represent the positions of bases in the intron. Introduced mutations are indicated by the dashed lines. For example, G at position 5 is mutated to A and mutant is called G5A, another mutant G5C has a mutation from G to C at 5<sup>th</sup> position.

All point mutants were transfected in CHO cells using polybrene, RNA was isolated after 36 hours and analyzed for spliced phenotypes. As expected, spliced and

unspliced product bands were not found in negative controls (Figure 22, lanes 2 and 3).



**B**



**Figure 22: *In vivo* splicing patterns of 5' splice site point mutants:** Point mutants were transfected in mammalian cells, RT-PCR was performed using the F2 and R2 primers. Panel A illustrates the affect of point mutations at positions 3 through 7 in the 5' splice site of intron 27 on splicing *in vivo*. Panel B depicts quantitation of relative splicing among the point mutants.

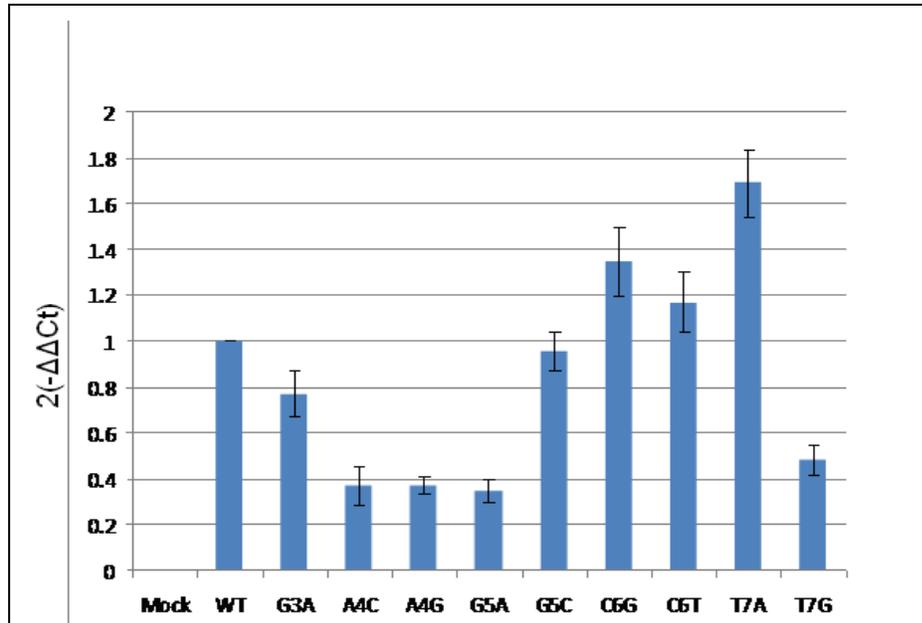
Both unspliced and spliced products were observed in the wild type as well as in the mutants. The wild-type spliced product is 263 bp long (Figure 22, lane 4). G3A mutant appears to splice better than the wild-type (lane 5).

This could be due to the fact that in adenosine is a preferred base at position 3 than G (though it is present in some introns). Splicing goes down in case of 4' position mutants (lanes 6 and 7) since adenosine is again a preferred base at the 4<sup>th</sup> position and C is seen in some introns coding miRNA. If we mutate A to C, splicing decreases but if A is mutated to G a further reduction in splicing is seen. In case of 5' position where guanine is the preferred nucleotide and was mutated to A (lane 8) and C (lane 9) negligible splicing was seen, indicating the importance of G at 5<sup>th</sup> position. As shown in the pictogram, positions 6 and 7 are quite degenerate and flexible for nucleotides, mutating C and T at 6<sup>th</sup> and 7<sup>th</sup> position respectively did not abolish splicing. However, splicing with 6<sup>th</sup> position mutants was as good as the wild-type (Figure 22, lanes 10 and 11). Splicing was little reduced as compared to the wild-type in the 7<sup>th</sup> position mutants as nucleotides A and G are also found at this position. The spliced and unspliced products obtained with the 4<sup>th</sup> and 5<sup>th</sup> position mutants (A4C, A4G, G5A and G5C) do not correspond to the total spliced and unspliced product quantity as the wild-type and other mutants. It is possible that splice site junction recognized by the spliceosome is upstream or downstream of the F2 and R2 primers and therefore, could not be amplified and detected. There is another possibility that the region of exons 27 and 28 was skipped during splicing and hence spliced and unspliced products were not observed.

### **3.10 Quantitation of Mature miR-208 Expression in Point Mutants**

qRT-PCR was performed on the same RNA that was used for *in vivo* splicing phenotype analysis for the 5' splice site point mutants so that miRNA expression could be compared with relative splicing. The relative expression of mature miR-208 normalized

to snoRNA202 was calculated and a graph was plotted (Figure 23). miR-208 expression reduces in G3A mutant as compared to the wild type. The reduction in miRNA



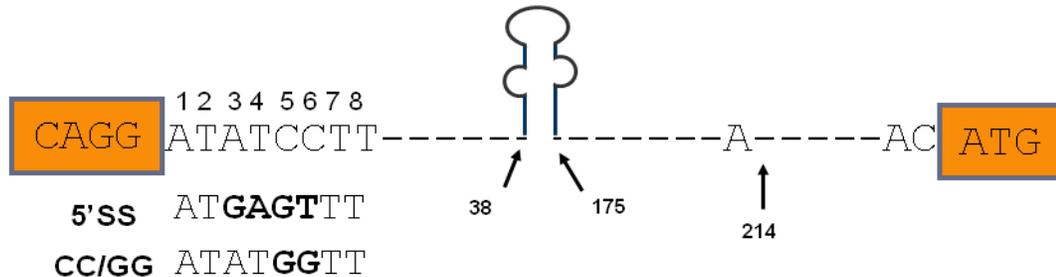
**Figure 23: Quantitative RT-PCR expression of miR-208 in point mutants:** The figure shows relative expression of miR-208 plotted as  $2^{-\Delta\Delta C_t}$  after normalizing to snoRNA202 expression levels and then to miR-208 levels from the wild-type.

expression is more significant in the A4C, A4G and G5A mutants. This indicated adenine and guanine at 4<sup>th</sup> and 5<sup>th</sup> positions is not only important for splicing (Figure 22) but also for miRNA processing. miR-208 levels from C6G, C6T and T7A mutants were higher than wild type levels indicating these mutations in the splice site enhanced the processing of miRNA. The level of relative splicing in 6<sup>th</sup> and 7<sup>th</sup> position mutants is also better than wild type suggesting these mutations not only increase splicing but also miRNA processing. qRT-PCR and the *in vivo* splicing analysis using RT-PCR show that presence of some nucleotides at a certain position is more favorable for pre-mRNA splicing or pre-miRNA processing to take place.

### 3.11 *In vivo* Splicing Phenotypes for U12-dependent miRNA System

p120 gene or human nucleolar proliferative antigen gene has a well identified U12-dependent intron, intron F, which has been characterized for both *in vivo* as well as *in vitro* splicing. Since no microRNA is known to exist in an U12-dependent intron, we were interested if U12-dependent intron would be compatible with miRNA processing if it is introduced in the intron upstream of branch site. Also we wanted to address if splicing of a U12-dependent intron would be affected by presence of miRNA.

We introduced a hsa-miR-208 encoding 137 bp fragment from intron 27 of

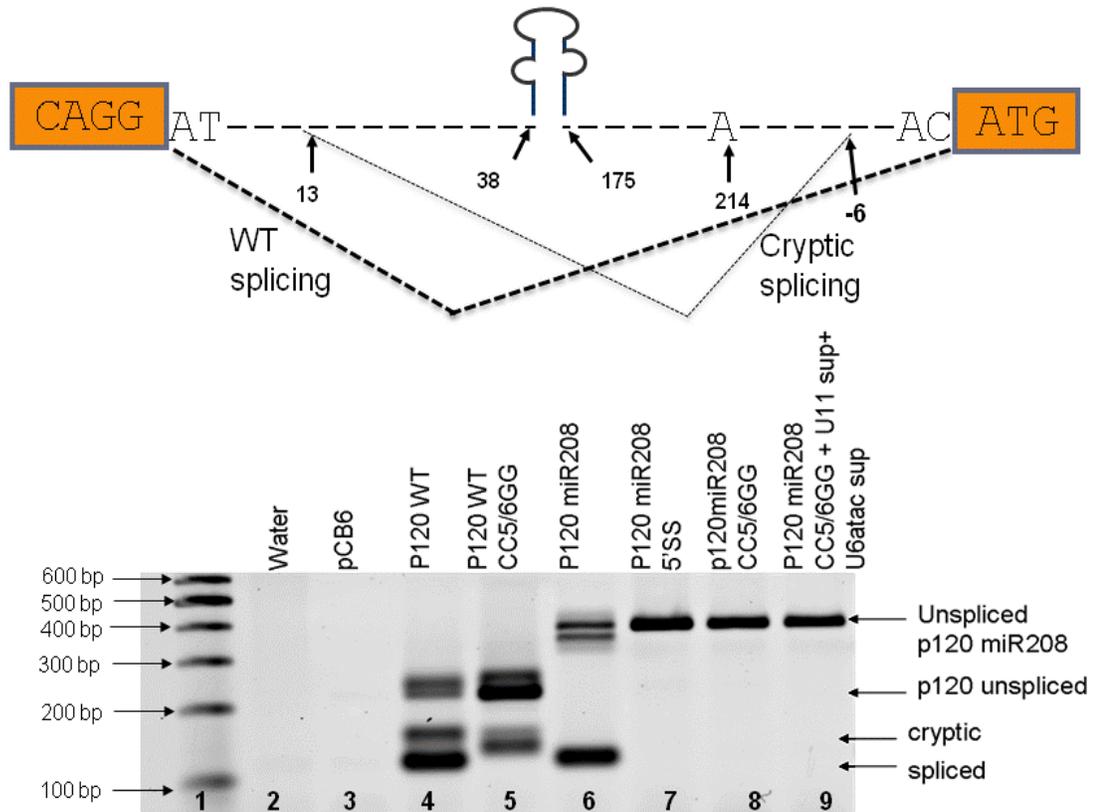


**Figure 24: Intron F of p120 minigene showing introduced miR-208 and mutations:** miR is represented by the stem-loop structure between 38 and 175 nucleotides, branch point Adenosine is marked at 214<sup>th</sup> base. Mutated bases are shown in bolds.

MYH6 in the intron F of p120 minigene reporter construct (Figure 24). We also mutated the 5' splice site region in intron F to identify the affects of these mutations and presence of miRNA on splicing and miRNA processing. It has been shown that if CC at 5<sup>th</sup> and 6<sup>th</sup> position in intron F of p120 minigene are mutated to GG a cryptic splicing event is activated as shown in Figure 25 (panel A) (Incorvaia and Padgett, 1998). It has also been shown that if U11 snRNA GG6/7CC and U6atac GG14/15CC are cotransfected with p120 intron F CC5/6GG mutations, it restores splicing as wild type. p120 minigene

construct and its mutants along with appropriate controls were transfected in CHO cells, isolated RNA and performed RT-PCR and the results are shown (Figure 25, panel B). Negative controls as expected did not show any bands (Figure 25, Panel B, lanes 2 and 3). *In vivo* spliced p120 wild type showed spliced, cryptic and unspliced bands (Figure 25, Panel B, lane 4). Lane 5 shows spliced variants for p120 CC5/6GG mutant where a cryptic splicing event takes place and hence spliced band was not observed. However, an increased level of unspliced product was also observed.

Splicing phenotypes for p120 miR-208 mutant are shown in lane 6 and in this

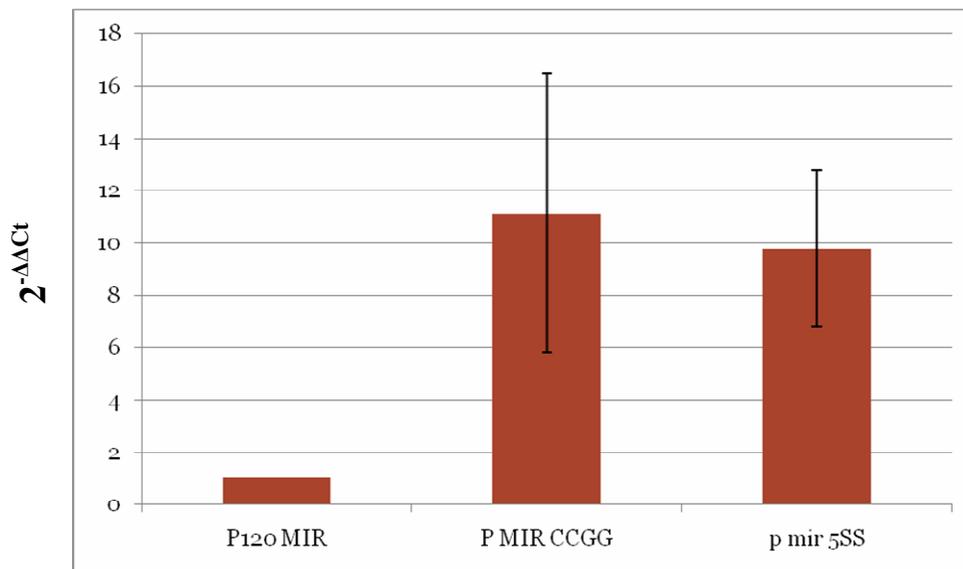


**Figure 25: *In vivo* splicing patterns of p120 mutants:** p120 mutants and wild-type was transfected in mammalian cells and RT-PCR analysis was done for spliced phenotypes. Panel A illustrates the site of miR-208 introduction and sites where cryptic or wild-type splicing takes place. Panel B shows splicing phenotypes of the p120 mutants. Marked with arrows are spliced, cryptic spliced and unspliced bands.

case cryptic splicing does not take place and only spliced or unspliced products were detected. Unspliced product was only observed upon the *in vivo* splicing analysis of p120 miR-208 5'SS and p120 miR-208 CC5/6GG mutants (Figure 25, panel B, lanes 7 and 8 respectively). Introduction of the 5' splice site mutations and/or introduction of miRNA in the intron completely abolished the splicing event. Co-transfection of U11 snRNA GG6/7CC and U6atac GG14/15CC snRNA with the p120 miR-208 CC5/6GG mutant (lane9) also did not rescue the splicing defect (as it does in p120 wild type, Incorvaia and Padgett, 1998).

### 3.12 Quantitation of Mature miR-208 Expression in U12-dependent miRNA System

qRT-PCR was performed on the same RNA that was used for *in vivo* splicing phenotype analysis for the p120 miR-208 wild type and splice site mutants so that



**Figure 26: Quantitative RT-PCR expression of miR-208 in p120 system:** The figure shows relative expression of miR-208 expressed from p120 minigene system, plotted as  $2^{-\Delta\Delta C_t}$  after normalizing to sno202 expression levels and then to miR-208 levels from the p120miR construct.

miRNA expression could be compared with relative splicing. The relative miR-208 expression normalized to sno202 and then to p120 miR is plotted (Figure 26). Since, microRNA is not known to exist in a U12-dependent intron; we were interested in finding if microRNA would be processed from U12-dependent intron. miR-208 expression for p120 miR208 CC5/6GG mutant is approximately 11 folds the level of p120 miR208 wild type and it is about 10 folds for p120 miR208 5' SS mutant (Figure 26). But if these results are compared with the *in vivo* splicing results (Figure 25), we do not observe a positive co-relation between the two processes as splicing is completely abolished in these mutants; however, miRNA processing is better than the wild type. It apparently appears that Drosha (microprocessor) is actively processing the miRNA to its precursor but splicing is not taking place. It could be due to the fact that the splicing machinery and miRNA processing machinery cannot act together in a U12-dependent intron. If this is the case then which component of spliceosome is of utmost importance and what does microprocessor and spliceosome share in common remains unanswered.

## CHAPTER IV

### CONCLUSIONS AND FUTURE DIRECTIONS

#### 4.1 Conclusions

Using the MYH6 minigene construct, we were able to utilize CHO cell-based transient transfection system to study pre-mRNA splicing and miRNA processing. miR-208 expresses from the intron 27 of MYH6 gene, in the heart and traceably in lungs. Our results show that MYH6 expresses in heart, kidney, lung, sparsely in brain but none in liver, indicating miR-208 expression is tissue-specific or regulated post-transcriptionally as MYH6 expresses in kidneys and lungs where miR-208 does not express. There could be some additional factors that are required for miR-208 expression, which are present in heart and lungs, aiding in its efficient processing. Since, miR-208 is located intronically, effect of miR sequence (and structure) on pre-mRNA splicing was studied. Mutations in the predicted stem-loop structure of miR-208 do not affect pre-mRNA splicing. But since the mutations were introduced in the predicted stem-loop regions of miR-208, namely, lower helix, upper helix and loop regions, the effect on miR expression level was expected. The *in vitro* splicing assay in cell-free extract using the stem loop mutants showed that the MYH6 pre-mRNA is a poor substrate for *in vitro* splicing but gave some

insight on miRNA processing. Pre-miR-208 was detected in the wild-type and loop mutants but it was not detected in upper helix and lower helix mutants suggesting importance of the helical regions in miRNA processing. We performed the quantitative RT-PCR to detect the expression levels of mature miR-208 from these mutants. Expression of miR-208 from loop mutant is approximately 30 folds compared to the wild type, on the other hand, its expression from upper helix mutant reduced 5 folds. Since, miR-208 is intronically located, splice site sequences could also influence miRNA processing so the 5' splice site region of intron 27 (intron harboring miR-208) was mutated. *In vivo* splicing showed that the U12 5'SS mutant and CC5/6GG mutant splice similar to the wild-type but 5'SS mutant activated a cryptic splicing event. DNA sequencing on the PCR product revealed that splicing takes place from a site located in the exon 27. The sequence is located 50 nucleotides upstream of wild-type site and is identical to the wild-type sequence. Level of mature miR-208 expression was detected by qRT-PCR studies and was reduced as compared to the wild-type levels. miR-208 expression from 5'SS was reduced approximately 5 folds as compared to the wild-type and negligible from the CC5/6GG mutant. To study how flexible and permissive is the 5' splice site to allow miRNA processing and splicing to occur, point mutations were introduced in 5' splice site of intron 27. G3A mutant splices better than the wild-type and miR-208 levels are comparable to the wild type. This could be due to the fact that adenosine is a preferred base at position 3 than G (though it is present in some introns). Splicing goes down in case of 4' position mutants (A4C and A4G) since adenosine is again a preferred base at the 4<sup>th</sup> position and C is seen in some introns coding miRNA. If A is mutated to C, splicing decreases but if A is mutated to G a further reduction in

splicing is seen. About 2 folds reduction in miR-208 expression was observed with both A4C and A4G mutants. In case of 5' position where guanine is the preferred nucleotide and was mutated to A or C, negligible splicing was detected, indicating the importance of G at 5<sup>th</sup> position. About 2 folds decrease in miR-208 expression was detected for G5A mutant; however, level of expression for G5C mutant was comparable to the wild-type. Positions 6 and 7 are quite degenerate and flexible for nucleotides, mutating C and T at 6<sup>th</sup> and 7<sup>th</sup> position respectively did not abolish splicing. However, splicing with 6<sup>th</sup> position mutants (C6G and C6T) was as good as the wild-type but miR-208 expression from both mutants was better than the wild type. Splicing was little reduced as compared to the wild-type in the 7<sup>th</sup> position mutants since nucleotides A and G are also found at this position in some miRNA containing introns. T7A expressed miR-208 about 1.7 folds compared to the wild-type, on the other hand, 2 folds decrease was observed in T7G mutant. It appears that G, A and G are required at 3<sup>rd</sup>, 4<sup>th</sup> and 5<sup>th</sup> positions respectively for both pre-mRNA splicing and miR-208 processing to occur. Till now, miRNAs have been detected in U2-dependent introns but none in U12-dependent introns. After introduction of miR-208 in intron F (U12-dependent) of p120 minigene and mutating 5' splice site of intron F, in vivo splicing and qRT-PCR analysis was done. We observed that p120 with miR-208 splices but its mutants (5'SS and CC5/6GG) failed to splice. However, miR-208 expression was 3 and 4 folds in CC5/6GG and 5'SS mutants respectively as compared to the p120 with miR-208. We did not observe a positive co-relation between the two processes as splicing is completely abolished in these mutants; however, miRNA processing is better than the wild type. It appears that Drosha (microprocessor) is actively processing the miRNA but splicing is not taking place if miR-208 is present in minor

class intron in case of the p120 miR-208 CC5/6GG and p120 miR-208 5'SS mutants. It could be due to the fact that the splicing machinery and miRNA processing machinery cannot act together in U12-dependent intron where miR-208 has been incorporated and 5' splice site has been mutated.

#### **4.2 Future Directions**

We want to establish if miRNA processing and splicing in both, U2-dependent or U12-dependent intron are inter-linked. It can be proven by the *in vitro* splicing assay in cell-free extracts. *In vitro* splicing assay also aids to determine various intermediates and hence understand basic mechanism coupling the two processes. Since the MYH6 substrate used before was splicing deficient, a new construct which would enable studying miRNA processing and pre-mRNA splicing is required. It can be achieved by introduction of certain cis-element sequences essential for splicing. We have detected the mature miR-208 expression levels in wild type as well as the mutants, but precursor miR-208 levels can also be detected using qRT-PCR. We want to develop an *in vivo* suppression assay system for intron 27 of MYH6 minigene. We would utilize the *in vivo* suppression assay system for intron F of p120 minigene and suppress splicing in CC5/6GG mutant using the U11 and U6atac snRNA suppressors. Finally, we would identify the protein/factor involved if pre-mRNA splicing and miRNA processing are coupled.

## BIBLIOGRAPHY

1. Bartel, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, 116(2): 281-97.
2. Baskerville, S. and Bartel, D.P. (2005). Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. *RNA*, 11: 241–247.
3. Bentwich, I. (2005). Prediction and validation of microRNAs and their targets. *FEBS Letters*, 579: 5904–5910.
4. Black, D.L. (2003). Mechanisms of alternative pre-messenger RNA splicing. *Annual Reviews of Biochemistry*, 72(1): 291–336.
5. Bohnsack, M.T., Czaplinski, K., Gorlich, D. (2004). Exportin 5 is a Ran-GTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA*, 10: 185–191.
6. Borchert, G.M., Lanier, W., Davidson, B.L. (2006). RNA polymerase III transcribes human microRNAs. *Nature Structural and Molecular Biology*, 13: 1097–1101.
7. Brain, S.D., Williams, T.J., Tippins, J.R., Morris, H.R., MacIntyre, I. (1985). Calcitonin gene related protein is a potent vasodilator. *Nature*, 313(5997): 54-56.
8. Breathnach, R., Benoist, C., O'Hare, K., Gannon, F., Chambon, P. (1978). Ovalbumin gene: evidence for a leader sequence in mRNA and DNA sequences at the exon-intron boundaries. *PNAS*, 75: 4853–4857.
9. Cai, X., Hagedorn, C.H., Cullen, B.R. (2004). Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA*, 10: 1957–1966.

10. Callis, T.E., Pandya, K., Seok, H.Y., Tang, R.H., Tatsuguchi, M., Huang, Z.P., Chen, J.F., Deng, Z., Gunn, B., Shumate, J., Willis, M.S., Selzman, C.H., Wang, D.Z. (2009). MicroRNA-208a is a regulator of cardiac hypertrophy and conduction in mice. *J. Clin. Invest.*, 119(9): 2772-2786.
11. Chen, C.Y., Gherzi, R., Ong, S.E., Chan, E.L., Raijmakers, R., Pruijn, G.J.M., Stoecklin, G., Moroni, C. (2001). AU binding proteins recruit the exosome to degrade ARE-containing mRNAs. *Cell*, 107(4): 451–464.
12. Chomczynski, P. and Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.*, 162: 156-9.
13. Copp, D.H. and Cheney, B. (1962). Calcitonin-a hormone from the parathyroid which lowers the calcium-level of the blood. *Nature*, 193: 381-2.
14. Cullen, B.R. (2004). Transcription and processing of human microRNA precursors. *Mol. Cell*, 16: 861–865.
15. Dietrich, R.C., Inoravia, R., Padgett, R.A. (1997). Terminal intron dinucleotide sequences do not distinguish between U2- and U12-dependent introns. *Mol. Cell*, 1: 151–160.
16. Friedman, R.C., Farh, K.K.H., Burge, C.B., Bartel, D.P. (2009). Most mammalian mRNAs are conserved targets of microRNAs. *Genome Research*, 19(1): 92–105.
17. Griffiths-Jones, S., Saini, H.K., van Dongen, S., Enright, A.J. (2008). miRBase: tools for microRNA genomics. *Nucleic Acids Res.*, 36: 154-158.
18. Grunstein, M. (1997). Histone acetylation in chromatin structure and transcription. *Nature*, 389: 349-352.

19. Hall, S.L. and Padgett, R. A. (1996). Requirement of U12 snRNA for in vivo splicing of a minor class of eukaryotic nuclear pre-mRNA introns. *Science*, 271: 1716–1718.
20. Hall, S.L., and. Padgett, R.A. (1994). Conserved sequences in a class of rare eukaryotic nuclear introns with non-consensus splice sites. *J. Mol. Biol.*, 239: 357–365.
21. Hartwell, L.H., Hood, L., Goldberg, M.L., Reynolds, A.E., Silver, L.M., Veres, R.C. (2004). *Genetics: From genes to genomes*, 5<sup>th</sup> Edition.
22. Hayashita, Y., Osada, H., Tatematsu, Y., Yamada, H., Yanagisawa, K., Tomida, S., Yatabe, Y., Kawahara, K., Sekido, Y., Takahashi, T. (2005). A polycistronic microRNA cluster, miR-17-92, is overexpressed in human lung cancers and enhances cell proliferation. *Cancer Research*, 65: 9628-9632.
23. Hirsch, P.F., Gauthier, G.F., Munson, P.L. (1963). Thyroid hypocalcemic principle and recurrent laryngeal nerve injury as factors affecting the response to parathyroidectomy in rats. *Endocrinology*, 73: 244–252.
24. Incorvaia, R. and Padgett, R.A., (1998). Base pairing with U6atac snRNA is required for 5' splice site activation of U12-dependent introns in vivo. *RNA*, 4: 709–718.
25. Jackson, I.J. (1991). A reappraisal of non-consensus mRNA splice sites. *Nucleic Acids Res.*, 19: 3795–3798.
26. Kandels-Lewis, S. and Seraphin, B. (1993). Role of U6 snRNA in 5' splice site selection. *Science*, 262: 2035-2039.
27. Kataoka, N. and Ohno, F.M. (2009). Functional association of the microprocessor complex with the spliceosome. *Mol. Cell Biol.*, 29(12): 3243-54.

28. Ketting, R.F., Fischer, S.E., Bernstein, E., Sijen, T., Hannon, G.J., Plasterk, R.H. (2001). Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes Dev.*, 15: 2654–2659.
29. Kim, V.N. and Nam, J.W. (2006). Genomics of microRNA. *Trends Genet.*, 22: 165–173.
30. Kim, V.N. (2005). Small RNAs: classification, biogenesis, and function. *Mol. Cell*, 19: 1–15.
31. Kim, Y.K. and Kim, V.N. (2007). Processing of intronic microRNAs. *EMBO J.*, 26: 775–783.
32. Kolossova, I. and Padgett, R.A. (1997). U11 snRNA interacts in vivo with the 5' splice site of U12-dependent (AU-AC) pre-mRNA introns. *RNA*, 3: 227–233.
33. Konarska, M.M. and Sharp, P.A. (1987). Interactions between small nuclear ribonucleoprotein particles in formation of spliceosomes. *Cell*, 49: 763-774.
34. Konarska, M.M., Padgett, R.A., Sharp, P.A. (1985). Trans splicing of mRNA precursors in vitro. *Cell*, 42(1): 165-171.
35. Lagos-Quintana, M., Rauhut, R., Lendeckel, W., Tuschl, T. (2001) Identification of novel genes coding for small expressed RNAs. *Science*, 294: 853–858.
36. Laios, A., O'Toole, S., Flavin, R., Martin, C., Kelly, L., Ring, M., Finn, S.P., Barrett, C., Loda, M., Gleeson, M., D'Arcy, T., McGuinness, E., Sheils, O., Sheppard, B., O'Leary, J. (2008). Potential role of miR-9 and miR-223 in recurrent ovarian cancer. *Molecular Cancer*, 7: 35.
37. Lee, R.C. and Ambros, V. (2001). An extensive class of small RNAs in *Caenorhabditis elegans*. *Science*, 294: 862–864.

38. Lee, R.C., Feinbaum, R.L., and Ambros, V. (1993). The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell*, 75: 843–854.
39. Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Radmark, O., Kim, S., Kim, V.N. (2003). The nuclear RNase III Droscha initiates microRNA processing. *Nature*, 425: 415–419.
40. Lee, Y., Jeon, K., Lee, J.T., Kim, S., Kim, V.N. (2002). MicroRNA maturation: stepwise processing and sub-cellular localization. *EMBO J.*, 21: 4663–4670.
41. Lee, Y., Kim, M., Han, J., Yeom, K.H., Lee, S., Baek, S.H., Kim, V.N. (2004). MicroRNA genes are transcribed by RNA polymerase II. *EMBO J.*, 23: 4051–4060.
42. Li, Q., Peterson, K.R., Fang, X., Stamatoyannopoulos, G. (2002). Locus control regions. *Blood*, 100: 3077-3086.
43. Li, Q.L., Zhou, B., Powers, P., Enver, T., Stamatoyannopoulos, G. (1990). Beta-globin locus activation regions: conservation of organization, structure, and function. *PNAS*, 87: 8207-8211.
44. Madhani, H.D. and Guthrie, C. (1994). Dynamic RNA-RNA interactions in the spliceosome. *Annu. Rev. Genet.*, 28: 1-26.
45. Montzka, K.A. and Steitz, J.A. (1988). Additional low-abundance human small nuclear ribonucleoproteins: U11, U12. *PNAS*, 85: 8885–8889.
46. Moore, M.J., Query, C.C., Sharp, P.A. (1993). Splicing of precursors to mRNA by the spliceosome in R.F. Gesteland and J.F. Atkins (ed.) "The RNA World" (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY), pp 303-358.

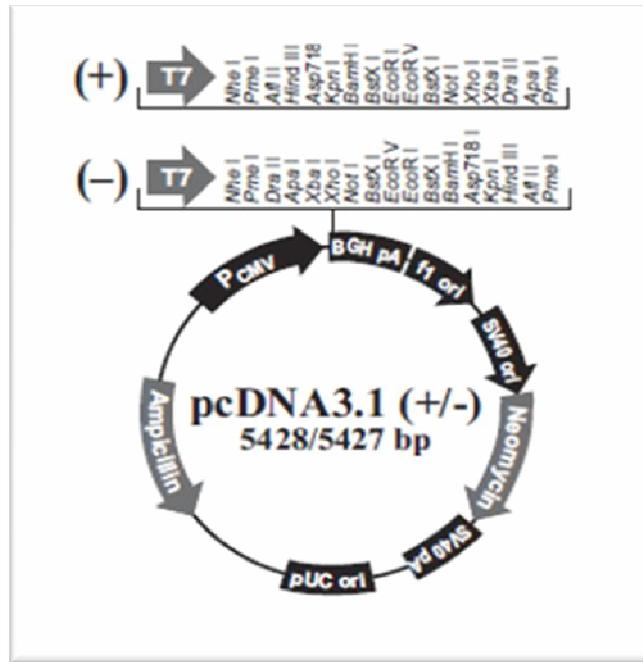
47. Morkin, E. (2000). Control of cardiac myosin heavy chain expression. *Microsc. Res. Tech.*, 50(6): 522-531.
48. Nilsen, T.W. (1993). Trans-splicing of nematode premessenger RNA. *Annu. Rev. Microbiol.*, 47: 413-40.
49. Nilsen, T.W. (1996). A Parallel Spliceosome. *Science*, 273: 1813.
50. Nilsen, T.W. (1998). RNA structure and function, RNA-RNA interactions in nuclear pre-mRNA splicing, eds. Simons, R.W., Grunberg-Manago, M., (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY), pp 279–307.
51. Poy, M.N., Eliasson, L., Krutzfeldt, J., Kuwajima, S., Ma, X., MacDonald, P.E., Pfeffer, S., Tuschl, T., Rajewsky, N., Rorsman, P., Stoffel, M. (2004). A pancreatic islet-specific microRNA regulates insulin secretion. *Nature*, 432: 226-230.
52. Poy, M.N., Hausser, J., Trajkovski, M., Braun, M., Collins, S., Rorsman, P., Zavolan, M., Stoffel, M. (2009). miR-375 maintains normal pancreatic alpha- and beta-cell mass. *PNAS*, 106(14):5813-8.
53. Proudfoot, N.J., Furger, A., Dye, M.J. (2002). Integrating mRNA processing with transcription. *Cell*, 108: 501-512.
54. Richards, A., Luccarini, C., Pope, F.M. (1997). The structural organization of LAMA4, the gene encoding laminin alpha4. *Eur. J. Biochem.*, 248: 15–23.
55. Rodriguez, A., Griffiths-Jones, S., Ashurst, J.L., Bradley, A. (2004). Identification of mammalian microRNA host genes and transcription units. *Genome Res.*, 14: 1902–1910.
56. Sharp, P.A. (1985). On the Origin of RNA Splicing and Introns. *Cell*, 42: 397-400.

57. Shatkin, A.J. and Manley, J.L. (2000). The ends of the affair: Capping and polyadenylation. *Nat. Struct. Biol.*, 7: 838-42.
58. Sikand, K., Slane, S.D., Shukla, G.C. (2009). Intrinsic expression of host genes and intronic miRNAs in prostate carcinoma cells. *Cancer Cell Int.*, 9: 21.
59. Siliciano, P.G. and Guthrie, C. (1988). 5' splice site selection in yeast: Genetic alterations in base pairing with U1 reveal additional requirements. *Genes and Dev.*, 2: 1258–1267.
60. Sperling, J., Azubel, M., Sperling, R. (2008). Structure and function of the pre-mRNA splicing machine. *Structure*, 16: 1605-1615.
61. Staley, J.P. and Guthrie, C. (1998). Mechanical devices in the spliceosome: Clocks, motors, springs and things. *Cell*, 92: 315-326.
62. Sutton, R.E. and Boothroyd, J.C. (1986). Evidence for trans splicing in trypanosomes. *Cell*, 47(4): 527-35
63. Takahashi, S., Onodera, K., Motohashi, H., Suwabe, N., Hayashi, N., Yanai, N., Nabesima, Y., Yamamoto, M. (1997). Arrest in primitive erythroid cell development caused by promoter-specific disruption of the GATA-1 gene. *J. Biol. Chem.*, 272(19): 12611-5.
64. Tang, G.Q. and Maxwell, E.S. (2008). *Xenopus* microRNA genes are predominantly located within introns and are differentially expressed in adult frog tissues via post-transcriptional regulation. *Genome Res.*, 18: 104–112.
65. Tarn, W.Y. and Steitz J.A. (1996 (b)). Highly diverged U4 and U6 small nuclear RNAs required for splicing rare AT-AC introns. *Science*, 273: 1824– 1832.

66. Tarn, W.Y. and Steitz, J.A. (1996 (a)). A novel spliceosome containing U11, U12, and U5 snRNPs excises a minor class (AT-AC) intron in vitro. *Cell*, 84: 801–811.
67. Tay, Y.M.S., Tam, W.L., Ang, Y.S., Gaughwin, P.M., Yang, H., Wang, W., Liu, R., George, J., Ng, H.H., Perera, R.J., Lufkin, T., Rigoutsos, I., Thomson, A.M., Lim, B. (2008). MicroRNA-134 modulates the differentiation of mouse embryonic stem cells, where it causes post-transcriptional attenuation of nanog and LRH1. *Stem Cells*, 26(1): 17 -29.
68. Tycowski, K.T., Kolev, N.G., Conard, N.K., Fok, V., Steitz, J.A. (2006). The ever-growing world of small nuclear ribonucleoproteins. In *The RNA World*, Third Edition, R.F. Gesteland, T.R. Cech, and J.F. Atkins, eds. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press), pp. 327–368.
69. vanRooij, E., Liu, N., Olson, E.N. (2008). MicroRNAs flex their muscles. *Trends Genet.*, 24(4): 159-66.
70. vanRooij, E., Sutherland, L.B., Qi, X., Richardson, J.A., Hill, J., Olson, E.N. (2007). Control of stress-dependent cardiac growth and gene expression by a microRNA. *Science*, 316(5824): 575-9.
71. Wang, Q., Li, Y.C., Wang, J., Kong, J., Qi, Y., Quigg, R.J., Li, X. (2008). miR-17-92 cluster accelerates adipocyte differentiation by negatively regulating tumor-suppressor Rb2/p130. *PNAS*, 105(8): 2889–2894.
72. Weiss, A. and Leinwand, L.A. (1996). The mammalian myosin heavy chain gene family. *Ann. Rev. of Cell Biol.*, 12: 417-39.
73. Will, C.L. and Luhrmann, R. (2001). Spliceosomal UsnRNP biogenesis, structure and function. *Curr. Opin. Cell Biol.*, 13: 290–301.

74. Will, C.L. and Luhrmann, R. (2006). Spliceosome structure and function. In *The RNA World*, Third Edition, R.F. Gesteland, T.R. Cech, and J.F. Atkins, eds. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press), pp. 369–400.
75. Wu, J.A. and Manley, J.L. (1989). Mammalian pre-mRNA branch site selection by U2 snRNP involves base pairing. *Genes Dev.*, 3: 1553-1561.
76. Yi, R., Qin, Y., Macara, I.G., Cullen, B.R. (2003). Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev.*, 17: 3011–3016.
77. Ying, S.Y. and Lin, S.L. (2009). Intron-mediated RNA interference and microRNA biogenesis. *Methods Mol. Biol.*, 487: 387-413.
78. Ying, S.Y., Chang, C.P., Lin, S.L. (2010). Intron-mediated RNA interference, intronic microRNAs, and applications. *Methods Mol. Biol.*, 629: 205-37.
79. Zhuang, Y. and Weiner, A.M. (1986). A compensatory base change in U1 snRNA suppresses a 5' splice site mutation. *Cell*, 46: 827–835.

## APPENDIX



**Appendix 1: Vector Map pcDNA3.1- :** It has ampicillin (Amp) as selection marker for *E.Coli* and Neomycin as selection in mammalian cells. It has a human cytomegalovirus (CMV) promoter. The multiple cloning site showing restriction enzyme sites is represented. Genomic region from exon 26 to 29 including introns from MYH 6 gene was cloned between Xba1 and Bam H1 sites.

**GAGAGTTGGCCCGGCAGCTAGAGGAAAAGGAGGCGCTAATCTCGCAGCTGACCC**  
**GGGGGAAGCTCTCTTATACCCAGCAAATGGAGGACCTCAAAGGCAGCTGGAGG**  
**AGGAGGGCAAG**gtgaggccagtgaggaggggtgggcaggcttgatggcagccct  
 ggggcaattcatctcagtgccagaaatggagcctggagctggaaagagtcctct  
 gcaagggaaagaccctccagtcctaggttctgcccctgcagctaagcgtcatttaa  
 tgccctcttttcttattcgttaaggggatgggggtgagcagactgggaaactcctca  
 aacagtgaggtgccacatcagcccacatggtgaataaggctgggcttggttgaa  
 gtactacataagaagagaatctagagaatggggcacagggagtcctcccacct  
 cctggtgccccccccctcccag**GCGAAGAACGCCCTGGCCCATGCACTGCAG**  
**TCGGCCCGGCATGACTGCGACCTGCTGCGGGAGCAGTACGAGGAGGAGACAGAG**  
**GCCAAGGCCGAGCTGCAGCGCTCCTGTCCAAGGCCAACTCGGAGGTGGCCAG**  
**TGGAGGACCAAGTATGAGACGGACGCCATTCAGCGGACTGAGGAGCTCGAAGAG**  
**GCCAA**gtgagctccagataccccctaacctgactctcagagaggaagggggcga  
 gaggacctgggggtggggacaggcaagtggtcatgagacggaagtggaagagac  
 aggaggaactcggaggggcaacagaagtgcttggaaagaaagcctgaactcttgc  
 tctgtgaactctggtgcccctgaccacttccctgtgacggggcagcttttggc  
**ccgggttatacctgatgctcacgtataagacgagcaaaaagcttgttggtcaga**  
 ggagctaccgtcgatcagcctgtgtgggggggtgagggcaggggggactgacacc  
 cagatgccactgcaggtagggaggacgcctgggcagcccgtgctggcggactct  
 gttccaggcatgagcaggctcagctcctgctaggtggacttacggtgtctcaa  
 ggagatatagggaggggggtggaaggaggtccaccaaggctccagtggtgcca  
 gtagagtcacacacacaccctccaccctcacctgggcag**AAAGAAGCTGGCCCA**  
**GCGGCTGCAGGATGCCGAGGAGGCCGTGGAGGCTGTTAATGCCAAGTGCTCCTC**  
**ACTGGAGAAGACCAAGCACCGGCTACAGAATGAGATAGAGGACTTGATGGTGGGA**  
**CGTAGAGCGCTCCAATGCTGCTGCTGCAGCCCTGGACAAGAAGCAGAGAACTT**  
**TGACAAG**gtggaccatgggcgggggccgcagccagcatgcagggcaagggggca  
 tgaggggttcagtgagaggccaaaggcaacctccttggaggtggaggaggagg  
 ctaagcccaggctcgggaccaggacagatcttggacatgcggctgaggctggg  
 ggctggggcactgggaagcaggagggtggggagctaaggctggggggctgaag  
 agtgagccttgtccccgggcag**ATCCTGGCCGAGTGAAGCAGAAGTATGAGGA**  
**GTCGCAGTCTGAGCTGGAGTCCTCACAGAAGGAGGCTCGCTCCCTCAGCACAGA**  
**GCTCTTCAAGCTCAAGAACGCCTACGAGGAGTCCCTGGAGCACCTAGAGACCTT**  
**CAAGCGGGAGAACAAGAACCTTCAGG**

**Appendix 2: Sequence of MYH 6:** The cloned genomic sequence of human Myosin Heavy Chain 6 gene. Exons are shown in bold and capital letters. Intronic sequences (intron 26-28) are shown in small letters. hsa-miR-208 is bolded in the intronic region and mature miRNA is underlined.

**Appendix 3: Sequences of Primers:** All primers are in 5' to 3' direction.

1. MYH Xho F: GCCTCGAGGAGAGTTGGCCCGGCAG
2. MYH BamH1 R: ACGGGATCCCCTGAAGGTTCTTGTC
3. F1: GGAAAAGGAGGCGCTAATCT
4. R1: GCGTCCGTCTCATACTTGGT
5. F2: GGAGCAGTACGAGGAGGAGA
6. R2: CGTCCACCATCAAGTCCTCT
7. F3: GGCTGTTAATGCCAAGTGCT
8. R3: TCGTAGGCGTTCTTGAGCTT
9. Exon Junction F: GAGGGCAAGGCGAAGAAC
10. Exon Junction R: GGCCAGGATCTTGTCAAAGT
11. 208 5'SS F: GAGGCCAACCGGGCTCCAG
12. 208 5'SS R: CTGGAGCCCGGTTGGCCTC
13. 208 U12 5'SS F: GAGGCCAAGTATCCTTCAGATACCC
14. 208 U125'SS R: GGGTATCTGAAGGATACTTGGCCTC
15. 208 U12BS F: CACACACATCCTTAACCCTCACC
16. 208 U12BS R: GGTGAGGGTTAAGGATGTGTGTG
17. 208 LH F: CTGTGACCGCTCGAAAAGGCCCGGG
18. 208 LH R: CCCGGGCCTTTTCGAGCGGCGTCACAG
19. 208 UH F: GGCCCGGGAATATCCTGATGCT
20. 208 UH R: AGCATCAGGATATTCCCGGGCC
21. Loop mutant: 5' CCCGGGTTATACCACTACGAGTCGTATAAGACGAG

22. 208 CC/GG primer: GAGGCCAAGTATGGTTCAG
23. MYH T7BamH1:  
ATATGGATCCTAATACGACTCACTATAGGGGGAGCAGTACGAGGAGGAGA
24. MYH Xba1 reverse: ACGGTCTAGACGTCCACCATCAAGTCCTCT
25. GAPDH Forward: GGAGAAACCTGCCAAGTATGATG
26. GAPDH Reverse: TGAAGTCGCAGGAGACAACCT
27. P120\_Not repair: GAGGAAGGCGGCCGCGAGAGGG
28. P120\_Pme Repair: GAGAGGGGTTTAAACGCCTTAG
29. Mir208 pri not Forward: GAATGCGGCCGCTGAACTCTGGCTGGCCCTGA
30. Mir208 pri pme Reverse: GCATGTTTAAACCACACAGGCTGATCGACGGT
31. P120 5'SS mut: GATGGAGCAGGATGAGTTTGCAGGGCAGAG
32. P120 CC5/6GG mut: TGGAGCAGGATATGGTTGCAGGGCAGAG
33. G3A: CCAAGTAAGCTCC
34. G3C: CCAAGTCAGCTCC
35. A4C: CAAGTGCCTCCA
36. A4G: CAAGTGGGCTCCA
37. G5A: AAGTGA ACTCCAG
38. G5C: AAGTGACCTCCAG
39. C6G: AGTGAGGTCCAGA
40. C6T: AGTGAGTTCCAGA
41. T7A: GTGAGCACCAGAT
42. T7G: GTGAGCGCCAGAT