Chapter IV

miRNA-proteomics: An Approach to Rapid Detection of Functional Role of microRNAs

This chapter is currently in communication as a research article.
4.1. INTRODUCTION

microRNAs are small regulatory RNA molecules produced by almost all eukaryotic cells, by enzymatic processing of stem loop precursor RNA molecules (Bartel, 2004; Krol et al., 2010). The mature microRNAs, of 19-21nt, bind to partially complementary target regions in messenger RNA molecules, typically in the 3’ untranslated region, in association with proteins of the miRNP complex (Bartel, 2004; Krol et al., 2010). The binding of miRNA to the target mRNA can result in target cleavage, destabilization through deadenylation of the target (Giraldez et al., 2006; Wu et al., 2006) and translational inhibition (Humphreys et al., 2005; Petersen et al., 2006). Barring a few exceptions, miRNA binding is known to down-regulate the expression of protein from the target. Each microRNA can potentially bind to hundreds of targets in the cell (Friedman et al., 2009). Each target transcript can harbor several different microRNA binding sites resulting in a complex miRNA-target interaction network that regulates translation of eukaryotic genes and consequently the proteomic profile of cells.

The differential expression of microRNAs in cancer, their ability to modulate apoptosis and the known interaction of miRNAs with oncogenes and tumor suppressors, taken together, suggest a significant role for microRNAs in cancer and apoptosis (Dews et al., 2006; Hermeking, 2007; Hermeking, 2010; El et al., 2011). Although identification of differentially expressed microRNAs in cancers is supported by rapid high-throughput methodologies (Ferdin et al., 2010), the identification of the targets through which these miRNAs exert their regulatory role is by and large done through experimental validation of individual miRNA-target pairs. Thomas et al. have reviewed methods for miRNA target identification (Thomas et al., 2010). The targets of microRNAs can be predicted by computational methods that look for complementarity to the microRNA sequence and a low free energy of binding that favors miRNA-target interaction (Bartel, 2009). The accessibility of the targeted region within the secondary structure of the messenger RNA and the conservation of the target site are often used to refine the computational predictions and reduce the number of potential false positives. miRNA-target interaction is usually validated through reporter assays wherein the activity of a reporter gene expressed in fusion to the 3’UTR being tested is monitored in cells that
miRNA-proteomics: An Approach to Rapid Detection of Functional Role of microRNAs

ectopically express the microRNA. MicroRNAs can impact stability and translation of the target mRNA eventually resulting in a net reduction in the steady state level of the protein product of the target. Therefore, several groups have tried to identify targets of miRNAs, by microarray based profiling of messenger RNA in cells over-expressing the microRNA (Frankel et al., 2008). Another approach to high-throughput identification of microRNA targets is to immunoprecipitate the miRNA-target complex using antibodies against the protein components of the miRNP machinery (Wen et al., 2011;Beitzinger and Meister, 2011). The co-immunoprecipitated miRNAs and targets can then be detected by hybridization to microarrays or sequencing. Finally computational prediction of targets is used in conjunction with the immunoprecipitation data to identify miRNA-target pairs. This approach, encompassing methods like High-throughput sequencing of RNA isolated by crosslinking immunoprecipitation (HITS-CLIP) (Chi et al., 2009) and Photoactivatable-Ribonucleoside-Enhanced Crosslinking and Immunoprecipitation (PAR-CLIP) (Hafner et al., 2010) is dependent on the quality and stringency of the immunoprecipitation and requires extensive standardization to minimize non-specific detection of targets. Moreover, this approach inherently assumes that the association of the miRNA with the target is sufficient for repression. Most recently, polysome profiling of mRNA species after over-expression of three different miRNA has been used to capture functional mRNA targets by virtue of their reduced occurrence in actively translating poly-ribosomal fractions (Guo et al., 2010). Irrespective of the mode of action of miRNA, through translational repression or active degradation and destabilization of RNAs, the target protein levels are expected to be reduced. Hence, proteomics can also be used to identify the role of miRNA-target interaction. This approach has been used with proteomics methods like stable isotope labeling with amino acids in cell culture (SILAC) (Vinther et al., 2006;Baek et al., 2008;Yang et al., 2010;Yan et al., 2011;Chen et al., 2011), pulsed SILAC (Selbach et al., 2008) and Isobaric tags for relative and absolute quantification (iTRAQ) (Yang et al., 2009) to identify the targets of microRNAs and their effect on the overall proteomic profile of cells.

Figure 4.1 describes the overall schema used for the carrying out SILAC experiments. Here SILAC was used to study the effect of two pro-apoptotic microRNAs, miR-34a
SILAC was chosen because this study was based on cultured cells and since SILAC involves metabolic labeling of the cells. Eukaryotic cells are dependent on external media for essential amino acids, so carbon or nitrogen in amino acid backbone of Arginine or Lysine is replaced with stable, non radioactive isotopes, $^{13}$C$_6$ or $^{15}$N$_4$. The incorporation of heavy isotopes in amino acids results in increase in molecular weight of proteins which is reflected in peptides after trypsinization. This difference in mass is exploited for quantification and identification of peptides and hence proteins. The advantage of the method is that there are no chemical modifications or reactions required after the completion of treatment and hence reduces sample to sample variations (Ong et al., 2002; Ong and Mann, 2006; Ong and Mann, 2007).

**Figure 4.1: Schema of SILAC experiments.** Lysate isolated from cells grown in different media is mixed and analyzed by mass spectrometry, to identify differentially expressed proteins.

miR-34a is a highly conserved pro-apoptotic miRNA that inhibits cell proliferation in normal cells and is induced by p53 in response to DNA damage (Hermeking, 2010). Loss of miR-34a has been shown in several cancers including neuroblastoma (Welch et al., 2007; Cole et al., 2008), pancreatic cancer (Chang et al., 2007) and non-small-cell lung cancer (Bommer et al., 2007). Some of the direct targets of miR-34a that allow it to inhibit cell proliferation include BCL2 (Bommer et al., 2007) and CDK4.
Ectopic expression of miR-34a is also known to induce cellular senescence in human fibroblasts (He et al., 2007) and apoptosis (Tarasov et al., 2007; Raver-Shapira et al., 2007; Chang et al., 2007; Welch et al., 2007). miR-29a is also a pro-apoptotic miRNA with critical roles in differentiation of neuronal cells. The known targets of miR-29a include members of the Bax family of anti-apoptotic proteins (Kole et al., 2011) and p85α (Park et al., 2009). The ability of miR-29a to repress CDC42 and p85α results in activation of p53 and induces apoptosis (Park et al., 2009). miR-29a is down-regulated in cancers, like hepatocellular carcinoma and acute myeloid leukemia (Garzon et al., 2009; Xiong et al., 2010) and during neuronal cell death. miR-29a can target Bace1, the gene that encodes beta-secretase, aberrant expression of which is implicated in Alzheimer’s disease (Hebert et al., 2008). Thus both miR-34a and miR-29a are pro-apoptotic miRNAs that interact closely with the p53 pathway to establish the balance between cell proliferation and cell death. Both miRNAs are involved in differentiation of neuronal cells and their aberrant expression results in cancers. Here miRNA-proteomics has been used to study the effects of over-expression and knock down of these miRNAs in HEK293T cells.

Our results show that both the miRNA have modest effects on protein level expression of their target genes. miRNA-proteomics was also then compared to each of the other methods available for miRNA-target identification and assessed their relative advantages and disadvantages. The consensus target lists from several computational target prediction tools was compared to the results of miRNA-proteomics, there is a significant enrichment of down-regulated targets. However, comparison with HITS-CLIP showed relatively lesser enrichment of down-regulated targets. Targets of both miRNAs showed enrichment of genes in functional classes related to their cellular role, i.e. regulation of apoptosis. However, there was minimal overlap in the affected proteins.

4.2. MATERIALS AND METHODS

4.2.1. Constructs

hsa-miR-34a was expressed from pSilencerTM 4.1-CMV neo vector (Ambion). The construction of clone and design of complementary lock nucleic acid (LNA)-modified oligonucleotides against miRNA has been described in (Ghosh et al., 2008). hsa-miR-
29a was expressed from pEGFP-N3 (Clontech). The construction of clone and design of LNA-modified oligonucleotides is described in (Ahluwalia et al., 2008). We have earlier shown that under these conditions the miRNA is 2.5 fold over-expressed in case of hsa-miR-34a (Ghosh et al., 2008) and 14 fold over-expressed in case of hsa-miR-29a (Ahluwalia et al., 2008).

4.2.2. Cell culture and SILAC

Human embryonic kidney cell line HEK293T (ATCC number CRL-11268) was cultured in high glucose Dulbecco's modified eagle's medium (DMEM, Invitrogen) supplemented with dialyzed fetal bovine serum. For preparing SILAC medium, SILAC™ Protein ID and Quantitation Media Kit (MS10030) from Invitrogen was used. A pool of cells was grown in medium containing $^{13}$C$_6$ Arginine and $^{13}$C$_6$ Lysine (Heavy medium) for 14 days. Complete incorporation of heavy isotope was confirmed using Mass spectrometry. Simultaneously cells were also grown in normal DMEM (Light medium), to maintain similar passage number. In case of hsa-miR-34a, heavy cells were transfected with antisense LNA-modified oligonucleotides (40nM). The light cells were transfected with pSilencer construct expressing hsa-miR-34a. In case of hsa-miR-29a, heavy cells were transfected with pEGFP-N3 construct expressing hsa-miR-29a, whereas light cells were transfected with antisense LNA-modified oligonucleotides (40nM). All transfections were performed in T25 culture flasks using lipofectamine (Invitrogen) as per manufacture’s protocol. The cells were harvested 24 hours after transfection in case of miR-29a and 36 hours post transfection in case of miR-34a.

4.2.3. Mass spectrometry

Harvested cells were then washed thrice with ice cold phosphate buffer saline (PBS). RIPA buffer (50 mM Tris-Cl pH 7.4, 150 mM NaCl, 2mM EDTA, 2% NP40 and 1X Protease Inhibitor) was used to prepare cell lysate. Debris was removed by centrifugation. The protein was quantified using bicinchoninic acid (BCA) method. Hundred microgram of each cell lysate was pooled for miR-34a and miR-29a separately and used for Methanol-Chloroform precipitation.
4.2.3.1. Multi-dimensional protein identification technology (MudPIT)

The protein pellet for miR-34a was dissolved in solution containing 0.2% ProteasMAX (Promega) and 4M Urea in 50 mM Ammonium bicarbonate. The protein was reduced (5mM tris(2-carboxyethyl)phosphine), alkylated (10 mM iodoacetamide) and trypsinized (1 microgram trypsin : 25 microgram protein) (Promega). Formic acid was added to final concentration of 5% of the peptide pool. The mixture was spun to remove the debris. The peptides were loaded on 250-µm i.d. (internal diameter) fused capillary column, packed with 2.5 cm long 5µm SCX (Luna, Phenomenex) and 2.5 cm of 10 µm Jupiter Reverse phase resin (Phenomenex). The peptides were desalted using this set-up. The SCX end of the this column was now attached to 100 µm i.d. fused capillary with a pulled tip and packed with 15 cm of 4 µm Jupiter Reverse phase resin (Phenomenex). This was the connected inline with Eksigent pump (Eksigent Technologies). The buffer solution that were used are: 5% acetonitrile with 0.1% formic acid (buffer A); 80% acetonitrile with 0.1% formic acid (buffer B), and 500 mM ammonium acetate with 5% acetonitrile and 0.1% formic acid (buffer C). For each sample seven steps chromatography was performed. For first step 70 minute gradient was used with 0 to 3 min – 100% buffer A, 3 to 10 min of 0 - 15% buffer B, 10 to 60 min of 15 -45% buffer B and 60 to 70 min of 45 – 100% buffer B and then back to 100% buffer A. For step 2 to 7, 180 minutes of gradient was used where 0 to 3 min – 100% buffer A, 3 to 8 min of salt bump (buffer C), 8 to 20 min of 0 - 15% buffer B, 20 to 150 min of 15 -45% buffer B and 150 to 170 min of 45 – 100% buffer B and then last 10 minutes of 100% buffer A. Six different salt bumps that were used are: 10, 20, 30, 40, 60 and 100%. The eluting peptides were then directly connected to hybrid LTQ linear ion trap-Orbitrap (ThermoFisher). The data was acquired in data dependent mode with a full scan between 400-2000 m/z at 60000 resolution in the Orbitrap. Full scan was succeeded by five ms2 scans for topmost peaks in linear ion trap with normalized collision energy of 35%. The parameters that were used for dynamic exclusion are as follows: repeat count - 2, repeat duration – 30, list size – 50, exclusion duration – 60.
4.2.3.2. Gel based method

The pellet for miR-29a was dissolved in RIPA buffer with 0.1% sodium dodecyl sulfate (SDS), which was loaded on 5 - 12% Bis-Tris gel. The gel was stained with Colloidal commassie blue solution. Sample lane was then cut from rest of the gel and sliced into ten pieces. Individual piece was then cut into smaller approximately one mm$^3$ size and washed in 50% acetonitrile to destain the pieces. The gel pieces were then dehydrated, reduced (100mM Dithiothreitol (DTT)), alkylated (50mM iodoacetamide) and trypsinized (1 microgram trypsin : 25 microgram protein) (Promega). The peptides were extracted by washing with 50% and 100% acetonitrile.

The peptides were then vacuum dried and resuspended in 100mM ammonium bicarbonate with 5% formic acid. The peptides were then loaded in split free nano-LC system (EASY-nLC; Proxeon Biosystems now Thermo Fisher Scientific) autosampler. The buffer solutions used for chromatography are: 2% acetonitrile with 0.1% formic acid (buffer A); 100% acetonitrile with 0.1% formic acid (buffer B). A gradient of 120 minutes was then used each fraction. The gradient was 0 to 108 min of 3 – 35% Buffer B, 108 to 110 min of 35 – 100% buffer B and last ten minutes of buffer B. The eluent peptides were coupled to hybrid LTQ linear ion trap-Orbitrap (ThermoFisher). The full scans were performed between 350 – 2000 m/z at 60000 resolution in Orbitrap. Each full scan was then followed by six data dependent ms2 scans in ion trap with normalized collision energy of 35%. The parameters that were used for dynamic exclusion are as follows: repeat count - 1, repeat duration – 20, list size – 500, exclusion duration – 40.

4.2.4. Western blotting

30 µg of protein for VDAC1 and 100 µg for VDAC2 was resolved on 15% SDS polyacrylamide gel electrophoresis (PAGE) and electro blotted onto a nitrocellulose membrane using TE 77 semidry transfer unit, Amersham Biosciences at 135mA for 2 hours. Antibodies for VDAC 1 (ab15895) (1:500) and VDAC 2 (ab47104) (1:500) were from abcam. The antibody for GAPDH (sc-32233) (1:1000) was from sigma. The western blots were imaged on the Odyssey Imager (Licor Biosciences).
4.2.5. Mass spectrometry data analysis

The acquired raw files were converted into ms1 and ms2 files using Rawextract 1.9.5. The data was searched with the ProLuCID algorithm (Xu et al., 2006) for human International Protein Index (IPI) database, concatenated with decoy database. Search parameters includes: Static modifications: carbamidomethylation at Cysteine (57.02146 amu) and Heavy modification at Lysine and Arginine (6.0204 amu) for heavy search; dynamic modifications: oxidation of methionine (15.9949 amu) and deamidation of Aspergine and Glutamine (0.984 amu) for both light and heavy search. DTASelect was used to organize and filter the search output from ProLuCID. Peptides with tryptic ends were filtered; false discovery rate at protein level was kept below 1% and overall mass accuracy of 5ppm (parts per million) was maintained. Proteins with at least two peptide hits were retained. And the data was used as input for quantitative analysis tool named Census (Park et al., 2008). The data was corrected if the mode of ratios was not zero. Arginine to proline conversion was corrected in Census. The quantified proteins were used for further analysis.

4.2.6. miRNA – target prediction

The prediction data for hsa-miR-29a and hsa-miR-34a was retrieved from http://mirecords.biolead.org/ (Xiao et al., 2009). In total there are ~31,000 and 33000 predicted targets for hsa-miR-29a and hsa-miR-34a respectively. Filtering criteria was used where at least four other prediction programs and TargetScan should predict a gene as target to derive a list of consensus list of targets for miRNAs. The RefSeq ids were converted to IPI ids using an online server (http://biit.cs.ut.ee/gprofiler/gconvert.cgi) (Reimand et al., 2007), to get a list of predicted list of potential targets.

4.2.7. miRNA – target interaction (HITS-CLIP)

HITS-CLIP data was retrieved for mmu-miR-34a in mm9 from http://ago.rockefeller.edu/tag_mm9.php. Of 736 total interactors, top 236 with cluster height >= 8 were used for further analysis. Mouse homologs were mapped to human using http://biit.cs.ut.ee/gprofiler/gorth.cgi which were then eventually converted to IPI ids using http://biit.cs.ut.ee/gprofiler/gconvert.cgi.
4.3. RESULTS

4.3.1. SILAC data generation

In order to study the effect of miR-34a and miR-29a, conditions for ectopic over-expression and knock down of both miRNAs were standardized. A pre-miRNA sequence along with flanking regions was cloned in two different vectors as mentioned in methods, and transiently transfected it into HEK293T cells. We have earlier shown that under these conditions the miRNA is 2.5 fold over-expressed in case of hsa-miR-34a (Ghosh et al., 2008) and 14 fold over-expressed in case of hsa-miR-29a (Ahluwalia et al., 2008). Locked nucleic acid modified antisense oligonucleotides were used to knock down the expression of both miRNAs (Figure 4.2).

![Figure 4.2: Experimental schema for miRNA-proteomics.](image)

Cells were transfected with either miRNA expressing vector or LNA modified oligonucleotide against miRNA. The lysate mix was analyzed on LTQ-ORBITRAP mass spectrometer. The data was analyzed as mentioned in materials and methods. The quantification data was then compared with other miRNA-target identification methods.
HEK293T cells have detectable levels of both miR-34a and miR-29a. The antisense based knock down of the miRNAs helped in clearing the endogenous miRNA and further enhancing the difference in expression level of the miRNAs between cells transfected with the over-expression construct and anti-sense oligonucleotides.

### 4.3.2. SILAC data analysis

SILAC based detection of the proteins was carried out in the transfected cells with either the over-expression construct or the LNA against the miRNA. The raw data files were processed as mentioned in materials and methods and summarized in figure 4.3.

![Flowchart of steps carried out to analyze the mass spectrometry data](image-url)

**Figure 4.3:** A flowchart of steps carried out to analyze the mass spectrometry data. The detailed steps and parameters used for data analysis are summarized in materials and methods.
2022 proteins were successfully detected, in the miR-34a experiment at 1% FDR at protein level and at least two peptides per protein. These accounted for 1952 non-redundant proteins. The entire list of proteins identified and quantified in our study can be found at (https://sites.google.com/a/rnabiology.org/beena-pillai/resources). In the miR-34a experiment, fold change in expression level of the protein was computed as the ratio of the peptide in miR-34a over-expression sample (labeled with the heavy isotope) to the corresponding peak in the sample wherein miR-34a was knocked down using LNA (light). On comparison of cells over-expressing miR-34a to cells where the endogenous miR-34a was cleared by anti-sense treatment, expression of 13% of the proteins was found to be changed by more than 50%. The bulk of these amounting to 9% of the total detected proteins were up-regulated while only 4% was down-regulated more than 1.5 fold (Figure 4.4A, C and table 4.1).

Figure 4.4: Proteomics quantification data. Log2 fold change distribution of the identified proteins in miR-34a (A) and miR-29a (B) SILAC experiments. The pie charts represent proportion of proteins differentially expressed with a fold change cut off of 1.5 in case of miR-34a (C) and miR-29a (D).
Table 4.1: Number of proteins differentially expressed in hsa-miR-34a proteomics experiment, with fold change cut off of 1.5 and 2

<table>
<thead>
<tr>
<th>Fold change</th>
<th>0.5 &gt; FC &gt; 2</th>
<th>0.667 &gt; FC &gt; 1.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>2022</td>
<td>2022</td>
</tr>
<tr>
<td>Unchanged</td>
<td>1908</td>
<td>1751</td>
</tr>
<tr>
<td>Up-regulated</td>
<td>93</td>
<td>181</td>
</tr>
<tr>
<td>Down-regulated</td>
<td>21</td>
<td>90</td>
</tr>
</tbody>
</table>

Although 3305 proteins were detected in total in the miR-29a experiment, only 1639 were non-redundant hits using similar quality control as miR-34a. The over-expression of miR-29a resulted in comparable number of proteins being up-regulated (14% of total) and down-regulated (17%) by 1.5 fold. Overall, miR-29a resulted in a larger impact on the proteome, with 31% of the detected proteins being affected (Figure 4.4B, D and table 4.2). An illustrative example of a differentially expressed protein detected in the study is shown in figure 4.5A. The peak corresponding to the peptide VYVGNLGTGAGK, mapped to the protein SFRS7, a serine/arginine (SR)-rich family of pre-mRNA splicing factors, implicated in alternative splicing and showed a significant up-regulation.

Table 4.2: Number of proteins differentially expressed in hsa-miR-29a proteomics experiment, with fold change cut off of 1.5 and 2

<table>
<thead>
<tr>
<th>Fold change</th>
<th>0.5 &gt; FC &gt; 2</th>
<th>0.667 &gt; FC &gt; 1.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>3315</td>
<td>3315</td>
</tr>
<tr>
<td>Unchanged</td>
<td>2931</td>
<td>2266</td>
</tr>
<tr>
<td>Up-regulated</td>
<td>226</td>
<td>448</td>
</tr>
<tr>
<td>Down-regulated</td>
<td>126</td>
<td>569</td>
</tr>
</tbody>
</table>
Figure 4.5: Differential expression of proteins and examples of validated miRNA-targets. A) Identification and quantification of SFRS7, one of the up-regulated proteins in miR-29a study is depicted as an example with software used at different steps. NOTCH2 and CD276 have been shown to be direct target of miR-34a and miR-29a respectively. B) Quantification based on ms1 data displayed in graphical format in census software for a peptide of NOTCH2 (B) and CD276 (C).

4.3.3. Validation and comparison with known targets

The differentially expressed proteins included the ion channel protein VDAC2. VDACs are a family of pore-forming, voltage-dependent, anion-selective channel proteins mainly located in the mitochondrial outer membrane. In mammals three isoforms of VDAC exist (De, V et al., 2010). The VDAC (voltage-dependent anion channel) plays a central role in apoptosis, participating in the release of apoptogenic factors including cytochrome c (Rostovtseva et al., 2005; Zalk et al., 2005). Both VDAC1 and VDAC2 were analysed to check the expression level of protein with over-expression and repression of miR-29a. Cells were transfected with LNA modified antisense oligonucleotide against miR-29a and miR-29a over-expressing construct. The LNA transfected sample showed 30% up-regulation of VDAC1 as compared to miRNA over-expressing cells. Whereas VDAC2 showed almost two fold up-regulation under similar condition (Figure 4.6).
Figure 4.6: Immunoblotting for Vdac1 and Vdac2. HEK293T cells were transfected with mock LNA, LNA modified anti-miR-29a, Vector or plasmid over-expressing miR-29a. Immunoblotting was performed as mentioned in materials and methods. A) Immunoblotting for Vdac1; B) Immunoblotting for Vdac2.

Known targets of the miRNA in the down-regulated set of proteins were explored. Notch1, Notch2 and CDK6 are reported to be targeted by miR-34a in Glioma and stem cells (Li et al., 2009c). Notch2 was found to be down-regulated, by 66% (Figure 4.5B and Table 4.3). Similarly, miR-29a is known to target and down-regulate CD276 in normal tissue samples (Xu et al., 2009). In agreement with this, CD276 was found to be down-regulated by about 80% in cells expressing miR-29a in our proteomics study (Figure 4.5C and Table 4.3).

Table 4.3: Protein quantification of previously known targets

<table>
<thead>
<tr>
<th>miRNA</th>
<th>IPI id</th>
<th>Gene description</th>
<th>Log₂ (Fold Change)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>hsa-miR-34a</td>
<td>IPI00297655.4</td>
<td>NOTCH2, Neurogenic locus notch homolog protein 2 precursor</td>
<td>-0.7485</td>
<td>(Li et al., 2009c)</td>
</tr>
<tr>
<td>hsa-miR-34a</td>
<td>IPI00219604.3</td>
<td>MAP2K1, Dual specificity mitogen-activated protein kinase kinase 1</td>
<td>0.1844</td>
<td>(Ichimura et al., 2010)</td>
</tr>
<tr>
<td>hsa-miR-29a</td>
<td>IPI00016786.1</td>
<td>CDC42, Isoform 2 of Cell division control protein 42 homolog</td>
<td>-0.304</td>
<td>(Park et al., 2009)</td>
</tr>
<tr>
<td>hsa-miR-29a</td>
<td>IPI00793688.1</td>
<td>CD276, 60 kDa protein</td>
<td>-0.8625</td>
<td>(Xu et al., 2009)</td>
</tr>
</tbody>
</table>
Both miR-34a and miR-29a are relatively well characterized with 21 and 11 reported targets found by Reporter-UTR assays (Xiao et al., 2009). However, there was very little overlap between these targets and the list of differentially expressed proteins identified from our study. For instance, out of the 11 reported targets of miR-29a, only 2 were found to be differentially expressed in the proteomics study.

4.3.4. Comparison of miR-34a SILAC with HITS-CLIP

Another experimental high-throughput approach for finding miRNA-targets is HITS-CLIP where miRNP complexes are pulled down and the enriched UTRs and miRNAs in the miRNP complex are identified by sequencing. After identifying the pool of miRNA and targets present in miRNPs, computational algorithms are used to predict potential binding between miRNA and target regions. There exists, currently, no HITS-CLIP data from human or mouse for miR-29a, while miR-34a targets have been identified using this method in mouse. 736 human homologs were found for the mouse targets from the HITS-CLIP data (Table 4.4). In order to filter data for high stringency, the gene list was sorted according to number of tags and selected all genes detected by at least 8 tags. The 236 genes thus identified were then mapped to its human homologs. The human homologs of these genes included 32 targets identified in our proteomics study of which 19 were down-regulated (Appendix IV).

Table 4.4: Predicted target proteins in HITS-CLIP study for hsa-miR-34a

<table>
<thead>
<tr>
<th>miRNA</th>
<th>hsa-miR-34a</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of HITS-CLIP targets (mouse homologs)</td>
<td>736</td>
</tr>
<tr>
<td>Common proteins (Highest cluster height)</td>
<td>32</td>
</tr>
<tr>
<td>FC &gt; 0</td>
<td>10</td>
</tr>
<tr>
<td>FC &lt; 0</td>
<td>19</td>
</tr>
</tbody>
</table>

4.3.5. Comparison of SILAC with computationally predicted targets

The lists of computationally predicted targets of both miRNAs to their respective lists of differentially expressed proteins from our proteomics study were compared. Computational prediction of targets against miRNAs typically leads to hundreds of
predicted targets and is widely held to be susceptible to false positive prediction. Consensus prediction of targets, considering only targets commonly predicted through at least five prediction programs including TargetScan was carried out to generate a list of targets for miR-34a and miR-29a and finally checked for the presence of these targets in our list of differentially expressed proteins identified by proteomics. The consensus target prediction approach led to 213 targets for miR-34 and 350 targets for miR-29a respectively (Table 4.5). The number of detected proteins in proteomics experiments is typically much lower than the total number of potential proteins in the genome and a computational prediction of targets should lead to many more targets thus proteins that were detected in our experiment were compared to the list of computationally predicted targets. Only 7 of the 213 predicted targets of miR-34a were detected in our experiments. Of the 7 predicted targets, however, six showed down-regulation (log2FC<0; Appendix V). The proteomics study for miR-29a identified 26 out of 350 predicted targets. 22 of these 26 predicted targets were found to be down-regulated (log2FC<0; Appendix VI).

Table 4.5: Comparison of computationally predicted targets to proteins found in miR-34a and miR-29a proteomics study

<table>
<thead>
<tr>
<th>miRNA</th>
<th>hsa-miR-34a</th>
<th>hsa-miR-29a</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of predicted targets (5 programs including TargetScan)</td>
<td>213</td>
<td>350</td>
</tr>
<tr>
<td>Number of targets predicted and detected by proteomics</td>
<td>7</td>
<td>26</td>
</tr>
<tr>
<td>FC &gt; 0</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>FC &lt; 0</td>
<td>6</td>
<td>22</td>
</tr>
</tbody>
</table>

In this study 100s of differentially expressed targets including 33(7+26) direct targets were found. The direct targets are largely down-regulated while the net effect of the over-expression of the microRNA in both cases is a larger number of up-regulated proteins than down-regulated proteins, implying that secondary effects account for a significant part of the differential expression. However, the changes in the proteome can provide valuable insights about the state of the cell following miRNA over-
expression and shed light on the overall effect of the miRNA. Functional classification of the differentially expressed proteins was carried out to understand the cellular role of these miRNAs as revealed through proteomics. Gene Ontology classification of miR-34a down-regulated targets revealed the occurrence of the GO term, “generation of energy and metabolites” and related terms to be over-represented (miR-34a-Appendix VII; miR-29a- Appendix VIII). Further, terms related to developmental process were also significantly frequent in agreement with the known role of miR-34a in development. Although this miRNA has been implicated in cell-cycle, the differentially expressed proteins did not contain well-known cell-cycle regulators. GO terms linked to mitochondrial activity and calcium signaling were over-represented in the target lists of both miRNAs highlighting their previously known role in apoptosis.

4.4. DISCUSSION

Currently, presence of hundreds of microRNAs in each eukaryotic organism is predicted or established through high-throughput experiments (Griffiths-Jones et al., 2006). Each of the microRNAs can theoretically bind to hundreds of targets, as predicted by algorithms that rely on the partial complementarity between 3’UTR and microRNA sequences (Friedman et al., 2009). More refined algorithms try to add further accuracy to the predictions by imposing penalties for lack of conservation and occurrence in highly structured regions that may mask the access to the site. Another approach to trim the false predicted targets has been to consider only consensus targets predicted by multiple programs (Hariharan et al., 2005). In spite of these approaches, the numbers of predicted targets remain well above the numbers that can be validated by reporter assays leaving the discovery of targets to miRNAs a tedious and slow process. Here the relevance of miRNA-proteomics in identifying direct targets and indirect effects of miRNA over-expression in cultured cells, by using two microRNAs, miR-29a and miR-34a was explored.

miR-34a is known to be induced by p53 in response to DNA damage and target anti-apoptotic proteins, thus mediating a pro-apoptotic effect. It has been also implicated
in regulation of the Notch-DLL pathway during fly development and regulation of SIRT1 in human cells (Yamakuchi et al., 2008). It is implicated in cancers due to its down-regulation in hepatocellular carcinoma and other cancers, suggesting a normal pro-apoptotic role that mediates cellular senescence during development and malignancy. miR-29a, on the other hand, is known to induce p53 activation by targeting CDC42, and also exert anti-apoptotic effects during brain development by keeping the expression of anti-apoptotic proteins in check.

Recently, two other groups (Chen et al., 2011; Kaller et al., 2011) have explored the effect of miR-34 over-expression on the proteomic profile of cells. Since the cell lines and methodology used is different in each study, a comparison cannot be used to quantitatively compare their experiments with the data collected here. Overlapping and differentially expressed proteins in all studies were checked qualitatively (Figure 4.7). Although each study had identified between 1000 and 2000 proteins, about 50% of the proteins were common with the other studies. In spite of hundreds of proteins detected in both experiments, the number of differentially expressed proteins was limited. MTHFD2, CHN1, RRM2 and ACADVL were down-regulated in our study and at least one of the other studies.

Figure 4.7: Comparative analysis of proteome profiles following miR-34 expression from three studies. Differentially expressed proteins from our study were compared to Chen et. al., 2011 and Kaller et. al. 2011. Numbers in the circles depict number of proteins detected in each experiment.
The comparison of targets from consensus bioinformatics predictions, HITS-CLIP, proteomics and reporter assays has not been done for any miRNA. The number of targets predicted following HITS-CLIP was found to be larger than the number of consensus target predictions. A total of 59% (19/32) of the HITS-CLIP targets queried by proteomics showed any level of down-regulation. These numbers suggest that HITS-CLIP may provide data on co-localization and enrichment of the miRNA and targets in miRNPs but the miRNA-target mapping maybe an over-estimate due to the use of computational target identification during HITS-CLIP data analysis. Thus co-localization of miRNA and target may not result in direct binding and further, binding may not be sufficient for down-regulation of the protein. In comparing computational predictions and proteomics data the majority of predicted targets, if detected, were down-regulated, suggesting that the consensus prediction of miRNA targets is indeed effective. Compared to computational predictions and HITS-CLIP, the major limitation of proteomics remains the low number of detected proteins and the inability to reliably detect modest effects. Since many miRNAs show only 30-60% reduction of target expression in Reporter-UTR assays, prevalent thresholds (like 2 fold) in proteomics studies are irrelevant to miRNA-proteomics studies.

A few other studies have previously tried to use proteomics to identify the effects of miRNAs. Vinther et. al first studied the effect of miR-1 (Vinther et al., 2006) in HeLa cells using SILAC technology. Subsequently, SILAC was applied to capture the effects of miR-124, miR-1 and miR-181 in HeLa cells and miR-223 in knock-out mice neutrophils (Baek et al., 2008). Pulsed SILAC was used to study the effect of miR-1, miR-155, miR-16, miR-30a and let-7b in HeLa cells (Selbach et al., 2008). The effect of miR-21 and miR-143 was also studied by miRNA-proteomics using iTRAQ (Yang et al., 2009) and SILAC (Yang et al., 2010) respectively. Selected reaction monitoring (SRM), a mass spectrometry based method has been used to validate computationally predicted targets in C. elegans (Jovanovic et al., 2010). Each of the previous reports independently concluded that the miRNAs showed modest effects both in number of differentially expressed proteins and fold changes in target expression. In this study two relatively well-characterized miRNAs have been used and we find that the vast majority of their effects are indeed modest in magnitude.
Since the current technology allows the detection of only a few thousand proteins of which 100s of proteins were differentially expressed, it is inferred that these two microRNAs affect the expression of a large number of target genes. A substantial fraction of these targets were potentially direct targets with predicted binding sites for the microRNAs.

Besides searching for direct targets of microRNAs, microRNA-proteomics provides an opportunity to compare the cellular effect of microRNAs in terms of the differentially expressed proteins. miR-29a/b and miR-34a/b, are both known to influence apoptosis and closely linked to p53 (Park et al., 2009; Hermeking, 2010). However, the evidence collected from different groups working on diverse systems is contrary to each other. On the one hand, several groups have established that the normal role of miR-34 is to trigger apoptosis, in response to transcriptional activation by p53 (Raver-Shapira et al., 2007). In keeping with this view, expression of miR-34 is reduced in several cancers including hepatocellular carcinoma (Li et al., 2009b) and ovarian cancer (Corney et al., 2010). According to independent reports, c-myc, a transcriptional repressor of miR-29a is also a target of miR-34a (Christoffersen et al., 2010). On the other hand, miR-29 has been implicated in cancers both as an oncomiR and tumor suppressor (Garzon et al., 2009; Han et al., 2010). Its reported ability to target DNMT1 and 3, resulting in global hypo-methylation and up-regulation of tumor suppressor genes supports a pro-apoptotic role in lung cancer cells (Fabbri et al., 2007) while its reduced expression in several myeloid cancers and its ability to cause AML in a mouse model suggests an anti-apoptotic role (Garzon et al., 2009; Han et al., 2010). A reduced expression of miR-29a is also associated with cell death in neuronal cells (Kole et al., 2011). Further, miR-29a can induce the activity of p53 by targeting CDC42 and p85alpha (Park et al., 2009). Here the effect of both microRNAs was compared in the same cellular context, by comparing the proteome of cells with over-expressed microRNA and reduced endogenous microRNA. Although only about 2000 proteins were detected in each study 961 proteins were detected in both studies, allowing a comparison of the proteomic changes triggered by over-expression of miR-34a and miR-29a. Majority of the proteins (51%; 493 out of 961), as expected, did not show any changes in either study. However, 5% of the proteins were up-regulated...
in both studies and 4% of them were down regulated in both studies. Only 1% of proteins showed anti-correlated changes, suggesting that both microRNAs have largely overlapping roles. Thus, miRNA-proteomics is proposed to be of limited utility in identifying direct targets, given the current limitations in number of proteins that are detected routinely but can be used for identifying the overlapping or distinct roles of miRNAs.