SRI VENKATESWARA INTERNSHIP PROGRAM FOR RESEARCH IN ACADEMICS (SRI-VIPRA)

SRI-VIPRA

Project Report of 2023: SVP- 2308

"Identification and Analysis of MicroRNA-137 Targets in Diseases"

IQAC
Sri Venkateswara College
University of Delhi
Benito Juarez Road, Dhaula Kuan
New Delhi -110021

SRI-VIPRA PROJECT 2023

Title : Identification and Analysis of MicroRNA-137 Targets in Diseases

Name of Mentor : Dr. Ravindra Varma **Name of Department :** Department of Biochemistry **Designation** : Assistant Professor	

List of students under the SRIVIPRA Project

S. No	Roll Number	Name of the student	Course	Photo	Signature
1	1221022	Lavanya Saini	B.Sc. (Hons.) Biochemistry		
2	1321013	Lavanya Tyagi	B.Sc. (Hons.) Biological Science		
3	1321008	Megha Chandilla	B.Sc. (Hons.) Biological Science		
4	1321026	Shivani Yadav	B.Sc. (Hons.) Biological Science		

Signature of Mentor

Certificate of Originality

This is to certify that the aforementioned students from Sri Venkateswara College have participated in the **SRI-VIPRA summer internship program** SVP-2308 titled **"Identification and analysis of microRNA-137 targets in diseases"**. The participants have carried out the research project work under my guidance and supervision from **15 June, 2023** to **15th September 2023.** The work carried out is original and carried out in a hybrid (offline and online) mode.

P. Ravindra Varma

Signature of Mentor

Acknowledgement

We would like to express our sincere thanks to the Principal of Sri Venkateswara College, University of Delhi (SVC, DU) for organizing the SRI VIPRA Program -2023. This program helped us to put our scientific understanding to application.

We would like to express our gratitude and thanks to our professor and mentor, Dr. Ravindra Varma for guiding us in this project. He provided us with his valuable suggestions throughout this process. He encouraged us to undertake this project from time to time.

We would also like to thank the coordinators of SRI VIPRA Program, 2023 for providing us the opportunity to be a part of this SRI VIPRA PROJECT - 2023.

We also would like to thank our family and friends for supporting us throughout this project.

We are a team of four students Lavanya Saini, Lavanya Tyagi, Megha Chandilla and Shivani Yadav who have worked together with the support and encouragement from the aforementioned people towards the successful completion of this project. Our consistent efforts combined with our dedication and teamwork helped us to achieve this.

TABLE OF CONTENTS

List of Tables and Figures

A.) FIGURES

B.) TABLES

S.No	Tile of Table
Table 1	Different microRNAs involved in different diseases
Table 2	List of mirna target prediction tools and used strategies
Table 3	List of ML prediction tools with kinds of strategies used
Table 4	List of Verified Targets derived from Literature Mining
Table 5	List of Verified Targets derived from miRTarBase Database
Table 6	List of disease wise target genes
Table 7	STRING analysis for KEGG PATHWAY
Table 8	STRING analysis for REACTOME PATHWAY
Table 9	STRING analysis for MOLECULAR FUNCTIONS
Table 10	STRING analysis for DISEASES
Table 11	STRING analysis for BIOLOGICAL PROCESSES

ABSTRACT

MicroRNAs are small RNA molecules associated with gene regulation through mRNA-silencing as a result of which they further affect multiple biological processes such as cell differentiation, programmed cell death etc. Studying microRNA biology can help us understand gene control, cell physiology and pathological processes.

MicroRNAs are known to have numerous disease-associated targets, but not many are experimentally verified. Our study focuses on microRNA-137 and investigates identification of multiple targets verified by different assays to create a microRNA target database. A total of 298 targets were obtained from published literature search as well as from mirTarBase database. The number of experimentally verified targets whose assays were known equated to be **255**. Out of these 255 targets, 134 were found from mirTarBase while **121** targets were found through research articles during literature mining. These targets are classified based on their association with diseases such as various cancers and other clinical diseases. We also identified information on the biological and molecular role of these targets. Bioinformatics analysis was performed for these 255 targets using STRING software in order to identify the statistically significant molecular functions, biological processes, KEGG and Reactome pathways. Many of these targets were associated with regulation of cancer gene network pathways suggesting their potential for further investigation as clinical targets for cancer diagnosis and therapy.

Integration studies related to microRNA targets identified in various diseases is extremely important for enhancing our knowledge on microRNA regulatory mechanisms. MicroRNA target verification and validation studies can help us in developing diagnostic as well as therapeutic assays useful for clinical management of various diseases.

1.0 INTRODUCTION

1.1 MicroRNA Biology

1.1.1 Introduction: For a long period of time, RNA was thought to contribute to gene expression simply by passing on the genetic information stored in the form of DNA to corresponding proteins. Later in the 1960s, a subset of RNAs was observed which controlled gene expression by turning certain genes on or off. These are now known as the MicroRNAs.

MicroRNAs form an extremely conserved group of small RNA molecules, composed of approximately 19-25 nucleotides. They are non-coding RNAs i. e. they are not seen to code for a protein. MicroRNAs are found to be extremely crucial to cell function. They show their involvement in multiple biological processes with damaging effects such as autoimmune diseases, viral infections and oncogenesis. They have sequence stretches complementary to corresponding mRNAs and act to decrease their expression. This is similar to the functioning of short interfering RNAs (siRNA) in mRNA splicing. The difference lies in the complexity of microRNA function as compared to the other 'on-off' model. [3, 4]

1.1.2 Biosynthesis: miRNAs precursor genes called mir-genes are found in the nucleus, and when transcribed, lead to the formation of miRNAs. The miRNA transcripts formed (primary miRNAs) then undergo the processes of splicing and capping. Their eventual processing to form pre-miRNA is preceded by formation of hairpin-like stem loop structure. A microprocessor complex involving an RNA III endonuclease- Drosha and a cofactor Pasha (DiGeorge syndrome critical region 8) perform the essential processing. Thus, this complex forms the pre-miRNAs. These are characterized by being 60 to 70 nucleotides long and having a 5′ phosphate and a 3′ nucleotide overhang. Exportin 5 (Ran transport receptor family) transport the pre-miRNA to cytoplasm where it is processed to shorter double stranded miRNA with the use of Dicer (RNase III endonuclease) and trans-activator RNA binding protein. The enzyme helicase unwinds the duplex to form mature miRNA which is single stranded. The mature miRNA is then incorporated within an RNA induced silencing complex and guided to the target mRNA. miRNAs are also produced by non-canonical processes which are Drosha independent.[4]

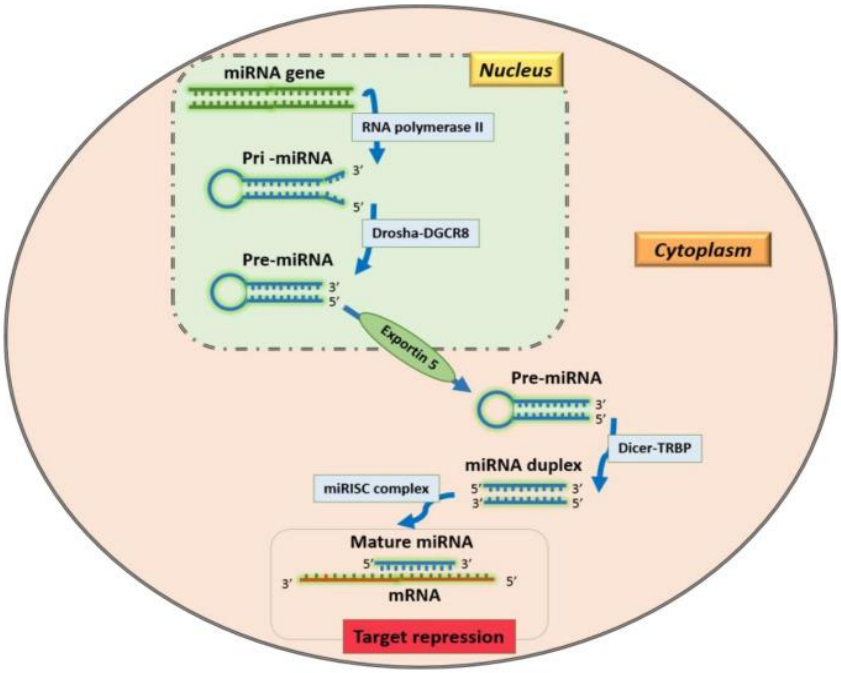


FIGURE 1: Canonical Pathway for microRNA production [60]

1.1.3 Functioning: As stated earlier, miRNAs can repress the corresponding target genes. Their mechanisms of action are diverse. MicroRNA might bind with the mRNA in the 2 to 8 nucleotide long seed region or base pair in the ~9-12 nucleotide long central region, consequently cleaving the mRNA and degrading it. It has been observed that microRNA-target interactions having higher degree of complementarity cause mRNA degradation whereas microRNA-target interactions with comparatively lesser complementarity result in translational repression. microRNAs generally work by decreasing expression of their target genes rather than completely silencing them. There are multiple studies of this mechanism of action, such as RNA degradation, inducing decapping or deadenylation etc. Even though the major pathway of miRNA action follows decreasing target expression by miR binding in the 3'UTR region , certain reports suggest this suppressive feature can be overcome under certain given conditions. In fact, microRNA and promoter binding can increase target expression levels.[2]

1.1.4 microRNA targeting: MicroRNAs are seen to regulate hundreds of mRNA targets and in turn can be regulated by numerous miRNAs. They regulated target expression by complementary base pairing to sequence motifs in the 3′ UTR of respective mRNAs. This base pairing might be perfect or near perfect. The complementary between sequences of miRNAs and their targets occurs at the 5'end of microRNA, called as the seed starting from the 5' end of miRNA, the seed incorporates nucleotides 2-8. There are multiple targets that don't possess this canonical seed binding site, for instance, the targets having G:U wobble pairing within the seed region. The target gene determines the length of the specific 3'UTR binding region. The length can be changed by alternative cleavage as well as PAS (Polyadenylation Signals). These signals are mostly composed

of a six-nucleotide sequence- 5′AAUAAA, however some single base variations might occur. A large number of these sequences follow the termination codon in a gene, allowing alternative usage. Alternative polyadenylation has the ability to change the length of the 3'UTR target region and it has also been studied that miRNA binding sites might be lost or gained by the same process of alternative splicing. Similarly, alternative splicing can cause changes in the miR binding sites due to their ability to alter the coding and untranslated regions of target mRNA.[2]

1.1.5 MiRNA Circulation: microRNAs are seen to be secreted into exosomes and possess the ability to communicate with other cells/tissues. When secreted, they might be taken up by cells from nearby tissue, or can be transported to distant sites in cases when vesicles reach circulation. It has been widely studied that can be released into extracellular fluid, which can, in fact, be used as markers for multiple diseases. These extracellular miRNAs can reach specific targets and act as regulators to alter cellular activities. Thus, it can be said that they have hormone-like actions.[2]

1.1.6 Regulation: As seen in other RNAs, miRNAs are also regulated by mechanisms like transcriptional activation, transcriptional inhibition, controlling degradation etc. More than half, i.e. ~52% of human miRNAs are found in the intergenic regions, about 40% are seen in intronic regions and the rest 8% lie in the exonic region of the genes. The host gene regulates the intronic miRNAs, which are usually processed from introns. Initiation of transcription occurs by upstream signaling and feedback loops might be formed due to miR genes targeting their own transcription factors. Precursor miRNAs with a 3' overhang is formed due to processing by Drosha/DGCR8 in the nucleus, and further processing is carried out by Dicer/TRBP in the cytoplasm which recognizes this 3′ overhang and performs further cleavage to give rise to the mature microRNA form. The alter processing differences as well as loop and miR secondary structure affect these both steps.[4]

Multiple studies and evidence exist which show that single nucleotide polymorphisms (SNPs) affect and thus regulate the functioning and biosynthesis of miRNA. SNPs arise due to alterations in the DNA sequences of either miR coding genes or binding sites. These polymorphisms tend to change the binding of miRNAs to target, like in terms of strength, and thereby strengthen or weaken the regulatory effects. Such changes are associated with abnormalities, such as giving rise to oncogenes.[1]

RNA editing is yet another method of miR regulation. It is a modification which occurs at a specific site which leads to a modulated product. About 16% of human pri-miRNAs undergo A-to-I editing. This process results in variations of target mRNA and a control layer to the regulatory process.[1]

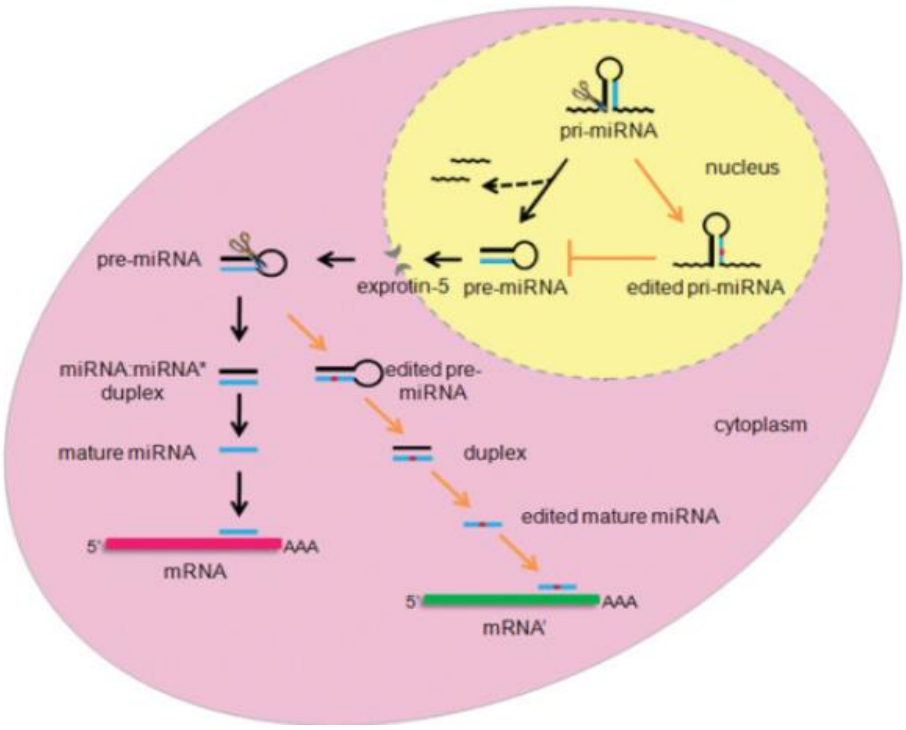


FIGURE 2 : Pathway of microRNA Editing [1]

1.2 MicroRNA in Diseases

MicroRNAs play various major roles in processes like cell proliferation, differentiation, apoptosis, metabolism, neuronal cell fate, neuronal gene expression, brain morphogenesis, muscle differentiation, stem cell division etc. Many miRNA expressions are disease specific this comes from the evidence that the gene expression profile of a non-neuron cell becomes more like that of a neuron when the neuron specific **miR-124** was artificially over expressed within. So, they are involved in various diseases like various miRNAs have been reported to be either upregulated or down regulated in different cancers. The upregulated miRNAs expression may function as oncogenes and under expressed expression may act like tumor suppressor genes; by this mechanism it regulates hallmarks of cancer such as angiogenesis, apoptosis and cell proliferation. **miR-1, miR-133** and **miR-208** are found to be overexpressed in heart and are chief regulators of heart development as well as myocyte differentiation whereas **miR-1** and **miR-133** were reported to be responsible for human heart failure. Many miRNAs are associated with the degeneration of skeletal muscles too like muscular dystrophies. Even Facioscapulohumeral muscular dystrophy (FSHD) could be distinguished from Duchenne muscular dystrophy based on the level of miRNAs-381 and miRNAs-382 expressions in FSHD patients. **miR-126** inhibits the expression of vascular cell adhesion molecule 1 (VCAM-1) which is important to intervene leukocyte adherence to endothelial cells. They are associated with neurodevelopmental disorders like intellectual disability, autism, attention deficit hyperactivity disorder (ADHD), epilepsy etc too. The loss of **miR-29 cluster** has been shown to be associated with increased (beta-amyloid precursor protein-

converting enzyme) BACE1/beta-secretase expression in sporadic Alzheimer's Disease patients. Likewise, miR-**298**, **miR-328** have been identified to directly interact with the 30-UTR of the BACE1 transcript that is involved in beta-amyloid production. The **miR-20a family (miR 20a, miR-17-5p and miR-106b)** has been reported to regulate the expression of Amyloid Precursor Protein (APP) thus providing probable associations for these microRNAs during Alzheimer's Disease development. In neuropsychiatric disorder of Schizophrenia, role of Brain-Derived Neurotrophic Factor (BDNF) is very much evident with GABA and glutamatergic neurons and bioinformatics findings have demonstrated that hsa-mir-1 and hsa-mir-206 have their target of the 30-UTR of BDNF. With this it can be said that miRNA could regulate BDNF protein synthesis by interfering with BDNF mRNA translation during brain development. Similarly in bone related diseases, like in osteoporosis, more than 50 miRNAs have been identified which target the bone remodeling process including **miR-21, miR-23-3p, miR-100, miR-124, miR-125b and miR-133.** [5,6,7,8]

The given below is a table representing some of the miRNAs associated with various diseases:

TABLE 1 : Different MicroRNAs involved in different diseases

DISEASES	miRNA
Breast cancer	miRNA-21, miR-155, miR-23, and miRNA-191, miR-205, miRNA- 145, miR-10b, and miR-125b
Ovary cancer	miR-199a, miR-140, miR-145, and miRNA-125bl , miRNA-200a, miR-200c, and miR-141
Endometrioid adenocarcinoma	miR-193a, 193b, miR-205, miR155 miRNA 200a, 200b, 200c
Colon cancer	miR-34, miR-let7, miR 143, miR 145, miRNA-133b, and miRNA-126, miR-let 7g, miR-21, miR-20a, miR-17- 19 family, miR 31, miR 135, miRNA-181b, and miRNA- 200c
AML	miR-155, hsa- miR- 191, 199a
CML	miR-17-5p, miR-173p, miR- 18a, miR-19a, miRNA-19b-1, miR-20a and miR- 92a-1
CLL	miR-21, miR 150, miR-155, miR- 15a, miR16, miR-29, miR143, miR-45, miR-30d, miRNA- let 7a, miR-181a
Oesophagus cancer	miRNA- 203, miR-194, miR- 192, miR-200c
Gastrointestinal cancer	miR-106b- 25, miR-15b, miR- 16
Lung cancer	has-mir-126*, miRNA-let 7, hsa-let-7a-2, let-7f-1, has-mir-21 and has-mir- 205, miRNA- 17-92

Bladder cancer	miR-2 23, miR- 26b, miRNA- 221, miR- 103-1, miR-185, miR-23 b, miR- 203, miRNA-17-5p, miRNA-23, miR- 205, miR-29c, miR-26a, miR-30c, miR- 30e-5p
Thyroid tumors -	
PC	miRNA- 146 b, miR- 221, miRNA- 222, miR- 181b, miR- 155, miR- 224
AC	miR-30d, miRNA-125b, miR- 26a, miR- 30a- 5p
Cardiac hypertrophy	miR-23a, miR-23b, miR-24, miRNA-195, miR-199a, and miR-214
Down syndrome	miR-99a, let-7c, miR-125b-2, miR-155 and miR-802
Alzheimer	miRNA-9, miR-128a, miR-125b
Rheumatic arthritis	miR-155, miRNA-146
Systemic lupus erythematosus	miRNA-189, miR-61, miR-78, miR-21, miR-142-3p, miRNA- 342, miR-299-3p, miR-198 and miR-298, miR-196a, miR-17-5p, miR-409-3p, miRNA-141, miR-383, miR- 112, and miR-184
Psoriasis	miRNA-203
Gastric cancer	miRNA-145
Liver cancer	miRNA-29b
VIRAL DISEASES	
HCV	miRNA-122, miR-155
HIV-1	miR-28, miR-125b, miR-150, miR-223, miR-382
Influenza virus	miR-21, miR-223
IMMUNE RELATED DISEASES	
Multiple sclerosis	miRNA-145, miR-34a, miR-155, miR-326
Systemic lupus erythematosus	miR-146a
Type II diabetes	miRNA-107, miRNA-144, miR-146a, miR-150, miRNA-182, miR-103

Nonalcoholic fatty liver disease	miRNA-200a, miRNA-200b, miR-429, miRNA-122, miR-451, miR-27
Non-alcoholic steatohepatitis	miRNA-29c, miR-34a, miR-155, miR-200b
Parkinson's disease	miRNA-30b, miR-30c, miR-26a, miR-133b, miR-184*, let-7
Alzheimer's disease	miR-29b-1, miRNA-29a, miR-9

1.3 MicroRNAs in Cancer

Cancer is a disease where some cells of the body (affected cells) lose the ability of contact inhibition and grow uncontrollably. Normal cells of the body respond to the signals that control normal cell behavior whereas cancer cells do not respond and grow uncontrollably. Under normal circumstances there is balance between cell proliferation and cell death and cellular homeostasis is maintained in cells in our body. But in the case of cancer, the cells continue to grow. As they grow, they may invade new tissues (malignant), thereby starting a new tumor there. Cancer can be malignant or benign depending on whether they can start a new tumor at a new site or not. Cancer is a complex disorder where not only genetic mutations but also the environment plays an important role. Cancer occurs in two major steps: initiation and progression. Initiation occurs due to genetic alteration (which can be genetic or environmental affect), later more mutations occur leading to tumor progression. Furthermore, cancer is very heterogeneous in nature and varies substantially in their behavior and response to treatment. Cancer can occur in most of the parts of the body. Lung cancer is the most common cancer in males and Breast Cancer is the most common cancer in females. Apart from this colon, liver, kidney, pancreatic, leukemia, glioblastoma lymphoma cancers are increasing these days and have become major health problems.

Several agents are known to cause cancer: radiation, chemical carcinogens like tobacco, diesel exhaust, smoking etc. In addition to these, certain viruses like Human papilloma virus (HPV) can also cause cancer (cervical cancer). In cancer, alterations in normal signaling pathways occur due to mutations and/or epigenetic changes. In 2000, Douglas Hanahan and Robert Weinberg published a paper describing six hallmarks of cancer and now over the years after two decades the number has increased to 14 which include-sustaining proliferative signaling, evading growth suppressors, resisting cell death, Hypoxia, enabling replicative immortality, inducing angiogenesis, Deregulating Cellular Metabolism, Tumor Promoting Inflammation, activating invasion and metastasis, Genome Instability & Mutation, Nonmutional Epigenetic Reprogramming and Polymorphic Microbiomes (1-3) [9]. Several scientists throughout the world

are busy understanding the underlying mechanisms responsible for the alterations hoping that this in turn will help them in providing meaningful targets for cancer therapeutics.

The most common genes that are found in cancer cells include p53, c-myc, Bcl-2 family members. In cancers there is downregulation of pro-apoptotic (Bax etc) and upregulation of anti-apoptotic (Bcl-2). In recent years, there have also been several studies that apart from dysregulated gene expression, dysregulated expression of miRNA(s) is also found in many human diseases including cancers [10]. It is important to note that miRNAs can act as either oncomirs or tumor suppressors depending on the genes they act on [11]. Genetic and epigenetic changes in miRNA-encoding genes affect their expression. For instance, alterations in genomic miRNA copy numbers and gene locations such as amplification of the gene, gene deletion or translocation affect the miRNA expression levels.

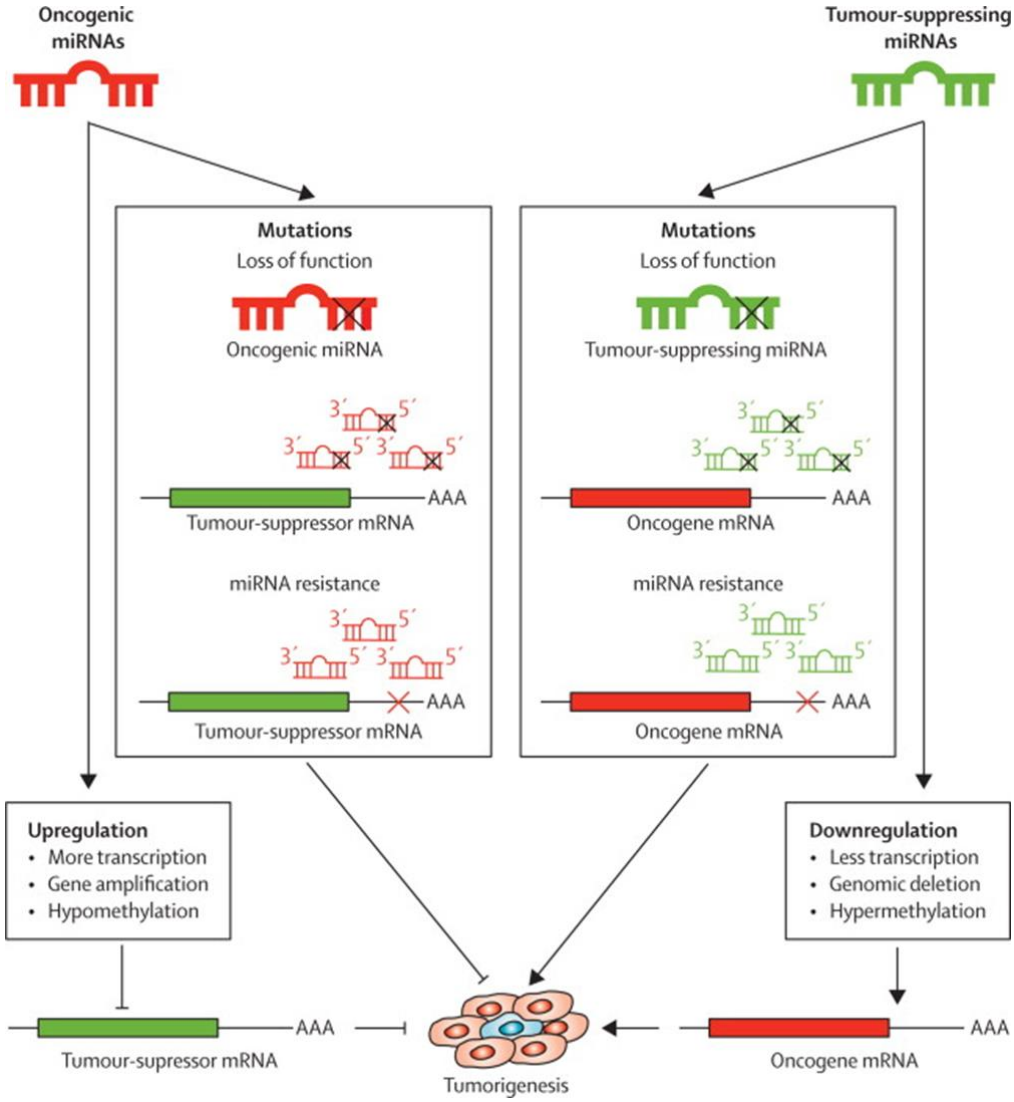


FIGURE 3 : Regulation of tumorigenesis by miRNAs : Tumorigenesis can be regulated by miRNAs at different levels. Upregulation of oncogenic miRNAs reduces expression of tumor-

suppressor protein, but downregulation of tumor-suppressing miRNAs results in an increased production of oncogenic protein. Loss-of-function mutations in tumor-suppressing miRNAs and mutation of the target section of oncogene mRNA can cause tumorigenesis, because expression of oncogenic proteins is no longer regulated. Loss-of-function mutations in oncogenic miRNAs and mutations in tumor-suppressor mRNA would increase expression of tumor-suppressor proteins and hence reduce tumorigenesis. miRNA-microRNA. [12]

1.4 Verification Assays for the miRNA Targets

MicroRNAs are post-transcriptional controllers that control mRNA stability and the restatement effectiveness of their target genes. Mature microRNAs are roughly 22- nucleotides in length. They intervene post-transcriptional gene regulation by binding to the amiss reciprocal sequences (a.k.a. microRNA nonsupervisory rudiments, MRE) in the target mRNAs. It's estimated that further than one- third of the protein- rendering genes in the human genome are regulated by microRNAs. The experimental styles to examine the interactiveness between the microRNA and its targeting point (s) in the mRNA are important for understanding microRNA functions.

1.4.1 LUCIFERASE ASSAY

The luciferase reporter assay is commonly used as a tool to study gene expression at the transcriptional level. It is widely used because it is convenient, relatively inexpensive, and gives quantitative measurements instantaneously.

Luciferases make up a class of oxidative enzymes found in several species that enable the organisms that express them to “bioluminesce,” or emit light. The most famous one of these enzymes is the firefly luciferase. Fireflies are able to emit light via a chemical reaction in which luciferin is converted to oxyluciferin by the luciferase enzyme. Some of the energy released by this reaction is in the form of light.

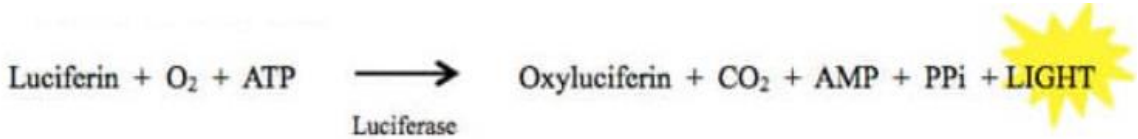


FIGURE 4 : Reaction for Luciferase Assay

This reaction is highly energetically efficient, meaning nearly all the energy put into the reaction is rapidly converted to light. This makes it extremely sensitive, which is great for a reporter assay. A luciferase assay is used to determine if a protein can activate or repress the expression of a target gene. The assay measures the light output from luciferase enzyme that is expressed under the control of the promoter. Luciferase enzymes are well suited for reporter assays because they are highly sensitive and can quantify even small changes in expression.

The aim of the assay is to demonstrate an interaction between the miRNA(s) of interest and the 3 ’ UTR of a particular gene. [54]

1.4.2 POLYMERASE CHAIN REACTION (qRT- PCR)

miRNAs play the part of post-transcriptional regulation by binding to target mRNAs; hence the target sequences were screened among mRNA deduced cDNAs in hybrid PCR. An oligo dT- 3 spots appendage manual was introduced into the 5'-end of mRNA-deduced cDNA during reverse transcription. This manual distinguished the mRNA-deduced cDNAs effectively from other DNAs or RNAs in modification. miRNA specific hybrid-primer was designed according to the miRNA sequence. The reverse and complementary sequence of the seed region of miRNA was located at the 3’end of the hybrid primer. Hybrid- PCR was projected as semi-nested PCR using the hybrid-primer and the external/ inner manuals homologous to the oligo dT- 3 spots appendage manual. In particular, the target mRNA of a given miRNA was determined by hybridization of the mongrel-manual to the sequence of mRNA- deduced cDNA. A low annealing temperature of 37 ℃ was applied in the first-round modification, so as to make mongrel- manual hybridize with apparent target sequences in a condition analogous to core body temperature. Also, an alternate round PCR with advanced annealing temperature of 55 ℃ was followed for further specific modification of sequences from apparent target mRNAs. Extension was long enough to avoid deficient modification. The products of modification were variable in length. [53]

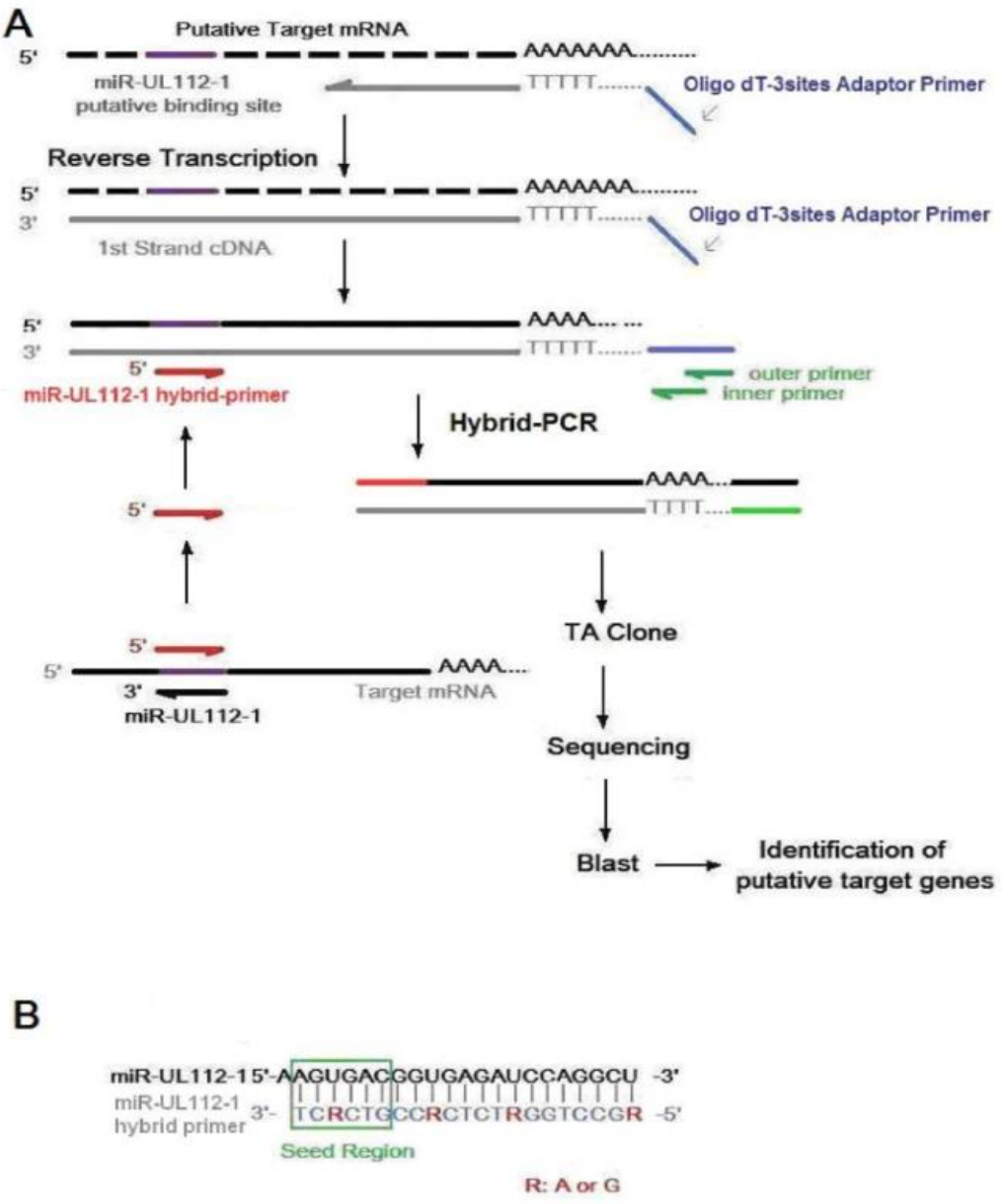


FIGURE 5 : Protocol of hybrid-PCR. (A) Schematic donation of principle and process designed for hybridl- PCR.(B) Diagram showing sequences of miR- UL112- 1 and miR- UL112- 1 mongrel manual. Positions marked by Red R meant arbitrary insertions of A or G. Seed region was indicated by the green box around nucleotide 2- 7 of miR- UL112- 1.

1.4.3 CLIP SEQUENCE

To further ameliorate the effectiveness of capture of miRNA targets, the Tuschl group proposed a modified protocol, photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation (PAR- CLIP), in which photoactivatable ribonucleoside analogues similar as 4-thiouridine(4-SU) or 6-thioguanosine (6- SG) are incorporated into RNAs before crosslinking.

These modified nucleotides can be efficiently crosslinked to proteins using UV A (365 nm). In addition, crosslinking- individual mutations (T-to-C or G-to-A,independently) are introduced during reverse transcription to allow determination of the list spots at close- to- nucleotide resolution. This protocol has been successfully used to identify not only miRNA targets, but also the RNA targets of numerous RNA-binding proteins. To achieve the asked single- nucleotide resolution in the identification of RBP targets, a system that takes advantage of the propensity of rear transcriptase to stop at the position of crosslinking has been proposed. This individual nucleotide resolution CLIP system (iCLIP) has only veritably lately been applied to the characterization of small RNA-guided relations.[57]

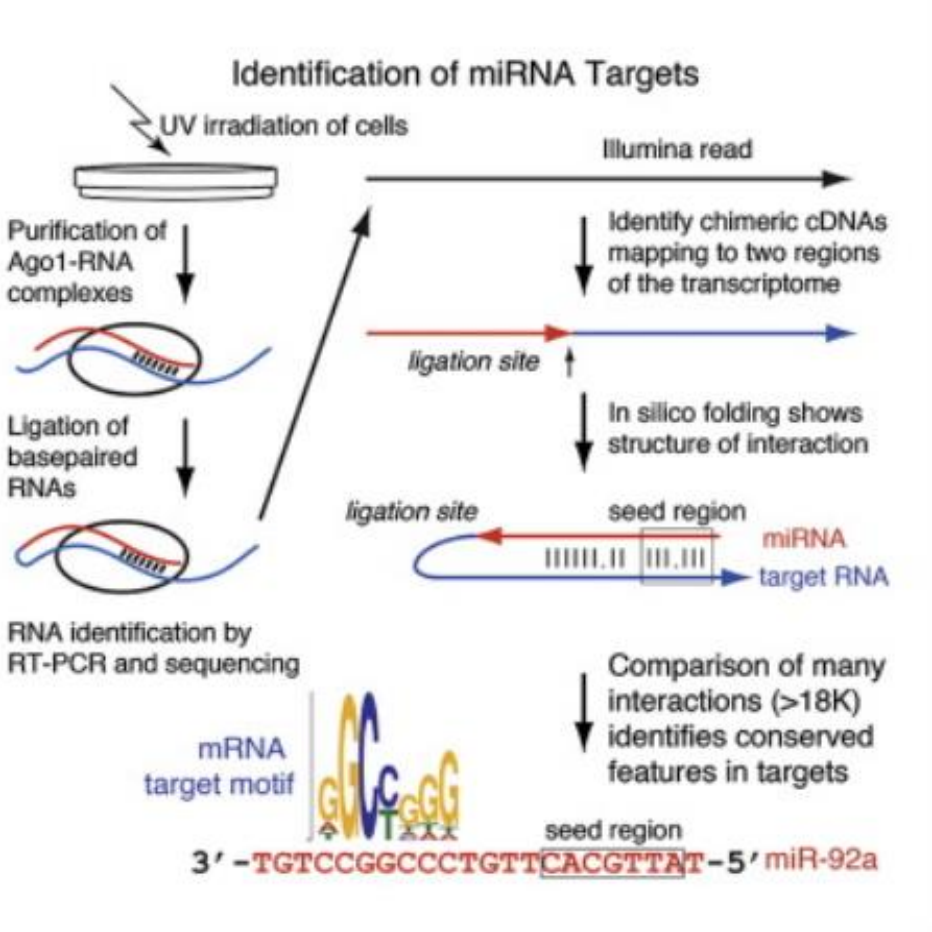


FIGURE 6 : Identification of miRNA target by clip sequence

1.4.4 WESTERN BLOT ANALYSIS

Western blotting is an important technique used in cell and molecular biology. By using a western blot, researchers are able to identify specific proteins from a complex mixture of proteins extracted from cells. The technique uses three elements to accomplish this task:
(1) separation by size,
(2) transfer to a solid support, and
(3) marking target protein using a proper primary and secondary antibody to visualize.

Western blot is often used in research to separate and identify proteins. In this technique a mixture of proteins is separated based on molecular weight, and thus by type, through gel electrophoresis. These results are then transferred to a membrane producing a band for each protein. The membrane is then incubated with label antibodies specific to the protein of interest. The unbound antibody is washed off leaving only the bound antibody to the protein of interest. The bound antibodies are then detected by developing the film. As the antibodies only bind to the protein of interest, only one band should be visible. The thickness of the band corresponds to the amount of protein present; thus, doing a standard can indicate the amount of protein present.

Western blot uses two different types of agarose gel: stacking and separating gel. The higher, stacking gel is slightly acidic (pH 6.8) and has a lower acrylamide concentration making a porous gel, which separates protein poorly but allows them to form thin, sharply defined bands. The lower gel, called the separating, or resolving gel, is basic (pH 8.8), and has a higher polyacrylamide content, making the gel's pores narrower. Protein is thus separated by their size more so in this gel, as the smaller proteins travel more easily, and hence rapidly, than larger proteins.

It is very important to be aware that the data produced with a western blot is typically considered to be semi-quantitative. This is because it provides a relative comparison of protein levels, but not an absolute measure of quantity. There are two reasons for this; first, there are variations in loading and transfer rates between the samples in separate lanes which are different on separate blots. These differences will need to be standardized before a more precise comparison can be made. Second, the signal generated by detection is not linear across the concentration range of samples. [58].

1.5 MicroRNA Prediction Tools

The regulatory network controlled by the microRNAs can be defined by learning about the miRNA and its target mRNA interaction. The first step to achieve this result is usually computational prediction, after which validation of this miRNA-mRNA interaction can be done experimentally. Multiple tools exist which facilitate computational analysis, and each of these tools uses a different methodology to predict the microRNA target. As mentioned earlier, a given miRNA can possess multiple potential target sites. It is not feasible to experimentally validate all of the considered targets in the laboratory due to expense and time constraints. However, the desired results can be achieved by applying a computational approach to predict microRNA potential targets, thus the

total number of target sites which need to be experimentally validated get reduced. miRNA target prediction tools can also aid in easier discovery of new interactions.

Predictive Approaches- There are multiple methods for developing the target prediction algorithms. They can fall into two major categories- miRNA and mRNA interaction-based algorithm derivation and Machine learning based statistical inference.

Features of de-novo Prediction:

- Seed pairing - The binding of target mRNA is based on the miRNA seed sequence. Even the special characteristics in close proximity of this seed region affect the process of gene repression, and as a result all these characters are usually included in the target prediction tools.Watson and Crick pairing is of immense importance in most of these algorithms.
- Thermodynamic stability -Target prediction tools assess the thermodynamic properties by analyzing the free energy of the predicted interaction. This evaluates the stability of miRNA-mRNA complexes.
- Evolutionary conservation -The principle followed here is that given interactions with advantageous biological functions tend to be selected during evolution.
- Accessibility of target site - The miRNA site and 3′-UTR of the considered target should be accessible. A few known algorithms rank the predicted targets on the basis of score relating to accessibility of binding site.
- Number of target sites in 3′-UTR - Prediction algorithms which consider multi-targeting massively reduce the number of sites which need to be experimentally verified by reducing the predicted sites number.

Computational tools using these strategies-

1. **miRanda-** it is among the first and frequently used algorithms. The complementarity between miRNAs and 3'-UTR regions, thermodynamic stability of the duplex structure, evolutionary conservation of the entire binding site, and its position within 3′-UTR as a final filter forms the basis for evaluation.
2. **TargetScan-** this tool predicts the target by searching for sites like 8mer,7mer,6mer which match the seed region of mirna. TargetScan then arrange the predicted targets on the basis of efficiency and cumulative feature scores. Beside this it also evaluates free energy by predicting the secondary structure.
3. **RNAhybrid-** the basis of evaluation of this tool is minimum free energy of hybridization for a long and short RNA, so the short sequence is hybridized to the best fitting part of the long sequence. The user can put restrictions according to its own concern.
4. **PITA-** it has different approaches as its main feature is target site accessibility. The free energy gained from miRNA–mRNA pair formation is evaluated and then mirna has access to the energy cost of making targets. The difference between these two parameters is computed. The sites around the seeds known as 'flank sites' are also considered for accessibility.

TABLE 2 : List of miRNA target prediction tools and used strategies [59]

SF: Sequence Features; TS: Thermo-Stability; EC: Evolutionary Conservation; SA: Site Accessibility**.**

Tool Name	SF	TS	EC	SA
miRanda	✓	✓	✓	
TargetScan	✓	✓	✓	✓
Rnahybrid		✓		
PITA		✓		✓

Feature of Machine learning-

- ML styles generally enable computers to dissect collected data in order to make data-driven models, discover statistically significant patterns and connections, and accordingly make prognostications on new data. In practice, ML algorithms are suitable to "learn" from sets of data and use the acquired knowledge to dissect analogous data and make prognostications grounded on the exemplifications that have been handed.
- In the case of the miRNA target identification, ML approaches don't relate to miRNA-mRNA features, similar to sequence or thermodynamic stability, but try to fete implicit miRNA targets by pertaining to miRNA – mRNA relations with proven natural significance. [59]

The general strategy for ML can be summarized as follows:

- For every miRNA, identify the presumed list point from validated mRNA targets (as positive) andnon-targets (as negative).
- Excerpt features from these relations (anyhow of whether they're functional or inoperative). Train a classifier to distinguish targets fromnon-targets.
- Test the classifier
- Use the classifier to sort unknown miRNA- mRNA relations as positive(target) or negative(non-target).

TABLE 3 : List of ML prediction tools with the kinds of used strategies. [59]

Tool Name	Algorithm	Positive	Negative	Features
MBSTar	Random Forest	MiRbase	Randomly generated	sequence, structure
NbmiRTar	Naïve Bayes	TarBase	Probability Randomization	sequence
TargetBoost	Genetic Programming	let-7, lin-4, miR-13a, bantam	Random string	sequence
DeepTarget	RNN	miRecords miRBase	Mocking in alignment	sequence
TargetMiner	SVM	miRecords	Randomly generated	seed
DeepMirTar	Autoencoder	miRecords	Mocking in alignment	sequence, structure, energy and other
DIANA-microT-CDS	microT-CDS algorithm	miRNA regulatory element in both the 3'-UTR and CDS		sequence, structure, energy and other
miRanda-mirSVR	SVR (similar to SVM)	set from transfection experiments		sequence, structure, energy and other

1.6 miRNA Target Identification

The most direct way to identify miRNA targets on a large scale is to quantify miRNA- induced changes in protein accumulation; many proteomics platforms have been developed for the qualitative and quantitative characterization of protein mixtures and post translational modifications.

1.6.1 PROTEOMIC APPROACHES:

1.6.1.1 Stable Isotope Labeling by Amino acids in Cell Culture (SILAC) – It is one of the proteomic approaches which is used to identify miRNA targets and investigate the consequences on protein level. It is a quantitative method based on whole proteome metabolic labeling with the use of stable isotope- labeled amino acids in cell culture. The principle behind this is the introduction of a mass difference between two proteomes which results in two versions of every peptide which can be easily distinguished in mass spectrometry-based analysis. The classic version of it was based on growing cells in media with natural light or heavy amino acids for several days or for approximately five cell doublings which allows the virtually complete incorporation of heavy amino acids into the cellular proteomes. It is used

in many cell systems like SW480 colon cancer, MiaPaCa-2 pancreatic cancer cells etc. The modern version of SILAC is adopted nowadays which is known as pulsed SILAC (pSILAC) to investigate miRNA proteomic effects. It includes growing two populations of cells for instance miRNA- treated and control cells which are pulse labeled with the two different heavy SILAC amino acids, they both are grown initially in standard media which contains normal light amino acids. After miRNA transfection, culture medium from control and miRNA transfected cells is replaced with SILAC medium containing either heavy amino acids or medium heavy amino acids. After the pulse labeling step, all the newly synthesized proteins will incorporate medium heavy or heavy amino acids while proteins that were synthesized before the labeling step will contain only the light amino acids and thus not considered for protein quantification.

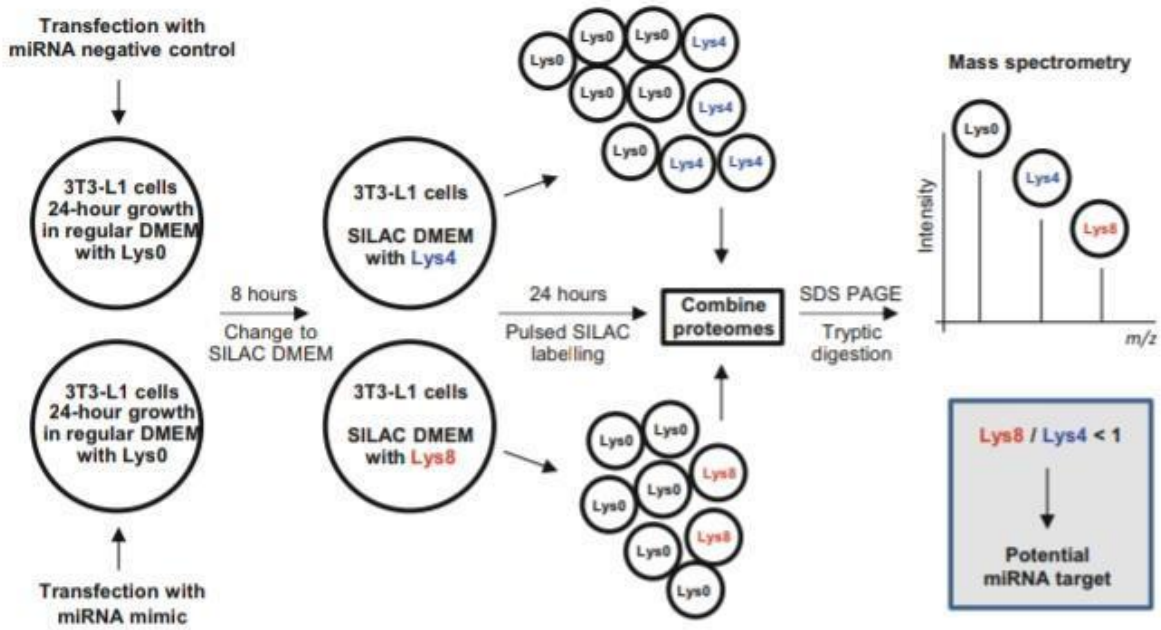


FIGURE 7 : Figure showing pSILAC experimental design. Cells are firstly cultivated in regular DMEM media containing light lysine (Lys 0) in two cell culture dishes for the time period of 24 h and then transfected with miRNA mimic of the interest or negative control. After 8 h, normal growth medium from both the cell populations is replaced with SILAC media which contains either heavy lysine [13C615N2 - l -lysine (Lys8); red font] and medium-heavy lysine [2H4- l -lysine (Lys4); blue font] for the miRNA mimic cells and miRNA negative control cells, respectively. After 24 h, protein extracts from both cell populations are combined and then subjected to SDS-PAGE separation, which is digested with trypsin and then submitted to high performance liquid chromatography tandem mass spectrometry (HPLC/MS/MS). The ratio of heavy and medium-heavy peptides reflects the differences in the translation of the corresponding protein induced by miRNA mimic whereas light peptides are ignored. A ratio of heavy (Lys8)/medium-heavy (lys4) peptides smaller than 1 indicates the potential miRNA target.

1.6.1.2 Two-dimensional differentiation in gel electrophoresis (2D-DIGE):

It is another approach which is applied for miRNA target identification, it includes electrophoresis on a single gel of two samples which are labeled with different fluorescent dyes, separating the proteins by iso-electric focusing and SDS PAGE which is followed by identification by mass spectroscopy. It was used to investigate miRNA- 21 targets in the MCF7 cells which are treated with an anti miR inhibitor of miRNA-21.

1.6.1.3 Translation profiling:

The technique of **polysome profiling** in which cycloheximide is used to trap the elongating ribosomes. After this, centrifugation through a sucrose gradient then separates mRNAs with no associated ribosomes from those which are bound to ribosomes and are presumably undergoing translation. The polysome profile of a mRNA provides information on the two key parameters of translation which are the fraction of mRNA species bound by at least one ribosome and the average number of the ribosomes bound per 100 bases of coding sequence. Poly(A)+ RNAs from bound as well as unbound pools are isolated, amplified, coupled to Cy5 and Cy3 dyes, respectively, and then competitively hybridized to DNA microarrays. Polysome profiling of HEK-293 T cells with and without over-expression of miR-124 was used to determine the relative contribution of translational repression and mRNA degradation in mediating the miRNA activity. [13,14,15,16]

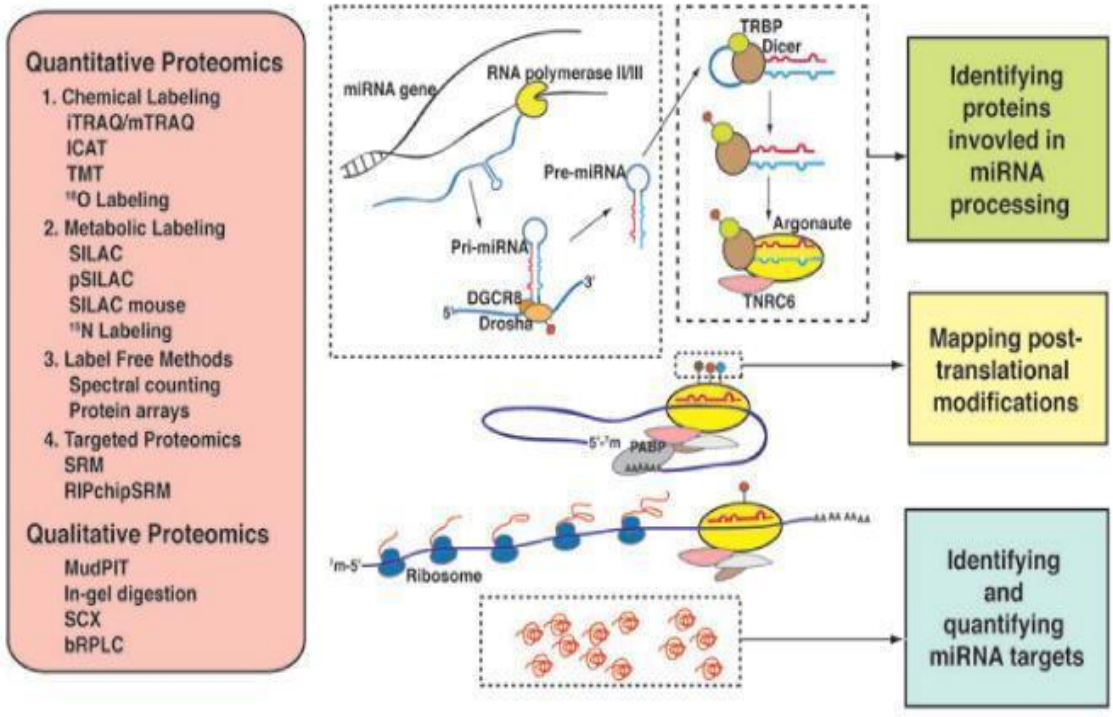
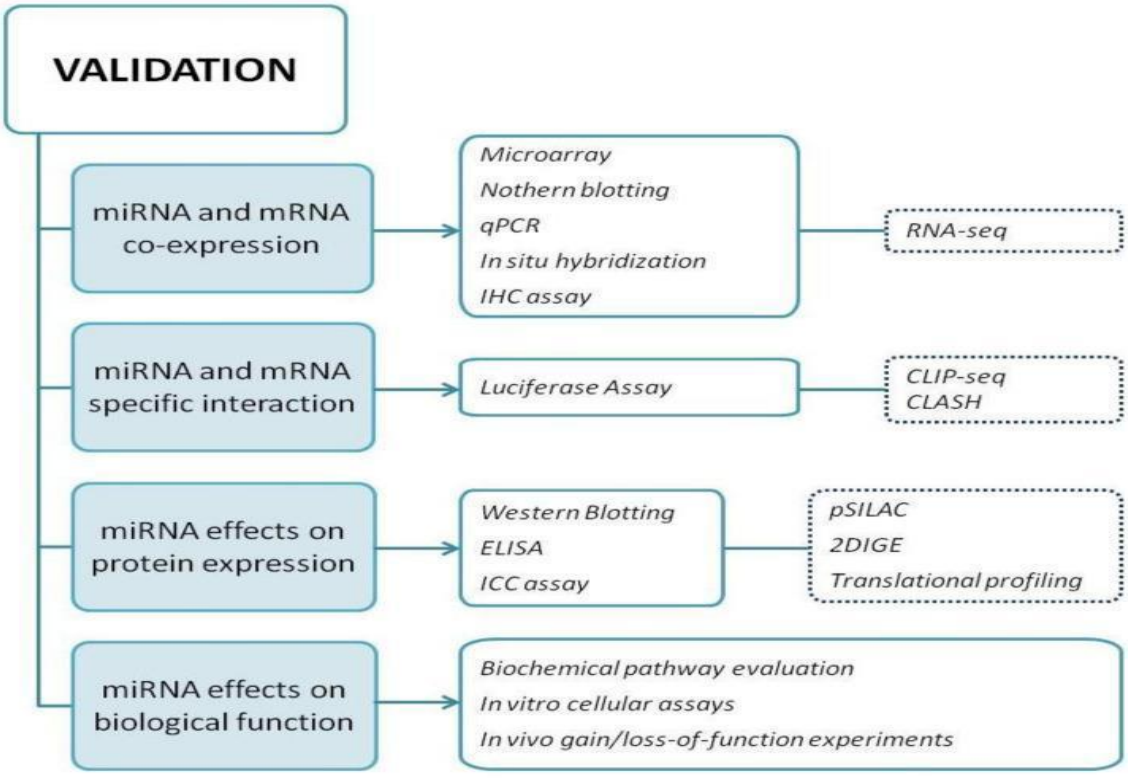


FIGURE 8 : showing proteomic strategies for studying the mechanism of miRNA processing as well as miRNA-mediated translation repression. Genes which are encoding miRNAs are

transcribed as pri-miRNAs with the help of RNA polymerase II and processed by the microprocessor complex consisting of Drosha and DGCR8. This complex first cleaves pri-miRNAs into approx 60–80 nt hairpin structures, which are called pre-miRNAs, that are transported out of the nucleus by exportin-5. In the cytoplasm, Dicer associates with TRBP and results in excision of the loop portion of pre-miRNAs generating double stranded mature miRNAs. Which are then loaded onto RNA interference silencing complex (RISC) which consists of Argonaute and TNRC6. Translation inhibition is done by base-pairing of guide strand miRNA to the 3′UTR of mRNA. Quantitative proteomics analysis includes different categories of strategies like chemical labeling, metabolic labeling, label-free methods and targeted proteomics. From them, iTRAQ and SILAC are the most frequently used. All these methods can enable not only characterization of the miRNA processing pathway components and their posttranslational modifications but also the discovery of miRNA targets.

In this flowchart, all experimental procedures that need to be performed in order to meet the four criteria required for miRNA validation are indicated. Dotted rectangles are for the methods belonging to High-Throughput technologies



-

FIGURE 9 :Experimental procedures that are performed in order to meet the four criteria required for miRNA validation.

1.6.2 RNA Sequencing :

RNA sequencing employs next-generation sequencing, also known as NGS. It reveals RNA in a sample and simultaneously quantifies RNA in a biological sample. Early RNA-sequencing techniques employed Sanger sequencing technology (innovative at the time yet low-throughput as well as costly). It is better than the older used methods since it helps to sequence DNA on a much larger scale at reduced costs [21].

An RNA-seq workflow involves several steps, they are:

1. RNA extraction; RNA sample is taken for analysis
2. Reverse transcription into Cdna : A Cdna is synthesized from a RNA by reverse transcription.
3. Adapted ligation: Cdna is then fragmented. Adaptors are added to each end of the fragment.
4. Amplification and Sequencing: Following processes of amplification, size selection, clean-up and quality checking, the cDNA library is analyzed using NGS, producing short sequences These sequences correspond either to all or a part of the fragment from which it had been derived.

On the basis of what type of RNA is being sequenced, they are [20] -

1. whole transcriptome RNA - sequencing develops sequences for all types of RNA present in a sample. It profiles the whole transcriptome and provides all required information about the gene expression of a cell and the nucleotide of the cell.
2. mRNAs-Seq

Only the mRNA of a cell is sequenced in this sequencing method. The mRNA is first isolated by poly- A chromatography (or magnetic beads) and formed into a poly- A library. Then, the library is directly sequenced or indirectly sequenced to get the mRNA sequence.

3. Single cell RNA-Seq

Single cell RNA sequencing involves sequencing the RNA of a single line/type. The entire transcriptome of the single cell is sequenced, a transcript library is created, and the entire library of RNA is sequenced.

4. SmallRNASeq

Small RNA-Seq is a sequencing method that involves sequencing small non-Coding RNAs from a cell. The most common small RNAs sequenced in this method are miRNA (miRNA), siRNA, and piRNA

5. tRNA-Seq is a method of isolating and sequencing tRNAs. Similarly, rRNA is a method of sequencing rRNAs. Both of these methods are rarely used.
6. Targeted RNA-sequencing is a method of sequencing a particular transcript of interest.

For the identification of miRNA targets, a high-throughput RNA sequencing (HTS) study is conducted to experimentally detect genes that are targeted and downregulated by the 25 microRNAs [19]. The HTS dataset is then combined with publicly available miRNAs target binding data for a systematic identification of miRNA targeting characteristics that are common to both miRNAs that bind to them and those that downregulate them. By incorporating these common characteristics into a machine learning model, an enhanced computational model is created for genome wide miRNA target predictions and is validated.

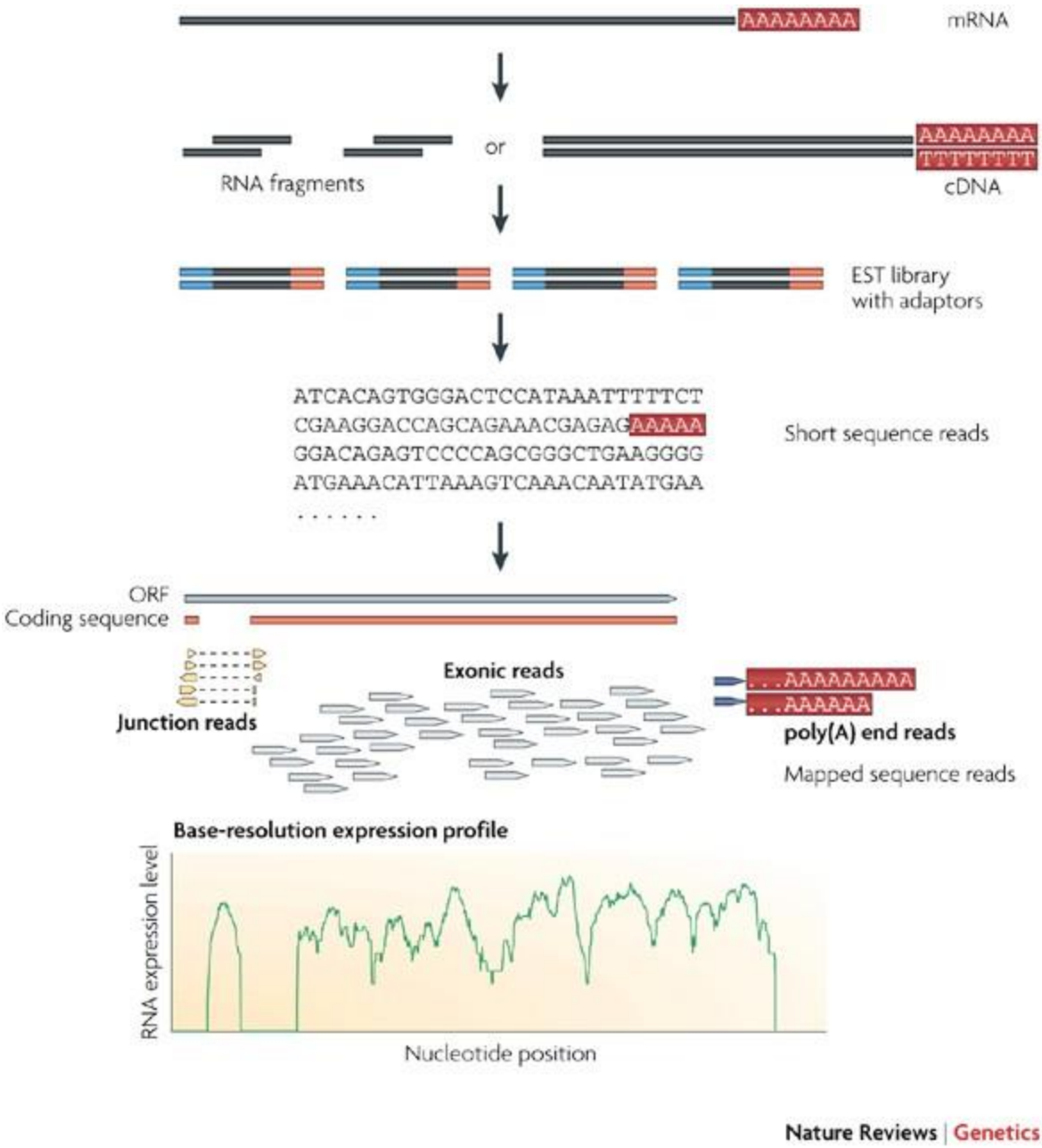


FIGURE 10 : **A typical RnA-seq experiment**. Briefly, long RNAs are first converted into a library of cDNA fragments through either RNA fragmentation or DNA fragmentation (see main text). Sequencing adaptors (blue) are subsequently added to each cDNA fragment and a short sequence is obtained from each cDNA using high-throughput sequencing technology. The resulting sequence reads are aligned with the reference genome or transcriptome, and classified as three types: exonic reads, junction reads and poly(A) end-reads. These three types are used to generate a base-resolution expression profile for each gene, as illustrated at the bottom; a yeast ORF with one intron is shown. [18]

1.6.3 MICROARRAY

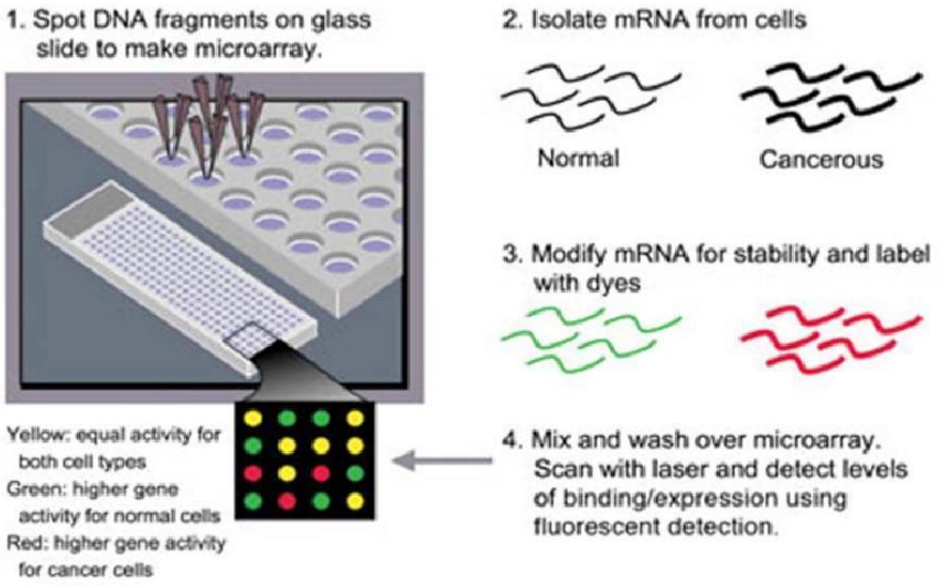


FIGURE 11 : representing the technique of microarray analysis.

Microarray analysis of gene expression is a very powerful tool for characterization of differential gene expression in many pathophysiologic processes.

The basic mechanism is that this RNA which is isolated from the tissue is hybridized to probes for the specific genes which are fixed in a grid in small microscopic spots. This technology rests on the ability to deposit many (tens of thousands) different DNA sequences on a small surface, usually a glass slide (often referred to as a "chip"). The different DNA fragments are arranged in rows and columns such that the identity of each fragment is known through its location on the array. Depending on the design of the experiment, the intensity of signal of the hybridization is normalized to the internal controls and other tissues to yield a result for each gene which tells the investigator whether a particular gene has an increased or decreased, or it is in the normal expression [22]. The result which is obtained for the sample, acts as an answer for each gene included in the microarray as to whether the gene is down or up regulated and this composite data represents a gene expression profile for the given sample.

When a large number of samples are tested, depending on the experimental designs, the use of this technique makes possible the establishment of gene expression profiles for any of the given disease states, the comparison of various subsets to determine the molecular predictors of clinical behavior, etc [23]. As the technology has advanced day by day, the sophistication of this analysis has likewise provided increasingly powerful ways of discriminating the molecular characteristics of disease states, including the identification of methylation status of genes (an epigenetic modification of expression) and alternative splicing. In addition, small variations in the DNA sequence that lead to different characteristics (such as skin color, facial features, or height) known as polymorphisms, which can cause or contribute to the development of many syndromes and diseases can be identified using microarray. This technology promises a more biologically based, individualized, and vastly improved standard which can have great clinical impact. [22,23]

2.0 OBJECTIVE :

The main objective of this study is to identify and analysis of microRNA-137 targets in diseases.

3.0 METHODOLOGY :

3.1 Literature Mining: The first step for the study was to conduct a search relating to microRNA-137 and its targets. For this, multiple tools were used - google scholar, PubMed, science direct and Sci-Hub. We searched for miRNA-137 targets using search terms : 'miRNA-137 targets for Cancer' (Google scholar) and 'Targets and MicroRNA 137' (PubMed). Following the initial literature search, 750 papers were extracted from PubMed, while approximately 800 papers from Google Scholar. The putative as well as verified targets of miRNA-137 that were identified from these papers were combined in an Excel Sheet. All the citations were downloaded and added to the sheet. The duplicates were then removed manually.

3.2 Searching Gene Symbols: In order to aid this search, we used an online database UniProt. From there, we searched for *Homo sapiens* genes and obtained their respective gene symbols.

3.3 Bioinformatic based extraction of experimentally verified targets: The miRTarBase was used to find and analyze the experimentally validated microRNA-target interactions. The key word 'hsa-mir-137' was put in the search engine and the results were downloaded onto an Excel sheet. The list incorporated target wise experimental assay used for its validation.

3.4 miRTarBase : The targets obtained from the literature mining as well as from the miRTarBase were **combined and sorted into categories**. Disease wise target segregation as well as target wise disease segregation was performed and results were summarized in a tabular form (in Excel sheet). Further, sorting based on their validation was done and targets were classified as predicted, identified or verified.

3.5 STRING analysis: The list of targets were put in the input search box under the 'multiple proteins' segment of the database. A detailed description of all individual targets was obtained. The settings relating to network type, active interaction sources, minimum required interaction score and network display options were modified wherever necessary. The biological function of each target and their mutual interactions were studied through the network obtained. The data related to the gene ontology, molecular function, KEGG pathways, Reactome pathway as well as disease-gene associations was obtained.

4.0 RESULTS AND DISCUSSION:

At the end of our study, a total of **298** targets were obtained and gathered from literature search as well as from mirTarBase database. The number of experimentally verified targets whose assays were known equated to be **255**. Out of these 255 targets, **134** were found from mirTarBase while **121** targets were found through articles during literature mining **(Table 4 & 5)**.

Table 4: Verified 121 microRNA-137 targets identified from literary search

Target	Dual-Luciferase Reporter assay	Luciferase assay	Western blot	qPCR	Micro array	NGS	pSILAC	CLIP-Seq	Other	Other assay
ACVR1	Yes	No	No	No	No	No	No	No	No	-
ADAMTS-5	Yes	No	No	No	No	No	No	No	No	-
AEG-1	No	Yes	Yes	Yes	No	No	No	No	No	-
AFM	Yes	No	No	Yes	No	No	No	No	No	-
AMPK1	No	Yes	Yes	Yes	No	No	No	No	No	-
APP	No	Yes	No	No	No	No	No	No	No	-
ARHGAP5	No	Yes	No	Yes	No	No	No	No	No	-
ASCT2	No	No	No	Yes	No	Yes	No	Yes	No	-
Atg13	Yes	No	No	No	No	No	No	No	No	-
ATG5	Yes	No	No	No	No	No	No	No	No	-
ATG7	No	No	No	Yes	No	No	No	No	No	-
AXL	No	No	Yes	No	No	No	No	No	No	-
BCL11A	Yes	No	No	No	No	No	No	No	No	-
BCL2L13	No	Yes	Yes	Yes	No	No	No	No	No	-
BDNF	No	Yes	Yes	Yes	No	No	No	No	No	-
BMI1	Yes	No	Yes	Yes	No	No	No	No	No	-
C10orf26	No	Yes	No	No	No	No	No	No	No	-
CACNA1C	No	Yes	No	No	No	No	No	No	No	-
CALN1	No	Yes	No	Yes	No	No	No	No	No	-

CAR	No	No	Yes	Yes	No	No	No	No	No	-
CASP3	No	Yes	No	No	No	No	No	No	Yes	site-directed mutagenesis assays
CDH	No	No	Yes	Yes	No	No	No	No	No	-
CDH1	No	No	Yes	Yes	No	No	No	No	No	-
CDK2	No	No	Yes	No	No	No	No	No	No	-
COL5A1	Yes	No	No	No	No	No	No	No	No	-
CORTBP1	No	Yes	No	No	No	No	No	No	Yes	-
CPLX1	No	No	Yes	No	No	No	No	No	No	-
CSMD1	No	Yes	No	No	No	No	No	No	No	-
CXCL12	Yes	No	Yes	No	No	No	No	No	No	-
DAAM1	No	Yes	No	No	No	No	No	No	No	-
DAG1	No	No	No	Yes	No	No	No	No	No	-
DCLK1	No	Yes	Yes	Yes	No	No	No	No	Yes	immunoblot analysis-2
DUSP4	No	No	Yes	Yes	No	No	No	No	No	-
E2F7	No	Yes	No	No	No	No	No	No	No	-
EDIL3	Yes	No	No	No	No	No	No	No	Yes	ELISA
EFEMP2	No	Yes	No	No	No	No	No	No	No	-
EFNB2	No	Yes	Yes	Yes	No	No	No	No	No	-
EGR2	No	No	No	Yes	No	No	No	No	No	-
EPS8	Yes	No	No	Yes	No	No	No	No	No	-
ERRA	No	Yes	No	No	No	No	No	No	No	-
FBI-1	No	Yes	Yes	Yes	No	No	No	No	No	Trans-well assay
FOXO1	Yes	No	No	No	No	No	No	No	No	-
FSTL1	No	Yes	No	No	No	No	No	No	No	-
FUN14	No	No	Yes	No	No	No	No	No	No	-
FXYD6	Yes	No	No	No	No	No	No	No	No	-
GREM1	Yes	No	No	No	No	No	No	No	No	-
GRIN2A	No	Yes	No	No	No	No	No	No	No	-

HCRT	No	Yes	No	No	No	No	No	No	No	-
HDAC9	No	No	No	Yes	No	No	No	No	No	-
HEY 2	No	No	Yes	Yes	No	No	No	No	No	-
IGFBP-5	No	Yes	No	No	No	No	No	No	No	-
IL-6	No	No	No	Yes	No	No	No	No	Yes	ELISA
KCC2	Yes	No	No	Yes	No	No	No	No	No	-
KCNMB2	No	No	Yes	No	No	No	No	No	No	-
KDM5B	No	Yes	Yes	No	No	No	No	No	No	-
KISS1	No	Yes	Yes	No	No	No	No	No	No	-
KLF12	No	No	Yes	Yes	No	No	No	No	Yes	-
LGR4	Yes	No	No	Yes	No	No	No	No	No	-
LRP6	No	Yes	No	Yes	No	No	No	No	Yes	-
LSD	No	Yes	No	No	No	No	No	No	No	-
LSD1	No	No	Yes	No	No	No	No	No	No	-
MAPK1	No	No	Yes	No	No	No	No	Yes	No	-
MAPK10	No	No	No	Yes	No	No	No	No	Yes	IHC
MCL1	No	No	No	Yes	No	No	No	No	No	-
MCP1	Yes	No	Yes	Yes	No	No	No	No	Yes	receptor internalizatio n assay
MECP2	No	No	No	Yes	No	No	No	No	No	-
MIB1	Yes	Yes	No	No	No	No	No	No	No	-
MMP2	Yes	No	Yes	Yes	No	No	No	No	No	-
MMP9	No	No	Yes	Yes	No	No	No	No	No	-
MSH2	No	No	No	No	No	No	No	No	Yes	Formalin-Fixed and Paraffin-Embedded (FFPE) Tissues RNA Isolation and Expression Assay
MSI1	Yes	No	No	No	No	No	No	No	No	-

MYO1C	No	Yes	Yes	Yes	No	No	No	No	No	-
NEUROD4	Yes	No	No	No	No	No	No	No	No	-
NF1	No	Yes	No	No	No	No	No	No	No	-
Notch 1	Yes	No	No	No	No	No	No	No	No	-
NOX4	Yes	No	Yes	No	No	No	No	No	No	-
NSF	No	No	No	Yes	No	No	No	No	No	-
OXR1	No	No	Yes	Yes	No	No	No	No	No	-
PAK2	No	Yes	Yes	No	No	No	No	No	No	-
PAQR3	No	Yes	Yes	No	No	No	No	No	No	-
PBX3	Yes	No	No	Yes	No	No	No	No	No	-
PGE2	Yes	No	No	No	No	No	No	No	No	-
PIK3R3	Yes	No	Yes	Yes	No	No	No	No	No	-
PTBP1	No	Yes	No	No	No	No	No	No	No	-
PTGFRN	No	Yes	Yes	Yes	No	No	No	No	No	
PTN	No	No	Yes	Yes	No	No	No	No	No	-
PTP4A3	No	No	Yes	Yes	No	No	No	No	No	-
Rac1	No	Yes	Yes	No	No	No	No	No	No	-
RANKL	Yes	No	Yes	Yes	No	No	No	No	No	-
RasGRF1	Yes	No	No	No	No	No	No	No	No	-
RB1	No	No	Yes	No	No	No	No	No	No	-
SCAMP1	No	Yes	Yes	Yes	No	No	No	No	No	-
SFRP4	Yes	No	Yes	Yes	No	No	No	No	Yes	-
SLC22A18	No	Yes	No	No	No	No	No	No	No	-
SMAD	Yes	No	No	No	No	No	No	No	No	-
SNAI1	No	No	No	Yes	No	No	No	No	No	-
SORL1	No	No	No	No	No	No	No	No	Yes	HITS-CLIP assay
SOX 9-b	No	No	No	Yes	No	No	No	No	No	-
SP1	No	No	No	No	No	No	No	No	No	-
SPT	No	Yes	Yes	Yes	No	No	No	No	Yes	ELISA
SRC1	No	No	No	No	No	No	No	No	No	-

SRC2	No	No	No	No	No	No	No	No	No	-
SRC3	No	No	No	No	No	No	No	No	No	-
STC1	No	No	No	No	No	No	No	No	No	-
TBX2	No	No	No	No	No	No	No	No	No	-
TFAM	No	No	No	Yes	No	No	No	No	No	-
TGFA	No	No	No	Yes	No	No	No	No	No	-
TINCR	No	No	No	No	No	No	No	No	No	-
TNFAIP1	Yes	Yes	Yes	Yes	No	No	No	No	No	-
TRIAP1	No	No	No	Yes	No	No	No	No	No	-
TRIM24	No	No	No	Yes	No	No	No	No	No	-
TRIM25	Yes	No	Yes	No	No	No	No	No	No	-
TRPC3	Yes	No	Yes	Yes	No	No	No	No	No	-
TWIST1	No	No	Yes	Yes	No	No	No	No	No	-
VKORC1	No	Yes	No	Yes	No	No	No	No	No	-
WIF1	No	No	No	Yes	No	No	No	No	No	-
WNT2B	Yes	No	No	No	No	No	No	No	No	-
WNT7A	No	No	Yes	No	No	No	No	No	No	-
YB-1	No	Yes	No	No	No	No	No	No	No	-
YWHAZ	No	No	Yes	No	No	No	No	No	No	-
ZBTB7A	No	No	Yes	Yes	No	No	No	No	No	-

Table 5: Verified microRNA-137 Targets from miRTarBase Database

Targets	Reporter assay	qPCR	Micro array	NGS	pSILAC	Other	CLIP-Seq	Western Blotting
TCF4	No	No	No	No	No	No	yes	No
SERPINA3	yes	No	No	No	No	No	No	No
JAG1	No	No	No	yes	No	No	yes	No
AKT2	yes	yes	No	yes	No	No	yes	yes
BMP7	No	No	No	yes	No	No	yes	No
BNIP3L	No	No	No	yes	No	No	yes	No
RUNX2	No	No	No	yes	No	No	yes	No
CDC42	yes	No	No	No	No	No	No	yes
CDK6	yes	yes	No	No	No	No	No	yes
CSE1L	yes	yes	No	No	No	No	No	yes
CTBP1	yes	yes	No	No	No	yes	No	yes
DBN1	yes	yes	No	No	No	yes	No	yes
E2F6	No	No	No	yes	No	No	yes	No
EGFR	No	No	No	yes	No	No	yes	No
ESRRA	yes	No	No	No	No	No	No	yes
EZH2	yes	yes	No	No	No	yes	No	yes
GLO1	No	No	No	yes	No	No	yes	No
NR3C1	No	No	No	yes	No	No	yes	No
TNC	yes	yes	No	No	No	yes	No	yes
IL6	No	yes	yes	No	No	No	No	yes
KCNA2	yes	No	No	No	No	No	No	No
KIT	yes	yes	No	No	No	No	No	yes
MEF2A	No	No	No	yes	No	No	yes	No
MITF	No	No	No	yes	No	No	yes	No

COX2	No	No	No	yes	No	No	yes	No
NOTCH1	No	No	No	yes	No	No	yes	No
YBX1	No	No	No	yes	No	No	yes	No
PTGS2	No	No	No	yes	No	No	yes	No
PXN	No	No	No	yes	No	No	yes	No
RORA	No	No	yes	No	No	No	No	No
SLC1A5	yes	yes	No	No	No	No	No	yes
AURKA	yes	yes	No	No	No	No	No	yes
TBX3	yes	yes	No	No	No	No	No	yes
TGFB2	No	yes	No	No	No	No	No	yes
ZNF24	yes	yes	No	No	No	No	No	yes
NCOA3	yes	yes	No	No	No	No	No	yes
CUL4A	yes	yes	No	No	No	No	yes	No
KLF4	No	No	No	yes	No	No	yes	No
KDM4A	No	yes	yes	No	No	No	No	yes
DELE1	yes	yes	No	No	No	No	No	yes
COQ7	No	No	No	yes	No	No	yes	No
NCOA2	No	No	No	yes	No	No	yes	No
GLIPR1	No	No	No	yes	No	No	yes	No
Nr1i3	yes	No	No	No	No	No	No	yes
KDM1A	yes	yes	No	No	No	No	No	yes
CCDC59	yes	No	No	No	No	No	No	No
ZNF804A	No	No	No	yes	No	No	yes	No
MTDH	No	No	No	yes	No	No	yes	No
FMNL2	No	No	No	yes	No	No	yes	No
FUNDC1	No	No	No	yes	No	No	yes	No
FOXK1	No	No	No	yes	No	No	yes	No
NUP43	No	No	No	yes	No	No	yes	No

ASGR2	yes	No	No	No	No	No	No	No
ATOX1	No	No	No	yes	No	No	yes	No
BCAT1	yes	yes	No	No	No	No	No	yes
CHRM3	No	No	No	yes	No	No	yes	No
CNN2	No	No	No	yes	No	No	yes	No
DR1	No	No	No	yes	No	No	yes	No
FKBP1A	yes	yes	No	No	No	No	No	yes
GOLGA3	yes	No	No	No	No	No	No	No
GTF2E1	yes	yes	No	No	No	No	No	yes
H3F3B	No	No	No	yes	No	No	yes	No
HOXC8	yes	yes	No	No	No	No	No	No
MET	No	No	No	yes	No	No	yes	No
NFYB	No	No	No	yes	No	No	yes	No
PAFAH1B2	No	No	No	yes	No	No	yes	No
PAX9	No	No	No	yes	No	No	yes	No
PLS1	No	No	No	yes	No	No	yes	No
RPL9	yes	yes	yes	No	No	No	No	yes
RREB1	No	No	No	yes	No	No	yes	No
SNRPE	No	No	No	yes	No	No	yes	No
SSFA2	No	No	No	yes	No	No	yes	No
TEF	yes	yes	No	No	No	yes	No	yes
TLR3	No	No	No	yes	No	No	yes	No
TRAM2	yes	yes	No	No	No	No	No	yes
ZC3H11A	No	No	No	yes	No	No	yes	No
AP5Z1	No	yes	No	No	No	No	No	yes
SUPT7L	No	No	No	yes	No	No	yes	No
HNRNPDL	yes	yes	yes	No	No	yes	No	yes
LHFPL2	No	No	No	yes	No	No	yes	No

TRIM13	No	No	No	yes	No	No	yes	No
EIF1	No	No	No	yes	No	No	yes	No
ZNF267	No	No	No	yes	No	No	yes	No
PAIP1	No	No	No	yes	No	No	yes	No
PAPD7	No	No	No	yes	No	No	yes	No
LIMCH1	No	No	No	yes	No	No	yes	No
GRAMD4	No	No	No	yes	No	No	yes	No
UBXN4	No	No	No	yes	No	No	yes	No
SLC39A14	No	No	No	yes	No	No	yes	No
BTBD1	No	No	No	yes	No	No	yes	No
RBM27	No	No	No	yes	No	No	yes	No
MRPS21	No	No	No	yes	No	No	yes	No
SNRK	No	No	No	yes	No	No	yes	No
TMA16	No	No	No	yes	No	No	yes	No
CTPS2	No	No	No	yes	No	No	yes	No
USP28	No	No	No	yes	No	No	yes	No
GPBP1L1	No	No	No	yes	No	No	yes	No
SPCS3	No	No	No	yes	No	No	yes	No
CIDEC	No	No	No	yes	No	No	yes	No
GIGYF1	No	No	No	yes	No	No	yes	No
BCL11B	No	No	No	yes	No	No	yes	No
NBEAL1	No	No	No	yes	No	No	yes	No
UBE2Z	No	No	No	yes	No	No	yes	No
LONRF3	No	No	No	yes	No	No	yes	No
PTGES2	No	No	No	yes	No	No	yes	No
KLHL15	yes	No	No	No	No	No	No	No
DDI2	yes	yes	No	No	No	No	No	No
SLX4	No	No	No	yes	No	No	yes	No

TXNDC17	No	No	No	yes	No	No	yes	No
SFT2D3	No	No	No	yes	No	No	yes	No
TMEM41A	yes	No	No	No	No	No	No	No
SFXN5	No	No	No	yes	No	No	yes	No
FNIP1	No	No	No	yes	No	No	yes	No
PCMTD1	No	No	No	yes	No	No	yes	No
COX20	yes	yes	No	No	No	No	No	yes
RFTN2	No	No	No	yes	No	No	yes	No
LYPD6	No	No	No	yes	No	No	yes	No
FUT11	No	No	No	yes	No	No	yes	No
ADAMTS15	No	No	No	yes	No	No	yes	No
AGO4	No	No	No	yes	No	No	yes	No
VMA21	No	No	No	yes	No	No	yes	No
YTHDF3	No	No	No	yes	No	No	yes	No
PPP1R37	No	No	No	yes	No	No	yes	No
ZNF326	No	No	No	yes	No	No	yes	No
EOGT	No	No	No	yes	No	No	yes	No
SREK1IP1	No	No	No	yes	No	No	yes	No
ZNF678	No	No	No	yes	No	No	yes	No
XKR7	No	No	No	yes	No	No	yes	No
SLCO4C1	No	No	No	yes	No	No	yes	No
ZNF772	No	No	No	yes	No	No	yes	No
H3F3C	No	No	No	yes	No	No	yes	No
LMOD2	No	No	No	yes	No	No	yes	No
ZFP62	No	No	No	yes	No	No	yes	No
SLC35E2B	No	No	No	yes	No	No	yes	No

A number of 27 of the targets found through papers were reported as predicted targets for whom no experimental methods were conducted, while 16 were identified targets for which some level of experimentation was conducted but no confirmed assays were performed.

The obtained targets were sorted as disease specific targets (targets common to a disease) and target wise diseases (wherein all diseases in which a particular target plays a role were listed with the target), as summarized in **Table 6 (below)**. From the recorded data, it was interpreted that few disorders had a high number of targets associated with them as in Breast Cancer (29), Glioma (41) and Schizophrenia (37).

TABLE 6: List of Disease-wise microRNA-137 Targets

DISEASES	TARGETS
CANCERS	
Gastrointestinal stromal tumor + Gastric Cancer	CDH, CDH1, KRT1, VIM GENE, TWIST1, AKT2, COL5A1, CUL4A, KLF12, MYO1C, PBX3, CDK6, CtBP1, EZH2, Cdc42, MYO1C, NUP43
Human colon carcinoma (cancer)	C21ORF66, CSDA, DBN1, DPYSL3, KIAA1671, PCNA, PPP1CB
Acute myeloid leukemia (cancer)	COL14A1, EIF4G1, HDAC3, HDAC8, MRPL49, PSMA2, REL, TRIM25, c-kit, EGFR (tyrphostin AG-1478), TOP1, MRPL49, REL,
Bladder cancer	PAQR3, PAK2,
Breast cancer	ASPH, BCL11A, BMP7, CCND1, CtBP1, DUSP4, EDIL3, FSTL1, ERRA, FUN-14, FUNDC1, GIGYF1, HEY2, JAG1, KCNMB2, KDM4A, KDM5B, NOTCH1, TBX3, SRC3, SORL1, SRC1, SRC2, WNT7A, YB-1, ZNF24, ASCT2, ZEB2, EPS8
Cervical cancer	GREM1, EZH2, Notch 1, ASCT2
Cholangiocarcinoma	WNT2B
Colon cancer	DCLK1, DCLK2, GLS, KDM1A (LSD1), TCF4
Colorectal cancer	APC, ARHGAP5, c-MET, DCLK1, FMNL2, KDM1A (LSD1), MECP27, TINCR, Cdc42, LSD1, MSH2, MSI1, TINCR, YB-1, ASCT2
Head and neck cancer	MTDH, Notch 1
Hepatocellular carcinoma	AFM, AKT2, Cdc42, EZH2, FBI-1, FOXO1, HEY 2, Notch 1, TUG1, EZH2 (liver cancer), FOXD3 (liver tumor), NR1I3 (liver disorder), P13K/AKT signaling pathway (liver tumor)

Human thyroid carcinoma (cancer) + Papillary thyroid carcinoma (cancer)	CXCL12, EGFR, CXCL12, SDF-1
Lymphoblastic leukemia (cancer)	KDM5B
Melanoma (cancer)	AURKA, CDH1, CDK-2, CDK6, COX2, CtBP1, GLO 1, GLS, MAPK1, MAPK3, MET (c-Met), MITF, PIK3R3, PRC2, TBX3, YB-1, MITF, MET (c-Met), PAK2
Metastasis of hypertrophic scar fibroblasts	MMP9
Oesophageal squamous cell carcinoma (cancer)	DAAM1, EZH2, PXN
Oral squamous cell carcinoma (cancer)	E2F6, NCOA2, RB1,
Ovarian cancer	AEG-1, CTCF, CXCL12, EZH2, FUNDC1, MCL1, SCAMP1, XIAP, SNAI1
Pancreatic cancer	ATG5 (Chemosensitises cells) KLF12, Notch1, MRGBP, PXN, GSTM4, ICOSLG, TNC, ROBO2
Multiple myeloma (cancer)	AURKA, BAD, CHEK2, MITF, PARP, MCL1
Osteosarcoma (cancer)	EFEMP2, FXYD6, TP53, TRIAP 1
Lung cancer + Nasopharyngeal Carcinoma (cancer)	BMP7, Cdc42, CDDP, CDK6, E2F7, H446, KIT/CD117, LSD1, PXN, SLC1A5, SLC22A18, SRA1, SRC3, TFAP2C, TGAF, AXL, CASP3, TWIST1
Prolactinomas tumorigenesis (cancer) + Pituitary adenoma (cancer)	KDM4A, AKT2, EGFR, WIF1
GBM / Glioma Cancer	BMI1, CDK6, hnRNPI, HOTAIRM, LRP6, PTGFRN, PTGS2, U251, COX2, EGFR, LSD1, PTP4A3, GLIPR1, EZH2, RND1, E2F3a, BCL2, CENPE, CSE1L, CXCL12, FOXC1, FOXK1, MSI1, NUCKS1, PDGFRa, Rac1, RasGRF1, RTVP-1, SLC1A5, SP1, STC1, TJP2, TRIM24, CAR, E2F7, GABRA1, C-kit, YB-1, TGF, RasGRF1, EZH2, KDM1A, CSE1L, NF1
Chondrogenic differentiation of DPSCs (dental pulp stem cells)	PTN
Tumorigenesis (cancer)	DELE1
Prostate cancer	EZH2, JNK3, KDM5B, LGR4, NOX4, TRIM24,
Retinoblastoma (cancer)	COX2, KDM4A, PGE2
Spinal cord cancer	CDK6

Uveal melanoma (cancer)	CDK -2, CDK-6, MAPK1, MAPK3
Laryngeal squamous cell carcinoma (cancer)	NCK1
Renal cell carcinoma (cancer)	PIK3AP1, ZBTB7A, ZFPM2-AS1
Doxorubicin (DXR) Resistance of Osteosarcoma Cells (cancer)	TRIAP 1
NON CANCEROUS DISEASES	
Chronic heart failure	SERPINA3
Organogenesis and differentiation	RORA
Steroid induced osteonecrosis of the femoral head	RUNX2
Psychosis	COX6A
Atrial fibrillation atrial fibrillation (AF)	VKORC1
VSMC Proliferation	TRPC3
Wake to sleep ratio	HCRT
Cognitive dysfunction	PTN
Glaucoma	SRC
Diabetic kidney diseases	Notch1
Ischemic brain injury	Notch1
abnormal brain development	TLX (NR2E1)
Neuropsychiatric disorders	KCC2, RB1 (p-Rb), SHANK2
Neurotoxicity	TNFAIP1
Neuronal cell injury	PTEN
spinal cord inflammation and oxidative stress	NEUROD4
Protects PC12 Cells Against Oxidative Stress	nNOS, CAPN2
Parkinson's disease	OXR1
Spinal cord injury	KDM4A, Notch1, SFRP4
Transient site-specific copy gain	KDM4A
Rheumatoid arthritis	COX2

Schizophrenia	ACSL6, ATXN1, BDNF, BRD1, CACNA1C, CALN1, CHGA, COX6A2, CPLX1, CSMD1, EFNB2, ERBB4, FXYD6, FZD3, GABRA1, GRIA1, GRIA4, GRIN2A, GRM5, GSK3B, HTR2C, IMPA2, LDLRAD4, MLC1, NR3C1, NRGN, SYN2, SYN3, SYT1, TCF4, WBP1L, ZNF804A, NUCKS1, CSE1L, MLC1, VEGFA
Rat brain amygdala	LSD
Respiratory disorders caused by COVID19	CCDC59
Muscle maintenance	DAG1
Cardiovascular diseases, Atherosclerosis	IGFBP-5, MCP1
Osteogenesis, Orthoartritis, Osteoblasts proliferation, Osteoporosis	RANKL, ADAMTS-5, KLF4, KDM4A
Obliterative bronchiolitis	MAX, SMAD
Neuropathic pain	KCNA2, TNFAIP1
Bovine myoblast proliferation	LncPRRX1, Cdc42
Microphthalmia	MITF
Ischemic brain injury	BNIP3L, EDIL3
Intervertebral disc degeneration	ACVR1
Pyruvate kinase M1/2(PKM) Expression (Biological Process)	PTBP1
Vasopressin response (Biological Process)	AQP2
Autophagy	ATG7
Nerve injury-induced neuropathic pain	KCNA2
Hypoxia- induced proliferation of pulmonary artery smooth muscle cells (Biological Process)	CAPN2
Brachial plexus root avulsion (BPRA)	CAPN2
Multiple sclerosis	HDAC9
Huntington disease	HTT
Intellectual disability	EZH2, KLF4, MITF
HG- induced vascular injury	AMPKa1

Myocardial ischemic injury	Nourin
Hepatic Dietary fatty acid composition defect	Cdc42, ELOVL4
AD- dementia	MEF2A
Alzheimer	APP, SNHG19, SPT, TNFR1, SPTLC1

The list of targets was then subjected to String analysis via the database, STRING. This resulted in production of tables based on: KEGG pathways, Reactome Pathways, Disease gene associations, Molecular Function, Biological process and many more criteria. The results were downloaded for the first 5 mentioned areas. All of them were sorted in increasing order of their respective False discovery rates (smallest to largest, since we need the ones with least error on top).

The **table for the KEGG pathway** showed 96 KEGG pathways. After their sorting, the top 10 categories were selected and the table for 10 is given below.

TABLE 7 : STRING analysis for KEGG

Term description	Observed gene count	Background gene count	Strength	False discovery rate	Matching proteins in your network (labels)
Pathways in cancer	30	515	0.71	1.78E-10	MAPK1,MMP2,MSH2,LRP6,CDH1,SMAD2,PIK3R3,CDK6,CDK2,RB1,EGFR,WNT7A,KIT,CTBP1,CASP3,MET,SP1,PTGS2,HEY2,WNT2B,NCOA3,MMP9,FOXO1,AKT2,CXCL12,IL6,MITF,TGFA,CDC42,NOTCH1
Breast cancer	15	146	0.96	6.24E-08	MAPK1,LRP6,PIK3R3,CDK6,RB1,EGFR,WNT7A,KIT,SP1,HEY2,WNT2B,NCOA3,AKT2,TNFSF11,NOTCH1
Endocrine resistance	11	94	1.02	2.86E-06	MAPK1,MMP2,PIK3R3,RB1,EGFR,SP1,N

					COA3,MMP9,SRC,AKT2,NOTCH1
AGE-RAGE signaling pathway in diabetic complications	11	96	1.01	2.86E-06	MAPK1,MMP2,CCL2,SMAD2,PIK3R3,NOX4,CASP3,FOXO1,AKT2,IL6,CDC42
EGFR tyrosine kinase inhibitor resistance	10	77	1.06	3.07E-06	MAPK1,PIK3R3,EGFR,AXL,MET,NF1,SRC,AKT2,IL6,TGFA
MicroRNAs in cancer	13	159	0.86	3.94E-06	MAPK1,PIK3R3,TNC,CDK6,EGFR,CASP3,MET,EZH2,PTGS2,MCL1,MMP9,BMI1,NOTCH1
Proteoglycans in cancer	14	194	0.81	4.49E-06	MAPK1,MMP2,TWIST1,SMAD2,PIK3R3,EGFR,WNT7A,CASP3,MET,WNT2B,MMP9,SRC,AKT2,CDC42
Gastric cancer	12	146	0.86	9.03E-06	MAPK1,LRP6,CDH1,SMAD2,PIK3R3,CDK2,RB1,EGFR,WNT7A,MET,WNT2B,AKT2
Pancreatic cancer	9	71	1.05	9.83E-06	MAPK1,SMAD2,PIK3R3,CDK6,RB1,EGFR,AKT2,TGFA,CDC42
Human papillomavirus infection	17	324	0.67	9.86E-06	MAPK1,PIK3R3,TNC,CDK6,CDK2,RB1,EGFR,WNT7A,TLR3,CASP3,PTGS2,HEY2,WNT2B,FOXO1,AKT2,CDC42,NOTCH1

Below is a description about the molecular functions of 5 such targets taken from KEGG, from under the first category **“Pathways in Cancer”**:

1. **MAPK1 -** It stands for mitogen-activated protein kinase (MAPK). This pathway (MAPK pathway) has a role in regulating some crucial cell functions like expression (of gene) in cells and cellular growth hence their survival. Any abnormality or defect in this pathway may lead to uncontrolled cell proliferation and resistance to apoptosis (like in cancers). [24]
2. **WNT2B** - Wnt Family Member 2B (WNT2B) belongs to WNT family of proteins which is known to play roles in cellular development and carcinogenesis. A study has reported that miRNA-137 suppresses the proliferation, migration, and invasion of cancer cells (cholangiocarcinoma cells) by targeting WNT2B [25].
3. **EGFR -** Epidermal growth factor receptor has been known to be related to malignancies (epithelial). Its activity enhances tumor growth and invasion hence metastasis [26].
4. **Notch1** - Being tightly connected to many signaling pathways that are involved in carcinogenesis, this signaling helps in cell fate determination. It also plays an important role in cell growth and division (proliferation), maturation (differentiation), and self-destruction (apoptosis) [27].
5. **CDK6 -** It controls how quickly cells divide and hence acts as an oncogenic kinase that regulates cell cycle [28].

The table for the **REACTOME** pathways showed 85 pathways. Sorting of the top 10 categories were selected.

TABLE 8 : STRING analysis for REACTOME PATHWAYS

Term description	Observed gene count	Background gene count	Strength	False discovery rate	Matching proteins in your network (labels)
Signal Transduction	73	2540	0.41	4.47E-11	GOLGA3,MAPK1,MMP2,CCL2,TNFAIP1,NR3C1,DUSP4,EGR2,SNAI1,H3-3B,CHRM3,LRP6,MIB1,CDH1,SMAD2,PIK3R3,CDK2,PXN,EGFR,APP,WNT7A,WIF1,KIT,FMNL2,CTBP1,HCRT,FSTL1,DUSP16,AXL,CASP3,PAK2,MET,EZH2,SP1,PTN,NUP43,PRKAA1,RAC1,PTBP1,NF1,TGFB2,KISS1,HEY2,WNT2B,COL5A1,NCOA3,MMP9,KDM4A,YBX1,SRC,BMI1,LGR4,FOXO1,TRPC3,ARHGAP5,AKT2,DBN1,DAAM1,CXCL12,YWHAZ,FKBP1A,KDM1A,IL6,NCOA2,TGFA,RASGRF1,HDAC9,BDNF,PAQR3,SLC1A5,MEF2A,CDC42,NOTCH1
Disease	53	1702	0.44	1.11E-08	MAPK1,NR3C1,MSH2,H3-3B,LRP6,MIB1,CDH1,SMAD2,PIK3R3,NOX4,CDK6,CDK2,RB1,EGFR,GTF2E1,APP,ADAMTS5,KIT,CTBP1,TLR3,DUSP16,PAK2,MET,EZH2,TRIM25,SP1,TRIM24,NUP43,ATG7,RAC1,NF1,HEY2,SRC,FGR,FOXO1,SLC22A18,AKT2,YWHAZ,FKBP1A,KDM1A,IL6,MECP2,SNRPE,TGFA,HDAC9,DAG1,FXYD6,TENT4A,RPL9,BCL11A,MYO1C,CDC42,NOTCH1
Diseases of signal transduction by growth factor receptors and second messengers	26	430	0.73	1.11E-08	MAPK1,LRP6,MIB1,SMAD2,PIK3R3,NOX4,EGFR,KIT,CTBP1,DUSP16,MET,TRIM24,ATG7,RAC1,NF1,HEY2,SRC,FOXO1,AKT2,FKBP1A,MECP2,TGFA,HDAC9,TENT4A,BCL11A,NOTCH1

RNA Polymerase II Transcription	46	1337	0.48	1.11E-08	MAPK1,AURKA,NR3C1,MSH2,NFYB,TWIST1,H3-3B,RORA,SMAD2,CDK6,CDK2,RB1,EGFR,GTF2E1,KIT,ZNF267,MET,EZH2,E2F7,SP1,GRIN2A,PRKAA1,MT-CO2,KDM5B,NR1I3,HEY2,RUNX2,YBX1,SRC,BMI1,FOXO1,TRPC3,BNIP3L,E2F6,AKT2,YWHAZ,ESRRA,IL6,MECP2,SNRPE,TGFA,HDAC9,BDNF,TRIAP1,ZC3H11A,NOTCH1
Gene expression (Transcription)	49	1476	0.47	1.11E-08	MAPK1,AURKA,NR3C1,MSH2,NFYB,TWIST1,H3-3B,RORA,SMAD2,CDK6,CDK2,RB1,EGFR,GTF2E1,KIT,ZNF267,MET,EZH2,E2F7,SP1,GRIN2A,NUP43,PRKAA1,MT-CO2,KDM5B,NR1I3,HEY2,RUNX2,YBX1,SRC,BMI1,FOXO1,TRPC3,BNIP3L,E2F6,AKT2,YWHAZ,ESRRA,IL6,MECP2,SNRPE,TGFA,HDAC9,BDNF,TFAM,TRIAP1,ZC3H11A,MYO1C,NOTCH1
Generic Transcription Pathway	43	1215	0.5	1.35E-08	MAPK1,AURKA,NR3C1,MSH2,NFYB,TWIST1,H3-3B,RORA,SMAD2,CDK6,CDK2,RB1,EGFR,KIT,ZNF267,MET,EZH2,E2F7,SP1,GRIN2A,PRKAA1,MT-CO2,KDM5B,NR1I3,HEY2,RUNX2,YBX1,SRC,BMI1,FOXO1,TRPC3,BNIP3L,E2F6,AKT2,YWHAZ,ESRRA,IL6,MECP2,TGFA,HDAC9,BDNF,TRIAP1,NOTCH1
ESR-mediated signaling	14	190	0.82	2.40E-05	MAPK1,MMP2,H3-3B,PIK3R3,EGFR,SP1,NCOA3,MMP9,SRC,AKT2,CXCL12,KDM1A,NCOA2,TGFA
Signaling by Receptor Tyrosine Kinases	22	521	0.57	5.68E-05	MAPK1,DUSP4,EGR2,PIK3R3,PXN,EGFR,KIT,AXL,PAK2,MET,PTN,RAC1,PTBP1,COL5A1,MMP9,SRC,AKT2,CXCL12,TGFA,BDNF,MEF2A,CDC42
MAPK family signaling cascades	17	322	0.67	7.07E-05	MAPK1,DUSP4,EGFR,KIT,DUSP16,PAK2,MET,RAC1,NF1,NCOA3,SRC,FOXO1,IL6,TGFA,RASGRF1,PAQR3,C

					DC42
Signaling by Interleukins	20	453	0.59	8.59E-05	MAPK1,MMP2,CCL2,DUSP4,TWIST1,RORA,PIK3R3,APP,CASP3,PAK2,PTGS2,MCL1,MMP9,FOXO1,YWHAZ,IL6,MEF2A,CNN2,HNRNPDL,CDC42

Out of the top 10 sorted, 3 pathways were analyzed thoroughly; signal transduction,diseases of signal transduction by growth factor receptors and second messengers, gene expression. Total 5 targets were studied in these criteria.

- **SIGNAL TRANSDUCTION :**

1. **CDC42-** The protein encoded by this gene is a small GTPase of the Rho-subfamily, which regulates signaling pathways that control diverse cellular functions including cell morphology, migration, endocytosis and cell cycle progression. This protein is highly similar to Saccharomyces cerevisiae Cdc 42, and is able to complement the yeast cdc42-1 mutant.
2. **TGFA**- This gene encodes a growth factor that is a ligand for the epidermal growth factor receptor, which activates a signaling pathway for cell proliferation, differentiation and development. This protein may act as either a transmembrane-bound ligand or a soluble ligand. This gene has been associated with many types of cancers, and it may also be involved in some cases of cleft lip/palate.

- **DISEASES OF SIGNAL TRANSDUCTION BY GROWTH FACTOR RECEPTORS AND SECOND MESSENGERS :**

3. **HEY2-** This gene encodes a member of the hairy and enhancer of split-related (HESR) family of basic helix-loop-helix (bHLH)-type transcription factors. The encoded protein forms homo- or hetero-dimers that localize to the nucleus and interact with a histone deacetylase complex to repress transcription. Expression of this gene is induced by the Notch signal transduction pathway. Two similar and redundant genes in mice are required for embryonic cardiovascular development, and are also implicated in neurogenesis and somitogenesis.
4. **TRIM24-** The protein encoded by this gene is a member of the tripartite motif (TRIM) family. The TRIM motif includes three zinc-binding domains, a RING, a B-box type 1 and a B-box type 2, and a coiled-coil region. The protein localizes to cytoplasmic bodies.

- **GENE EXPRESSION :**

5. **FOXO1-** a cell specific core transcription factor for endometrial remodeling and homeostasis during menstrual cycle and early pregnancy. FOXO1 inhibits prostate cancer cell proliferation via suppressing E2F1 activated NPRL2 expression.

The table for **Molecular Functions** showed multiple categories. Top ten values were selected.

TABLE 9 : STRING analysis for MOLECULAR FUNCTIONS

Term description	Observed gene count	Background gene count	Strength	False discovery rate	Matching proteins in your network (labels)
Protein binding	142	7242	0.24	1.47E-12	GOLGA3,MAPK1,AURKA,MMP2,CCL2,TNFAIP1,NR3C1,MSH2,IGFBP5,NFYB,TBX2,TWIST1,EGR2,NEUROD4,SNAI1,H3-3B,MSI1,TBX3,SORL1,LRP6,RORA,BTBD1,CDH1,SMAD2,PIK3R3,CSE1L,NOX4,TNC,CDK6,TEF,CACNA1C,CDK2,RB1,PXN,EGFR,CXADR,APP,ADAMTS5,WNT7A,WIF1,KIT,FMNL2,STC1,CTBP1,HCRT,SLX4,EDIL3,TLR3,DUSP16,AXL,AGXT,CPLX1,EFEMP2,CASP3,PAK2,MET,EZH2,E2F7,ZBTB7A,TRIM25,FOXK1,SP1,SUPT7L,PLS1,MTDH,TRIM24,PTN,SHANK2,PRKAA1,ATG7,RAC1,DCLK1,TGFB2,KISS1,KDM5B,PTGS2,HEY2,MCL1,WNT2B,DR1,RUNX2,COL5A1,NCOA3,MMP9,KDM4A,YBX1,SRC,FGR,CUL4A,BMI1,FOXO1,SLC22A18,BNIP3L,E2F6,ARHGAP5,DBN1,DAAM1,CXCL12,BMP7,YWHAZ,NSF,TCF4,TNFSF11,ZNF24,FKBP1A,KDM1A,ESRRA,IL6,SLC12A5,ITPRID2,MITF,MECP2,NBEAL1,NCOA2,CTPS2,TGFA,ACVR1,RASGRF1,HDAC9,SFRP4,BDNF,WBP1L,DDI2,TFAM,FNIP1,LIMCH1,ATOX1,ATG13,PAFAH1B2,DAG1,TRIAP1,MEF2A,CNN2,SCAMP1,TENT4A,EFNB2,EPS8,BCL11A,MYO1C,CDC42,NOTCH1,GREM1
Binding	187	12838	0.11	5.45E-07	HOXC8,GOLGA3,MAPK1,AURKA,MMP2,CCL2,TNFAIP1,AFM,NR3C1,MSH2,IGFBP5,NFYB,TBX2,TWIST1,EGR2,NEUROD4,SNAI1,H3-3B,CHRM3,CCDC59,MSI1,TBX3,S

					ORL1,LRP6,RORA,MIB1,BTBD1,CDH1,SMAD2,PIK3R3,CSE1L,NOX4,TNC,RBM27,CDK6,TEF,CACNA1C,CDK2,RB1,PXN,EGFR,GTF2E1,CXADR,APP,ADAMTS5,WNT7A,WIF1,KIT,FMNL2,STC1,CTBP1,HCRT,SLX4,FSTL1,SNRK,EDIL3,TLR3,DUSP16,ZNF267,AXL,AGXT,PAIP1,ZNF804A,CPLX1,EFEMP2,CASP3,PAK2,MET,CHDH,EZH2,COQ7,E2F7,ZBTB7A,TRIM25,FOXK1,SP1,GRIN2A,SUPT7L,PLS1,MTDH,TRIM24,PTN,SHANK2,PTGES2,PRKAA1,ATG7,ASGR2,GPBP1L1,TRIM13,RAC1,PTBP1,BCL11B,NF1,DCLK1,UBE2Z,MT-CO2,PAX9,TGFB2,KISS1,KDM5B,PTGS2,NR1I3,HEY2,MCL1,FUNDC2,WNT2B,DR1,RUNX2,LONRF3,COL5A1,NCOA3,MMP9,KDM4A,YBX1,GLO1,PBX3,SRC,FGR,CUL4A,BMI1,KLF12,FOXO1,TRPC3,RREB1,SLC22A18,BNIP3L,E2F6,ARHGAP5,AKT2,DBN1,DNASE1L3,VKORC1,DAAM1,CALN1,CXCL12,BMP7,YWHAZ,NSF,TCF4,TNFSF11,ZNF24,FKBP1A,KDM1A,ESRRA,IL6,SLC12A5,ITPRID2,MITF,MECP2,NBEAL1,NCOA2,SNRPE,CTPS2,TGFA,ACVR1,RASGRF1,HDAC9,SFRP4,BDNF,WBP1L,DDI2,EIF1,TFAM,FNIP1,LIMCH1,ATOX1,ATG13,PAFAH1B2,DAG1,SLC1A5,TRIAP1,SERPINA3,MEF2A,CNN2,MRPS21,SCAMP1,HNRNPDL,TENT4A,ZC3H11A,EFNB2,EPS8,RPL9,BCL11A,MYO1C,CDC42,NOTCH1,GREM1
Double-stranded DNA binding	48	1661	0.41	2.11E-06	HOXC8,NR3C1,MSH2,NFYB,TBX2,TWIST1,EGR2,NEUROD4,SNAI1,H3-3B,TBX3,RORA,SMAD2,TEF,RB1,EGFR,APP,ZNF267,EZH2,E2F7,ZBTB7A,FOXK1,SP1,TRIM24,BCL11B,PAX9,KDM5B,NR1I3,HEY2,DR1,RUNX2,YBX1,PBX3,BMI1,KLF12,FOXO1,RREB1,E2F6,TCF4,ZNF24,ESRRA,MITF,MECP2,NCOA2,TFAM,MEF2A,HNRNPDL,BCL11A

Transcription factor binding	27	587	0.61	2.11E-06	NFYB,TBX2,TWIST1,EGR2,TBX3,RORA,SMAD2,RB1,CTBP1,EZH2,ZBTB7A,SP1,MTDH,TRIM24,HEY2,DR1,RUNX2,NCOA3,SRC,YWHAZ,TCF4,KDM1A,NCOA2,HDAC9,TFAM,MEF2A,BCL11A
Enzyme binding	55	2084	0.37	2.11E-06	MAPK1,AURKA,TNFAIP1,NR3C1,MSH2,TBX2,EGR2,SNAI1,SORL1,CDH1,SMAD2,CSE1L,NOX4,RB1,PXN,EGFR,APP,KIT,FMNL2,SLX4,DUSP16,CASP3,PAK2,MET,ZBTB7A,SP1,PTN,RAC1,PTGS2,HEY2,NCOA3,KDM4A,YBX1,SRC,FGR,CUL4A,FOXO1,SLC22A18,DAAM1,YWHAZ,NSF,KDM1A,SLC12A5,NBEAL1,ACVR1,HDAC9,WBP1L,FNIP1,ATG13,MEF2A,EPS8,BCL11A,MYO1C,CDC42,NOTCH1
Sequence-specific double-stranded DNA binding	45	1564	0.41	5.79E-06	HOXC8,NR3C1,MSH2,NFYB,TBX2,TWIST1,EGR2,NEUROD4,SNAI1,H3-3B,TBX3,RORA,SMAD2,TEF,RB1,APP,ZNF267,EZH2,E2F7,ZBTB7A,FOXK1,SP1,TRIM24,BCL11B,PAX9,KDM5B,NR1I3,HEY2,DR1,RUNX2,YBX1,PBX3,BMI1,KLF12,FOXO1,RREB1,E2F6,TCF4,ZNF24,ESRRA,MITF,NCOA2,TFAM,MEF2A,BCL11A
RNA polymerase II cis-regulatory region sequence-specific DNA binding	38	1196	0.45	7.36E-06	NR3C1,NFYB,TBX2,TWIST1,EGR2,NEUROD4,SNAI1,H3-3B,TBX3,RORA,SMAD2,TEF,APP,ZNF267,EZH2,E2F7,ZBTB7A,FOXK1,SP1,TRIM24,BCL11B,PAX9,NR1I3,HEY2,RUNX2,PBX3,BMI1,KLF12,FOXO1,RREB1,E2F6,TCF4,ZNF24,ESRRA,MITF,NCOA2,MEF2A,BCL11A
Sequence-specific DNA binding	46	1666	0.39	9.81E-06	HOXC8,NR3C1,MSH2,NFYB,TBX2,TWIST1,EGR2,NEUROD4,SNAI1,H3-3B,TBX3,RORA,SMAD2,TEF,RB1,APP,ZNF267,EZH2,E2F7,ZBTB7A,FOXK1,SP1,TRIM24,BCL11B,PAX9,KDM5B,NR1I3,HEY2,DR1,RUNX2,YBX1,PBX3,BMI1,KLF12,FOX

					O1,RREB1,E2F6,TCF4,ZNF24,ESRRA,MITF,MECP2,NCOA2,TFAM,MEF2A,BCL11A
Transcription regulator activity	50	1931	0.36	1.23E-05	HOXC8,NR3C1,NFYB,TBX2,TWIST1,EGR2,NEUROD4,SNAI1,TBX3,RORA,SMAD2,TEF,RB1,PXN,CTBP1,ZNF267,EZH2,E2F7,ZBTB7A,TRIM25,FOXK1,SP1,SUPT7L,MTDH,TRIM24,TRIM13,BCL11B,PAX9,KDM5B,NR1I3,HEY2,RUNX2,NCOA3,PBX3,KLF12,FOXO1,RREB1,E2F6,TCF4,ZNF24,KDM1A,ESRRA,MITF,MECP2,NCOA2,HDAC9,TFAM,MEF2A,BCL11A,NOTCH1
Transcription cis-regulatory region binding	42	1506	0.39	2.62E-05	HOXC8,NR3C1,NFYB,TBX2,TWIST1,EGR2,NEUROD4,SNAI1,H3-3B,TBX3,RORA,SMAD2,TEF,RB1,APP,ZNF267,EZH2,E2F7,ZBTB7A,FOXK1,SP1,TRIM24,BCL11B,PAX9,NR1I3,HEY2,DR1,RUNX2,PBX3,BMI1,KLF12,FOXO1,RREB1,E2F6,TCF4,ZNF24,ESRRA,MITF,NCOA2,TFAM,MEF2A,BCL11A

Two from the above table analysed were **'Transcription factor binding'** and **'Transcription regulator activity'** were analyzed. Some major targets therein were as follows:

- **TRANSCRIPTION REGULATOR ACTIVITY :**

1. **Hoxc8** - This gene is responsible for controlling migration of cells as well as their differentiation.It is also associated with pattern formation. The overexpression of this gene in humans can be a causative reason for epithelial ovarian cancer. Its gene product might also be involved in regulating cartilage differentiation.[34]
2. **EGR2** -EGR2 is associated with the pathogenicity of TH17 cells during autoimmune arisen neuroinflammation. EGR2 as a transcription factor regulates the migration and pathogenic behavior of the TH17 cells in the CNS.In addition to that, EGR2 also contributes to sustaining anti-tumor responses in exhausted CD8 + T cells.[35]
3. **KLF12** - Also known as activator protein-2 alpha, KLF12 is a transcriptional factor which is regulated developmentally. It has an important involvement in gene expression regulation in the process of vertebrate development as well as oncogenesis. It promotes the advance of breast cancer by negatively regulating p21 transcription in both p53-dependent and p53-independent manner.[36]

- **TRANSCRIPTION FACTOR BINDING :**

1. **RB1 -** The *RB1* gene is responsible for production of the pRB protein. pRB is a tumor suppressor, and functions by regulating cell growth and preventing fast paced or uncontrolled cell division. It can also prevent triggering of DNA replication by some proteins, this further regulates cell division and stops tumor growth.In addition to these functions, pRB can influence apoptosis as well as survival of cells.[37]
2. **TBX2 -** It is a member of the TBX subfamily in the T-box transcription factors. It possesses multiple important functions such as in controlling cell proliferation, programmed cell death, cell expansion etc. Additionally, TBX2 contributes to preservation of mitochondrial function as well as cholesterol synthesis in bovine cumulus cells.[38]

The table for **diseases** also showed multiple diseases. First ten were selected. (Table below)

TABLE 10 : STRING analysis for DISEASES

Term description	Observed gene count	Background gene count	Strength	False discovery rate	Matching proteins in your network (labels)
Disease of anatomical entity	97	4798	0.25	1.38E-06	AQP2,MMP2,CCL2,AFM,NR3C1,MSH2,TWIST1,EGR2,H3-3B,CHRM3,SORL1,LRP6,RORA,MIB1,CDH1,SMAD2,NOX4,TNC,CDK6,CACNA1C,CDK2,RB1,EGFR,APP,ADAMTS5,WNT7A,KIT,HCRT,SLX4,SNRK,EDIL3,TLR3,AXL,AGXT,CPLX1,EFEMP2,CASP3,MET,EZH2,ZBTB7A,GRIN2A,PLS1,PRKAA1,ATG7,RAC1,BCL11B,NF1,PCMTD1,MT-CO2,PAX9,TGFB2,KISS1,ATG5,DR1,RUNX2,COL5A1,NCOA3,MMP9,FOXO1,TRPC3,RREB1,SLC22A18,SLC39A14,CIDEC,AKT2,DBN1,DNASE1L3,VKORC1,CXCL12,BMP7,YWHAZ,NSF,TCF4,TNFSF11,IL6,SLC12A5,MITF,MECP2,NCOA2,SNRPE,TGFA,ACVR1,OXR1,SFRP4,BDNF,FNIP1,DAG1,SERPINA3,HNRNPDL,T

					ENT4A,TINCR,EPS8,RPL9,BCL11A,AP5Z1,NOTCH1,GREM1
Disease	115	6291	0.21	2.05E-06	AQP2,AURKA,MMP2,CCL2,AFM,NR3C1,MSH2,TBX2,TWIST1,EGR2,H3-3B,CHRM3,TBX3,SORL1,LRP6,RORA,MIB1,CDH1,SMAD2,CSE1L,NOX4,TNC,CDK6,CACNA1C,CDK2,RB1,SFXN5,EGFR,APP,ADAMTS5,WNT7A,KIT,HCRT,SLX4,SNRK,EDIL3,TLR3,AXL,AGXT,ZNF804A,CPLX1,EFEMP2,CASP3,MET,CHDH,EZH2,COQ7,ZBTB7A,GRIN2A,KLHL15,PLS1,TRIM24,SHANK2,PRKAA1,ATG7,RAC1,BCL11B,NF1,PCMTD1,MT-CO2,PAX9,TGFB2,KISS1,KDM5B,ATG5,FUNDC2,DR1,RUNX2,COL5A1,NCOA3,MMP9,SRC,FGR,BMI1,FOXO1,TRPC3,RREB1,SLC22A18,SLC39A14,CIDEC,AKT2,DBN1,DNASE1L3,VKORC1,CXCL12,BMP7,YWHAZ,NSF,TCF4,TNFSF11,IL6,SLC12A5,MITF,MECP2,NCOA2,SNRPE,TGFA,ACVR1,OXR1,SFRP4,BDNF,FNIP1,ATOX1,DAG1,SERPINA3,HNRNPDL,TENT4A,TINCR,EPS8,RPL9,BCL11A,CDC42,AP5Z1,NOTCH1,GREM1
Cancer	35	978	0.5	3.63E-06	AURKA,MSH2,H3-3B,MIB1,CDH1,CDK6,CDK2,RB1,EGFR,KIT,SNRK,EDIL3,AXL,CASP3,MET,EZH2,TRIM24,PRKAA1,RAC1,NF1,SRC,FGR,BMI1,FOXO1,RREB1,SLC22A18,AKT2,TNFSF11,IL6,MITF,NCOA2,TGFA,ACVR1,SERPINA3,NOTCH1
Disease of cellular proliferation	37	1101	0.47	4.60E-06	AURKA,AFM,MSH2,H3-3B,MIB1,CDH1,CDK6,CDK2,RB1,EGFR,KIT,SNRK,EDIL3,AXL,CASP3,MET,EZH2,TRI

					M24,PRKAA1,RAC1,NF1,SRC,FGR,BMI1,FOXO1,RREB1,SLC22A18,AKT2,TNFSF11,IL6,MITF,NCOA2,TGFA,ACVR1,SERPINA3,NOTCH1,GREM1
Organ system cancer	28	757	0.52	5.31E-05	MSH2,H3-3B,MIB1,CDH1,CDK6,CDK2,RB1,EGFR,KIT,SNRK,EDIL3,CASP3,MET,EZH2,PRKAA1,RAC1,NF1,FOXO1,RREB1,SLC22A18,TNFSF11,IL6,MITF,NCOA2,TGFA,ACVR1,SERPINA3,NOTCH1
Cell type cancer	20	451	0.59	0.00027	MSH2,H3-3B,CDH1,CDK6,RB1,EGFR,KIT,SNRK,EDIL3,MET,EZH2,NF1,FOXO1,RREB1,TNFSF11,MITF,TGFA,ACVR1,SERPINA3,NOTCH1
Connective tissue disease	26	774	0.47	0.00069	MMP2,CCL2,TWIST1,H3-3B,LRP6,ADAMTS5,WNT7A,TLR3,EFEMP2,EZH2,NF1,RUNX2,COL5A1,NCOA3,MMP9,CIDEC,AKT2,DNASE1L3,BMP7,TNFSF11,IL6,MITF,NCOA2,ACVR1,SFRP4,SERPINA3
Musculoskeletal system disease	32	1154	0.39	0.0018	MMP2,CCL2,TWIST1,EGR2,H3-3B,LRP6,ADAMTS5,WNT7A,KIT,TLR3,EFEMP2,EZH2,NF1,RUNX2,COL5A1,NCOA3,MMP9,FOXO1,CIDEC,AKT2,DNASE1L3,BMP7,TNFSF11,IL6,MITF,NCOA2,ACVR1,SFRP4,DAG1,SERPINA3,HNRNPDL,TENT4A
Gastrointestinal system cancer	11	161	0.78	0.0019	MSH2,CDH1,CDK6,RB1,EGFR,KIT,EDIL3,CASP3,MET,PRKAA1,NF1
Central nervous system disease	32	1199	0.37	0.0031	AQP2,EGR2,H3-3B,SORL1,RORA,MIB1,EGFR,APP,HCRT,CPLX1,CASP3,ZBTB7A,GRIN2A,ATG7,NF1,MT-

					CO2,ATG5,MMP9,TRPC3,SLC39A14,DBN1,YWHAZ,NSF,TCF4,IL6,SLC12A5,MECP2,ACVR1,OXR1,BDNF,SERPINA3,AP5Z1

From the above table, two analyzes were **cancer** and **disease of cellular proliferation**. Some major targets therein were as follows:

1. **CDK2-** Cyclin-dependent kinase 2 (CDK2), it is a gene that encodes a member of the serine/threonine protein kinase family which functions in cell cycle regulation. This protein is important for the progression from G1 to S phase. CDK2 interacts with and phosphorylates proteins in pathways such as DNA damage, intracellular transport, protein degradation, signal transduction, DNA and RNA metabolism, translation etc. CDK2 and its regulatory subunits are deregulated in many of the human cancers. [39,40]

2. **CASP3 -** Caspase-3 (*CASP3*) is a major mediator of apoptosis activated during the cellular exposure to cytotoxic drugs, radiotherapy, or immunotherapy. It is very often used as a marker for efficacy of cancer therapy. It also has non-apoptotic roles such as promotion of tumor relapse and tumor angiogenesis etc. [41]
3. **RAC1-** Ras-related C3 botulinum toxin substrate 1 (Rac1) it is a member of the Rac family of guanosine triphosphate phosphohydrolases (GTPase), it is involved in number of functions like cellular plasticity, migration and invasion, cellular adhesions, cell proliferation, apoptosis, reactive oxygen species (ROS) production and in inflammatory responses. [42]
4. **AKT2-** serine/threonine kinase Akt2, is a known proto-oncogene whose elevated expression positively correlates with aggressiveness of cancer and poor survival rates. Amplification and overexpression of Akt2 is frequently observed in a number of human tumors, including prostate, ovarian, breast and pancreatic etc. [43]
5. **KIT**- c-Kit, is a receptor tyrosine kinase, involved in intracellular signaling, its mutated form plays a crucial role in occurrence of some of the cancers. It has led to the concept that inhibiting c-Kit kinase activity can be used as a target for cancer therapy. The promising results of the inhibition of c-Kit for treatment of cancers have been observed in some of the cancers which includes gastrointestinal stromal tumor, acute myeloid leukemia, melanoma, and etc., and these results have encouraged attempts towards improvement of using c-Kit as a capable target for cancer therapy. [44]

The **table for biological processes**, too, showed multiple processes. List of first ten targets is given below.

TABLE 11 : STRING analysis for BIOLOGICAL PROCESSES

Term description	Observed gene count	Background gene count	Strength	False discovery rate	Matching proteins in your network (labels)
Positive regulation of metabolic process	114	3847	0.42	1.33E-21	MAPK1,AURKA,CCL2,TNFAIP1,NR3C1,MSH2,NFYB,TBX2,TWIST1,EGR2,SNAI1,TBX3,SORL1,LRP6,RORA,CDH1,SMAD2,PIK3R3,NOX4,TNC,CDK6,TEF,CDK2,RB1,EGFR,APP,WNT7A,KIT,CTBP1,HCRT,SLX4,TLR3,ZNF267,AXL,PAIP1,ZNF804A,CASP3,PAK2,MET,EZH2,COQ7,E2F7,TRIM25,FOXK1,SP1,GRIN2A,SUPT7L,MTDH,TRIM24,PTGES2,PRKAA1,ATG7,GPBP1L1,TRIM13,RAC1,PTBP1,BCL11B,MT-CO2,PAX9,TGFB2,KDM5B,PTGS2,NR1I3,HEY2,DR1,RUNX2,NCOA3,MMP9,KDM4A,YBX1,PBX3,SRC,FGR,BMI1,KLF12,LGR4,FOXO1,RREB1,BNIP3L,AKT2,BMP7,NSF,TCF4,TNFSF11,ZNF24,FKBP1A,KDM1A,ESRRA,IL6,GRAMD4,MITF,MECP2,NCOA2,TGFA,ACVR1,SFRP4,BDNF,WBP1L,TFAM,FNIP1,LIMCH1,ATG13,PAFAH1B2,DAG1,TRIAP1,MEF2A,CNN2,BCL2L13,TENT4A,BCL11A,MYO1C,CDC42,NOTCH1,GREM1
Positive regulation of biological process	146	6207	0.32	1.95E-21	MAPK1,AURKA,MMP2,CCL2,TNFAIP1,NR3C1,MSH2,IGFBP5,NFYB,TBX2,TWIST1,EGR2,NEUROD4,SNAI1,H3-3B,CHRM3,TBX3,SORL1,LRP6,RORA,MIB1,CDH1,SMAD2,PIK3R3,NOX4,TNC,CDK6,TEF,CDK2,RB1,PXN,EGFR

					,APP,WNT7A,WIF1,KIT,STC1,CTBP1,HCRT,SLX4,EDIL3,TLR3,ZNF267,AXL,PAIP1,ZNF804A,EFEMP2,CASP3,PAK2,MET,EZH2,COQ7,E2F7,TRIM25,FOXK1,SP1,GRIN2A,SUPT7L,PLS1,MTDH,TRIM24,PTN,SHANK2,PTGES2,PRKAA1,ATG7,GPBP1L1,TRIM13,RAC1,PTBP1,BCL11B,NF1,UBE2Z,MT-CO2,PAX9,TGFB2,KISS1,KDM5B,PTGS2,NR1I3,HEY2,MCL1,ATG5,WNT2B,DR1,RUNX2,NCOA3,MMP9,KDM4A,YBX1,PBX3,SRC,FGR,CUL4A,BMI1,KLF12,LGR4,FOXO1,TRPC3,RREB1,LHFPL2,BNIP3L,SLC39A14,AKT2,DBN1,VKORC1,CXCL12,BMP7,NSF,TCF4,TNFSF11,ZNF24,FKBP1A,KDM1A,ESRRA,IL6,GRAMD4,MITF,MECP2,NCOA2,TGFA,ACVR1,RASGRF1,HDAC9,SFRP4,BDNF,WBP1L,TFAM,FNIP1,LIMCH1,PTP4A3,ATG13,PAFAH1B2,DAG1,TRIAP1,MEF2A,CNN2,BCL2L13,TENT4A,EFNB2,EPS8,BCL11A,MYO1C,CDC42,NOTCH1,GREM1
Positive regulation of macromolecule metabolic process	107	3533	0.43	8.70E-21	MAPK1,AURKA,CCL2,TNFAIP1,NR3C1,MSH2,NFYB,TBX2,TWIST1,EGR2,SNAI1,TBX3,SORL1,LRP6,RORA,CDH1,SMAD2,PIK3R3,NOX4,TNC,CDK6,TEF,CDK2,RB1,EGFR,APP,WNT7A,KIT,CTBP1,SLX4,TLR3,ZNF267,PAIP1,ZNF804A,CASP3,PAK2,MET,EZH2,COQ7,E2F7,TRIM25,FOXK1,SP1,GRIN2A,SUPT7L,MTDH,TRIM24,PTGES2,PRKAA1,ATG7,GPBP1L1,TRIM13,RAC1,PTBP1,BCL11B,PAX9,TGFB2,KDM5B,PTGS2,NR1I3,HEY2,DR1,RUNX2,NCOA3,MMP9,KDM4A,YBX1,PBX3,SRC,FGR,BMI1,KLF12,

					LGR4,FOXO1,RREB1,AKT2,BMP7,NSF,TCF4,TNFSF11,ZNF24,FKBP1A,KDM1A,ESRRA,IL6,GRAMD4,MITF,MECP2,NCOA2,TGFA,ACVR1,SFRP4,BDNF,WBP1L,TFAM,FNIP1,LIMCH1,DAG1,TRIAP1,MEF2A,CNN2,BCL2L13,TENT4A,BCL11A,MYO1C,NOTCH1,GREM1
Positive regulation of cellular process	136	5584	0.33	1.58E-20	MAPK1,AURKA,MMP2,CCL2,TNFAIP1,NR3C1,MSH2,IGFBP5,NFYB,TBX2,TWIST1,EGR2,NEUROD4,SNAI1,H3-3B,TBX3,SORL1,LRP6,RORA,MIB1,CDH1,SMAD2,PIK3R3,NOX4,TNC,CDK6,TEF,CDK2,RB1,PXN,EGFR,APP,WNT7A,WIF1,KIT,CTBP1,HCRT,SLX4,EDIL3,TLR3,ZNF267,AXL,PAIP1,ZNF804A,EFEMP2,CASP3,PAK2,MET,EZH2,COQ7,E2F7,TRIM25,FOXK1,SP1,GRIN2A,SUPT7L,PLS1,MTDH,TRIM24,PTN,SHANK2,PTGES2,PRKAA1,ATG7,GPBP1L1,TRIM13,RAC1,PTBP1,BCL11B,NF1,UBE2Z,MT-CO2,PAX9,TGFB2,KISS1,KDM5B,PTGS2,NR1I3,HEY2,MCL1,DR1,RUNX2,NCOA3,MMP9,KDM4A,YBX1,PBX3,SRC,FGR,CUL4A,BMI1,KLF12,LGR4,FOXO1,TRPC3,RREB1,BNIP3L,SLC39A14,AKT2,DBN1,CXCL12,BMP7,NSF,TCF4,TNFSF11,ZNF24,FKBP1A,KDM1A,ESRRA,IL6,MITF,MECP2,NCOA2,TGFA,ACVR1,RASGRF1,HDAC9,SFRP4,BDNF,TFAM,FNIP1,LIMCH1,PTP4A3,ATG13,PAFAH1B2,DAG1,TRIAP1,MEF2A,TENT4A,EFNB2,EPS8,BCL11A,MYO1C,CDC42,NOTCH1,GREM1
Negative regulation of	124	4736	0.37	2.26E-20	USP28,HOXC8,MAPK1,AURKA,MMP2,CCL2,TNFAIP1,N

cellular process					R3C1,MSH2,IGFBP5,DUSP4,TBX2,TWIST1,SNAI1,H3-3B,TBX3,SORL1,LRP6,RORA,MIB1,CDH1,SMAD2,PIK3R3,NOX4,CDK6,CDK2,RB1,EGFR,CXADR,APP,WNT7A,WIF1,KIT,STC1,CTBP1,HCRT,SLX4,FSTL1,TLR3,DUSP16,AXL,PAIP1,EFEMP2,CASP3,PAK2,MET,EZH2,COQ7,E2F7,ZBTB7A,FOXK1,KLHL15,MTDH,TRIM24,PTN,SHANK2,PRKAA1,ATG7,RAC1,PTBP1,BCL11B,NF1,DCLK1,UBE2Z,PAX9,TGFB2,KISS1,KDM5B,PTGS2,NR1I3,HEY2,MCL1,ATG5,DR1,RUNX2,COL5A1,MMP9,KDM4A,YBX1,GLO1,SRC,FGR,CUL4A,BMI1,KLF12,LGR4,FOXO1,RREB1,BNIP3L,E2F6,AKT2,CXCL12,BMP7,YWHAZ,TNFSF11,ZNF24,FKBP1A,KDM1A,IL6,GRAMD4,MITF,MECP2,NCOA2,TGFA,ACVR1,OXR1,HDAC9,SFRP4,BDNF,PAQR3,FNIP1,LIMCH1,ATG13,DAG1,TRIAP1,MEF2A,CNN2,TENT4A,EFNB2,EPS8,BCL11A,CDC42,NOTCH1,GREM1
Positive regulation of nitrogen compound metabolic process	99	3166	0.44	7.90E-20	MAPK1,AURKA,TNFAIP1,NR3C1,MSH2,NFYB,TBX2,TWIST1,EGR2,SNAI1,TBX3,SORL1,LRP6,RORA,CDH1,SMAD2,PIK3R3,NOX4,TEF,CDK2,RB1,EGFR,APP,WNT7A,KIT,CTBP1,SLX4,TLR3,ZNF267,PAIP1,CASP3,PAK2,MET,EZH2,COQ7,E2F7,TRIM25,FOXK1,SP1,GRIN2A,SUPT7L,MTDH,TRIM24,PTGES2,PRKAA1,ATG7,GPBP1L1,TRIM13,RAC1,PTBP1,BCL11B,MT-CO2,PAX9,TGFB2,PTGS2,NR1I3,HEY2,DR1,RUNX2,NCOA3,MMP9,YBX1,PBX3,SRC,BMI1,KLF12,LGR4,FOXO1,RREB1,AKT2,BMP7,NSF,T

					CF4,TNFSF11,ZNF24,FKBP1A,KDM1A,ESRRA,IL6,GRAMD4,MITF,MECP2,NCOA2,TGFA,ACVR1,BDNF,WBP1L,TFAM,FNIP1,LIMCH1,DAG1,TRIAP1,MEF2A,BCL2L13,TENT4A,BCL11A,MYO1C,NOTCH1,GREM1
Positive regulation of cellular metabolic process	97	3114	0.44	3.37E-19	MAPK1,AURKA,TNFAIP1,NR3C1,MSH2,NFYB,TBX2,TWIST1,EGR2,SNAI1,TBX3,LRP6,RORA,CDH1,SMAD2,PIK3R3,NOX4,TEF,CDK2,RB1,EGFR,APP,WNT7A,KIT,CTBP1,SLX4,TLR3,ZNF267,AXL,PAIP1,CASP3,PAK2,MET,EZH2,COQ7,E2F7,TRIM25,FOXK1,SP1,SUPT7L,MTDH,TRIM24,PTGES2,PRKAA1,GPBP1L1,TRIM13,RAC1,PTBP1,BCL11B,MT-CO2,PAX9,TGFB2,PTGS2,NR1I3,HEY2,DR1,RUNX2,NCOA3,MMP9,YBX1,PBX3,SRC,FGR,KLF12,LGR4,FOXO1,RREB1,BNIP3L,AKT2,BMP7,NSF,TCF4,TNFSF11,ZNF24,KDM1A,ESRRA,IL6,MITF,MECP2,NCOA2,TGFA,ACVR1,BDNF,TFAM,FNIP1,LIMCH1,ATG13,PAFAH1B2,DAG1,TRIAP1,MEF2A,TENT4A,BCL11A,MYO1C,CDC42,NOTCH1,GREM1
Animal organ development	99	3246	0.43	3.76E-19	HOXC8,AQP2,MAPK1,AURKA,MMP2,CCL2,NR3C1,MSH2,IGFBP5,TBX2,TWIST1,EGR2,NEUROD4,SNAI1,H3-3B,TBX3,LRP6,RORA,MIB1,BTBD1,CDH1,SMAD2,PIK3R3,NOX4,TNC,CDK6,CACNA1C,RB1,EGFR,CXADR,APP,ADAMTS5,WNT7A,KIT,STC1,SNRK,TLR3,AXL,EFEMP2,CASP3,MET,EZH2,E2F7,ZBTB7A,FOXK1,GRIN2A,PLS1,PTN,SHANK2,BCL11B,NF1,DCLK1,MT-

					CO2,PAX9,TGFB2,KDM5B,PTGS2,HEY2,ATG5,WNT2B,RUNX2,COL5A1,MMP9,GLO1,PBX3,SRC,FGR,CUL4A,BMI1,LGR4,SLC39A14,ARHGAP5,AKT2,VKORC1,CXCL12,BMP7,YWHAZ,TNFSF11,FKBP1A,KDM1A,ESRRA,IL6,MITF,MECP2,TGFA,ACVR1,HDAC9,SFRP4,WBP1L,FNIP1,PAFAH1B2,DAG1,MEF2A,CNN2,CSMD1,EFNB2,CDC42,NOTCH1,GREM1
Negative regulation of biological process	129	5313	0.33	6.63E-19	USP28,HOXC8,MAPK1,AURKA,MMP2,CCL2,TNFAIP1,NR3C1,MSH2,IGFBP5,DUSP4,TBX2,TWIST1,SNAI1,H3-3B,TBX3,SORL1,LRP6,RORA,MIB1,CDH1,SMAD2,PIK3R3,NOX4,CDK6,CDK2,RB1,EGFR,CXADR,APP,ADAMTS5,WNT7A,WIF1,KIT,STC1,CTBP1,HCRT,SLX4,FSTL1,TLR3,DUSP16,AXL,PAIP1,EFEMP2,CASP3,PAK2,MET,EZH2,COQ7,E2F7,ZBTB7A,TRIM25,FOXK1,GRIN2A,KLHL15,MTDH,TRIM24,PTN,SHANK2,PRKAA1,ATG7,TRIM13,RAC1,PTBP1,BCL11B,NF1,DCLK1,UBE2Z,PAX9,TGFB2,KISS1,KDM5B,PTGS2,NR1I3,HEY2,MCL1,ATG5,DR1,RUNX2,COL5A1,MMP9,KDM4A,YBX1,GLO1,SRC,FGR,CUL4A,BMI1,KLF12,LGR4,FOXO1,RREB1,BNIP3L,E2F6,AKT2,CXCL12,BMP7,YWHAZ,TNFSF11,ZNF24,FKBP1A,KDM1A,IL6,GRAMD4,MITF,MECP2,NCOA2,TGFA,ACVR1,OXR1,HDAC9,SFRP4,BDNF,PAQR3,FNIP1,LIMCH1,ATG13,DAG1,TRIAP1,SERPINA3,MEF2A,CNN2,TENT4A,EFNB2,EPS8,BCL11A,CDC42,NOTCH1,GREM1
System	105	3867	0.38	4.63E-17	HOXC8,AQP2,MAPK1,AUR

development					KA,MMP2,CCL2,NR3C1,MSH2,IGFBP5,TBX2,TWIST1,EGR2,NEUROD4,SNAI1,H3-3B,CHRM3,MSI1,TBX3,LRP6,RORA,MIB1,BTBD1,CDH1,SMAD2,PIK3R3,NOX4,TNC,CDK6,CACNA1C,RB1,EGFR,CXADR,APP,WNT7A,KIT,STC1,SNRK,TLR3,AXL,EFEMP2,CASP3,PAK2,MET,EZH2,E2F7,ZBTB7A,GRIN2A,PLS1,PTN,SHANK2,RAC1,BCL11B,NF1,DCLK1,TGFB2,KDM5B,PTGS2,HEY2,ATG5,WNT2B,RUNX2,COL5A1,MMP9,GLO1,PBX3,SRC,FGR,CUL4A,BMI1,LGR4,FOXO1,SLC39A14,AKT2,DBN1,VKORC1,CXCL12,BMP7,YWHAZ,TCF4,TNFSF11,ZNF24,FKBP1A,KDM1A,ESRRA,IL6,SLC12A5,MITF,MECP2,TGFA,ACVR1,RASGRF1,HDAC9,SFRP4,BDNF,WBP1L,FNIP1,PAFAH1B2,DAG1,MEF2A,CNN2,CSMD1,EFNB2,CDC42,NOTCH1,GREM1

Two categories were analyzed from the table above, they are **positive regulation of metabolic process and that of cellular process.** Some major targets therein are as follows:

- **POSITIVE REGULATION OF METABOLIC PROCESS**

1. **KDM1A :** KDM1A is known to play a role in cancer stem cell self-renewal, cell proliferation and differentiation, as well as epithelial-mesenchymal transition and in metastasis. KDM1A is found to be overexpressed in several cancers, where it is known to inhibit cell differentiation and on the other hand, enhance cell proliferation and aggressiveness [45,46].
2. **TWIST1 :** It helps to promote tumor invasion as well as metastases by promoting EMT. It is done by turning-down the expression of epithelial specific proteins (E-cadherin) and by upregulating the expression of mesenchymal markers (N-cadherin, the vimentin and the smooth-muscle actin) [47,48].

- **POSITIVE REGULATION OF CELLULAR PROCESS**

1. **MSH2 :** This gene is responsible for production of a protein that plays a role in DNA repair. Defect in this gene affects the production of this protein and thereby increases the

risk of getting cancer. Many cancers are hence related to defective MSH2 gene. Functional MSH2 is a classical tumor suppressor [49,50].

2. **AXL :** This is a tyrosine kinase receptor which also acts as an oncogene (promotes cancer development - increases proliferation, survival, invasion, and migration in cancer cells). It also leads to the development of resistance to different types of therapies such as chemotherapy, radiotherapy etc. [51]
3. **CCL2 :** Studies have shown that it can activate tumor cell growth and proliferation. On interaction with CCR2, CCL2 promotes the migration of cancer cells and it also recruits immunosuppressive cells to the tumor microenvironment, favoring development of many cancers [52].

5.0 CONCLUSION:

It is thus seen that miR-137 plays a pivotal role in multiple diseases by regulating a number of gene targets. As inferred in the case of cancer , it is mostly seen to act as a tumor suppressor.
Through our study, we could find out about multiple experimentally verified direct targets of miR-137 and thus created a miR-137 targets database. We also analyzed their biological and molecular roles. This further help us understand the pathway of action of miRNA regulation in each of these diseases.

In conclusion, understanding the regulation of microRNA-137 can definitely help us find diagnostic and therapeutic pathways for numerous associated diseases and also aid in understanding the miRNA regulatory network. The data incorporating validated miRNA targets is extremely important in order to understand their action as well as to further develop the basis of miRNA target identification, considering the fact that more and more prediction tools are now using and learning from confidently validated miRNA-mRNA interactions.

6.0 REFERENCES :

1) Cai, Y., Yu, X., Hu, S., & Yu, J. (2009). A brief review on the mechanisms of MIRNA regulation. *Genomics, Proteomics & Bioinformatics*, *7*(4), 147–154. https://doi.org/10.1016/s1672-0229(08)60044-3
2) O'Brien, J. A., Hayder, H., Zayed, Y., & Peng, C. (2018). Overview of MicroRNA biogenesis, mechanisms of actions, and circulation. *Frontiers in Endocrinology*, *9*. https://doi.org/10.3389/fendo.2018.00402
3) Mohr, A. M., & Mott, J. L. (2015). Overview of MicroRNA biology. *Seminars in Liver Disease*, *35*(01), 003–011. https://doi.org/10.1055/s-0034-1397344
4) Ranganathan, K., & Sivasankar, V. (2014). MicroRNAs - Biology and clinical applications. *Journal of Oral and Maxillofacial Pathology*, *18*(2), 229. https://doi.org/10.4103/0973-029x.140762
5) Ardekani, A. M. (2010, December 1). *The role of MicroRNAs in human diseases*. PubMed Central (PMC). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3558168/#S0004title
6) Yu, L., & Kowdley, K. V. (2012). MicroRNAs in common human diseases. *Genomics, Proteomics & Bioinformatics*, *10*(5), 246–253. https://doi.org/10.1016/j.gpb.2012.07.005
7) Dwivedi, S., Purohit, P., & Sharma, P. (2019). MicroRNAs and Diseases: Promising biomarkers for diagnosis and therapeutics. *Indian Journal of Clinical Biochemistry*, *34*(3), 243–245. https://doi.org/10.1007/s12291-019-00844-x
8) Eisenberg, I., Eran, A., Nishino, I., Moggio, M., Lamperti, C., Amato, A. A., Lidov, H. G., Kang, P. B., North, K., Mitrani-Rosenbaum, S., Flanigan, K. M., Neely, L. A., Whitney, D., Beggs, A. H., Kohane, I. S., & Kunkel, L. M. (2007). Distinctive patterns of microRNA expression in primary muscular disorders. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(43), 17016–17021. https://doi.org/10.1073/pnas.0708115104
9) Hanahan, D., & Weinberg, R. A. (2000). The hallmarks of cancer. Cell, 100(1), 57–70.
10) Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. Cell, 144(5), 646–674.
11) Hanahan D. (2022). Hallmarks of Cancer: New Dimensions. Cancer discovery, 12(1), 31–46.
12) Kong, Y. W., Ferland-McCollough, D., Jackson, T. J., & Bushell, M. (2012). microRNAs in cancer management. *The Lancet. Oncology*, *13*(6), e249–e258. https://doi.org/10.1016/S1470-2045(12)70073-6
13) Duque-Guimarães, D. E., De Almeida-Faria, J., Ong, T. P., & Ozanne, S. E. (2016b). Pulsed SILAC as a approach for miRNA targets identification in cell culture. In *Methods in molecular biology*. https://doi.org/10.1007/978-1-4939-6730-8_11
14) Riolo, G., Cantara, S., Marzocchi, C., & Ricci, C. (2020). MiRNA targets: From prediction tools to experimental validation. *Methods and Protocols*, *4*(1), 1. https://doi.org/10.3390/mps4010001

15) Huang, T., Pinto, S. M., & Pandey, A. (2012). Proteomics for understanding miRNA biology. *Proteomics*, *13*(3–4), 558–567. https://doi.org/10.1002/pmic.201200339
16) Thomson, D., Bracken, C. P., & Goodall, G. J. (2011). Experimental strategies for microRNA target identification. *Nucleic Acids Research*, *39*(16), 6845–6853.https://doi.org/10.1093/nar/gkr330
17) https://pubmed.ncbi.nlm.nih.gov/?term=Targets+and+MicroRNA+137&filter=pubt.review&show_snippets=off&size=200 (for literature mining)
18) Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nature reviews. Genetics*, *10*(1), 57–63. https://doi.org/10.1038/nrg2484
19) Liu, W., & Wang, X. (2019). Prediction of functional microRNA targets by integrative modeling of microRNA binding and target expression data. *Genome biology*, *20*(1), 18. https://doi.org/10.1186/s13059-019-1629-z
20) https://microbenotes.com/rna-sequencing-principle-steps-types-uses/
21) https://www.technologynetworks.com/genomics/articles/rna-seq-basics-applications-and-protocol-299461#D2
22) Korth, M. J., & Law, G. L. (2016). Systems Virology. In *Elsevier eBooks* (pp. 141–156). https://doi.org/10.1016/b978-0-12-800964-2.00011-2
23) Subair, S. (2015). *Microarray Analysis and Gene Expression : A simplified Review*. https://www.semanticscholar.org/paper/Microarray-Analysis-and-Gene-Expression-%3A-A-Review-Subair-Aleisa/a75f33de77e9b7e220c102245fda6fb7bddf3245
24) https://www.genentechoncology.com/pathways/cancer-tumor-targets/mapk.html#:~:text=The%20mitogen%2Dactivated%20protein%20kinase,%2C%20cellular%20growth%2C%20and%20survival.&text=Abnormal%20MAPK%20signaling%20may%20lead,proliferation%20and%20resistance%20to%20apoptosis.&text=Research%20into%20the%20MAPK%20pathway,be%20important%20in%20some%20cancers.
25) Yao, Y., Jiao, D., Liu, Z., Chen, J., Zhou, X., Li, Z., Li, J., & Han, X. (2020). Novel miRNA Predicts Survival and Prognosis of Cholangiocarcinoma Based on RNA-seq Data and *In Vitro* Experiments. *BioMed research international*, *2020*, 5976127. https://doi.org/10.1155/2020/5976127
26) Sasaki, T., Hiroki, K., & Yamashita, Y. (2013). The role of epidermal growth factor receptor in cancer metastasis and microenvironment. BioMed research international, 2013, 546318. https://doi.org/10.1155/2013/546318
27) Gharaibeh, L., Elmadany, N., Alwosaibai, K., & Alshaer, W. (2020). Notch1 in Cancer Therapy: Possible Clinical Implications and Challenges. Molecular pharmacology, 98(5), 559–576. https://doi.org/10.1124/molpharm.120.000006
28) https://www.breastcancer.org/treatment/targeted-therapy/what-are-cdk46-inhibitors
29) https://www.ncbi.nlm.nih.gov/gene/998
30) https://www.ncbi.nlm.nih.gov/gene/7039
31) https://www.ncbi.nlm.nih.gov/gene/23493
32) https://www.genecards.org/cgi-bin/carddisp.pl?gene=TRIM4

33) https://www.ncbi.nlm.nih.gov/gene/2308#:~:text=FoxO1%20is%20a%20cell%2Dspecific,cell%20lymphoma%20lack%20transcriptional%20activity
34) Matías, F. L. &. J. (n.d.). *HOXC8 (homeobox C8)*. https://atlasgeneticsoncology.org/gene/45912/hoxc8-(homeobox-c8)#:~:text=Transcriptional%20regulation%3A%20As%20a%20transcription,et%20al.%2C%202005 , *HOXC8 homeobox C8 [Homo sapiens (human)] - Gene - NCBI*. (n.d.). https://www.ncbi.nlm.nih.gov/gene/3224
35) *Egr2 early growth response 2 [Mus musculus (house mouse)] - Gene - NCBI*. (n.d.). https://www.ncbi.nlm.nih.gov/gene/13654#:~:text=GeneRIFs%3A%20Gene%20References%20Into%20Functions,tumour%20infiltrating%20CD8%20%2B%20T%20cells
36) *KLF12 KLF transcription factor 12 [Homo sapiens (human)] - Gene - NCBI*. (n.d.). https://www.ncbi.nlm.nih.gov/gene/11278
37) *RB1 gene: MedlinePlus Genetics*. (n.d.). https://medlineplus.gov/genetics/gene/rb1/#:~:text=The%20RB1%20gene%20provides%20instructions,or%20in%20an%20uncontrolled%20way.
38) Li, S., Jiang, H., Liu, Z., Yu, W., Cai, X., Liu, C., Xie, W., Quan, F., Gao, W., Kim, N., Yuan, B., Chen, C., & Zhang, J. (2022). *TBX2* affects proliferation, apoptosis and cholesterol generation by regulating mitochondrial function and autophagy in bovine cumulus cell. *Veterinary Medicine and Science*, *9*(1), 326–335. https://doi.org/10.1002/vms3.1009
39) *CDK2 - My Cancer Genome*. (n.d.). https://www.mycancergenome.org/content/gene/cdk2/#:~:text=CDK2%20is%20altered%20in%200.16,prevalence%20of%20alterations%20%5B3%5D.
40) Tadesse, S., Anshabo, A. T., Portman, N., Lim, E., Tilley, W. D., Caldon, C. E., & Wang, S. (2020). Targeting CDK2 in cancer: challenges and opportunities for therapy. *Drug Discovery Today*, *25*(2), 406–413. https://doi.org/10.1016/j.drudis.2019.12.001
41) Zhou, M., Liu, X., Li, Z., Huang, Q., Li, F., & Li, C. Y. (2018). Caspase-3 regulates the migration, invasion and metastasis of colon cancer cells. *International Journal of Cancer*, *143*(4), 921–930. https://doi.org/10.1002/ijc.31374
42) Marei, H., & Malliri, A. (2016). Rac1 in human diseases: The therapeutic potential of targeting Rac1 signaling regulatory mechanisms. *Small GTPases*, *8*(3), 139–163. https://doi.org/10.1080/21541248.2016.1211398
43) Rychahou, P. G., Kang, J., Gulhati, P., Doan, H. Q., Chen, L. A., Xiao, S. Y., Chung, D. H., & Evers, B. M. (2008). Akt2 overexpression plays a critical role in the establishment of colorectal cancer metastasis. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(51), 20315–20320. https://doi.org/10.1073/pnas.0810715105

44) Babaei, M., Kamalidehghan, B., Saleem, M., Huri, H. Z., & Ahmadipour, F. (2016). Receptor tyrosine kinase (c-Kit) inhibitors: a potential therapeutic target in cancer cells. *Drug Design Development and Therapy*, *Volume 10*, 2443–2459. https://doi.org/10.2147/dddt.s89114
45) Antona, A., Leo, G., Favero, F., Varalda, M., Venetucci, J., Faletti, S., Todaro, M., Mazzucco, E., Soligo, E., Saglietti, C., Stassi, G., Manfredi, M., Pelicci, G., Corà, D., Valente, G., & Capello, D. (2023). Targeting lysine-specific demethylase 1 (KDM1A/LSD1) impairs colorectal cancer tumorigenesis by affecting cancer cells stemness, motility, and differentiation. *Cell death discovery*, *9*(1), 201. https://doi.org/10.1038/s41420-023-01502-1
46) Jiang, Q., Stachelscheid, J., Bloehdorn, J., Pacholewska, A., Aszyk, C., Grotenhuijs, F., Müller, T., Onder, O., Wagle, P., Herling, C. D., Kleppe, M., Wang, Z., Coombes, K. R., Robrecht, S., Dalvi, P. S., Plosnita, B., Mayer, P., Abruzzo, L. V., Altmüller, J., Gathof, B., … Herling, M. (2023). Oncogenic role and target properties of the lysine-specific demethylase KDM1A in chronic lymphocytic leukemia. *Blood*, *142*(1), 44–61. https://doi.org/10.1182/blood.2022017230)
47) Jérémy Bastid, Claire Ciancia, Alain Puisieux, Stéphane Ansieau (December 2009). Role of TWIST proteins in cancer progression. Atlas of Genetics and cytogenetics in oncology and hematology. https://atlasgeneticsoncology.org/deep-insight/20081/role-of-twist-proteins-in-cancer-progression
48) Liu, T., Zhao, X., Zheng, X., Zheng, Y., Dong, X., Zhao, N., Liao, S., & Sun, B. (2020). The EMT transcription factor, Twist1, as a novel therapeutic target for pulmonary sarcomatoid carcinomas. *International journal of oncology*, *56*(3), 750–760. https://doi.org/10.3892/ijo.2020.4972
49) https://medlineplus.gov/genetics/gene/msh2/#:~:text=The%20MSH2%20gene%20provides%20instructions,in%20preparation%20for%20cell%20division.
50) Martin, S. A., McCabe, N., Mullarkey, M., Cummins, R., Burgess, D. J., Nakabeppu, Y., Oka, S., Kay, E., Lord, C. J., & Ashworth, A. (2010). DNA polymerases as potential therapeutic targets for cancers deficient in the DNA mismatch repair proteins MSH2 or MLH1. *Cancer cell*, *17*(3), 235–248. https://doi.org/10.1016/j.ccr.2009.12.046
51) Wium, M., Ajayi-Smith, A. F., Paccez, J. D., & Zerbini, L. F. (2021). The Role of the Receptor Tyrosine Kinase Axl in Carcinogenesis and Development of Therapeutic Resistance: An Overview of Molecular Mechanisms and Future Applications. *Cancers*, *13*(7), 1521. https://doi.org/10.3390/cancers13071521
52) Xu, M., Wang, Y., Xia, R., Wei, Y., & Wei, X. (2021). Role of the CCL2-CCR2 signaling axis in cancer: Mechanisms and therapeutic targeting. *Cell proliferation*, *54*(10), e13115. https://doi.org/10.1111/cpr.13115)
53) Mahmood, T., & Yang, P. C. (2012). Western blot: technique, theory, and trouble shooting. *North American journal of medical sciences*, *4*(9), 429–434. https://doi.org/10.4103/1947-2714.100998

54) Gäken, J., Mohamedali, A. M., Jiang, J., Malik, F., Stangl, D., Smith, A. E., Chronis, C., Kulasekararaj, A. G., Thomas, N. S., Farzaneh, F., Tavassoli, M., & Mufti, G. J. (2012). A functional assay for microRNA target identification and validation. *Nucleic acids research*, *40*(10), e75. https://doi.org/10.1093/nar/gks145
55) Jin, Y., Chen, Z., Liu, X., & Zhou, X. (2013). Evaluating the microRNA targeting sites by luciferase reporter gene assay. *Methods in molecular biology (Clifton, N.J.)*, *936*, 117–127. https://doi.org/10.1007/978-1-62703-083-0_10
56) Huang, Y., Qi, Y., Ruan, Q., Ma, Y., He, R., Ji, Y., & Sun, Z. (2011). A rapid method to screen putative mRNA targets of any known microRNA. *Virology journal*, *8*, 8. https://doi.org/10.1186/1743-422X-8-8
57) Mittal, N., & Zavolan, M. (2014). Seq and CLIP through the miRNA world. *Genome biology*, *15*(1), 202. https://doi.org/10.1186/gb4151
58) Tomasello, L., Cluts, L., & Croce, C. M. (2019). Experimental Validation of MicroRNA Targets: Analysis of MicroRNA Targets Through Western Blotting. *Methods in molecular biology (Clifton, N.J.)*, *1970*, 341–353. https://doi.org/10.1007/978-1-4939-9207-2_19
59) Riolo, G., Cantara, S., Marzocchi, C., & Ricci, C. (2020). miRNA Targets: From Prediction Tools to Experimental Validation. *Methods and protocols*, *4*(1), 1. https://doi.org/10.3390/mps4010001 - Table
60) Riolo, G., Cantara, S., Marzocchi, C., & Ricci, C. (2020). MiRNA targets: From prediction tools to experimental validation. *Methods and Protocols*, *4*(1), 1. https://doi.org/10.3390/mps4010001