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

SRI-VIPRA

Project Report of 2024: SVP-2414

"Oncomeric MicroRNAs and their Target Data Analysis in Glioblastoma Tumors by Bioinformatics Approach"

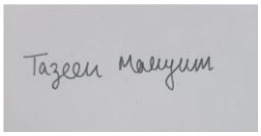
IQAC
Sri Venkateswara College
University of Delhi
Benito Juarez Road, Dhaula Kuan, New Delhi
New Delhi -110021

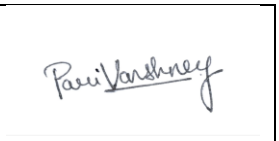
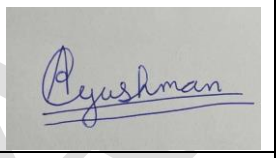
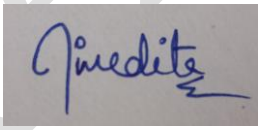
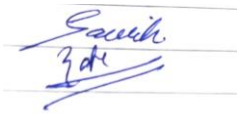
SRI-VIPRA PROJECT 2024

Title: Oncomeric MicroRNAs and their Target Data Analysis in Glioblastoma Tumors by Bioinformatics Approach

Details of the Mentors	
Dr. Ravindra Varma Polisetty **Assistant professor** **Department of Biochemistry**	**Dr. Nimisha Sinha** **Assistant professor** **Department of Biochemistry**
	

List of students under the SRI-VIPRA Project

S.No	Photo	Name of the student	Roll number	Course	Signature
1		Ananya Datta	1222001	B.Sc.(H) Biochemistry	
2		Tazeen Maryum	1222014	B.Sc.(H) Biochemistry	
3		Sanjana Kumari	1222047	B.Sc.(H) Biochemistry	

4		Pari Varshney	1222051	B.Sc.(H) Biochemistry	
5		Ayushman Saini	1223012	B.Sc.(H) Biochemistry	
6		Nivedita Kalia	1223044	B.Sc.(H) Biochemistry	
7		Gaurika Bahl	1223063	B.Sc.(H) Biochemistry	
8		Durani Sikha Kakoty	1323001	B.Sc.(H) Biological Science	
9		Samarjeetsingh Chandel	1123036	B.Sc.(P) Life Sciences	

Signature of Mentor

Signature of Mentor

Certificate of Originality

This is to certify that the aforementioned students from Sri Venkateswara College have participated in the summer project SVP-2414 titled "**Oncomeric MicroRNAs and their Target Data Analysis in Glioblastoma Tumors by Bioinformatics Approach"**. The participants have carried out the research project work under my guidance and supervision from 1st July, 2024 to 30th September 2024. The work carried out is original and carried out in an online/offline/hybrid mode.

P. Ravindra Varma

Signature of Mentor

Nimisha

Signature of Mentor

Acknowledgements

We would like to extend our deepest gratitude to Dr. Ravindra Varma and Dr. Nimisha Sinha of the Biochemistry Department at Sri Venkateswara College, University of Delhi, for providing us with the invaluable opportunity to undertake this summer internship project. Their unwavering guidance, mentorship, and support have been instrumental in our academic growth and development.

We are also profoundly thankful to the entire Biochemistry Department team for their technical expertise and unwavering assistance throughout the project. Their patience, willingness to share their knowledge, and collaborative spirit have been invaluable assets to our research endeavours.

We would also like to express our sincere regards to our esteemed principal, Prof. Vajala Ravi, for his unwavering support and provision of the necessary resources to make this research project a reality. His commitment to fostering a conducive environment for academic excellence has been instrumental in our success."

TABLE OF CONTENTS

1.0 INTRODUCTION

Hallmarks Of Cancer

The hallmarks of cancer are six biological capacities acquired during the multifaceted development of human tumors. These hallmarks serve as a guiding principle for understanding the intricate complexities of neoplastic disease. They encompass the sustaining of proliferative signalling, the evasion of growth suppressors, the resistance to cell death, the enabling of replicative immortality, the induction of angiogenesis, and the activation of invasion and metastasis. Tumors contain various recruited, seemingly normal cells that contribute to acquiring hallmark traits by creating the "tumor microenvironment."

Beneath these hallmarks lie genome instability, which generates the genetic diversity that accelerates their acquisition, and inflammation, which fosters multiple hallmark functions. In the past decade, conceptual advancements have introduced two emerging hallmarks of potential generality to this list: the reprogramming of energy metabolism and the evasion of immune destruction.

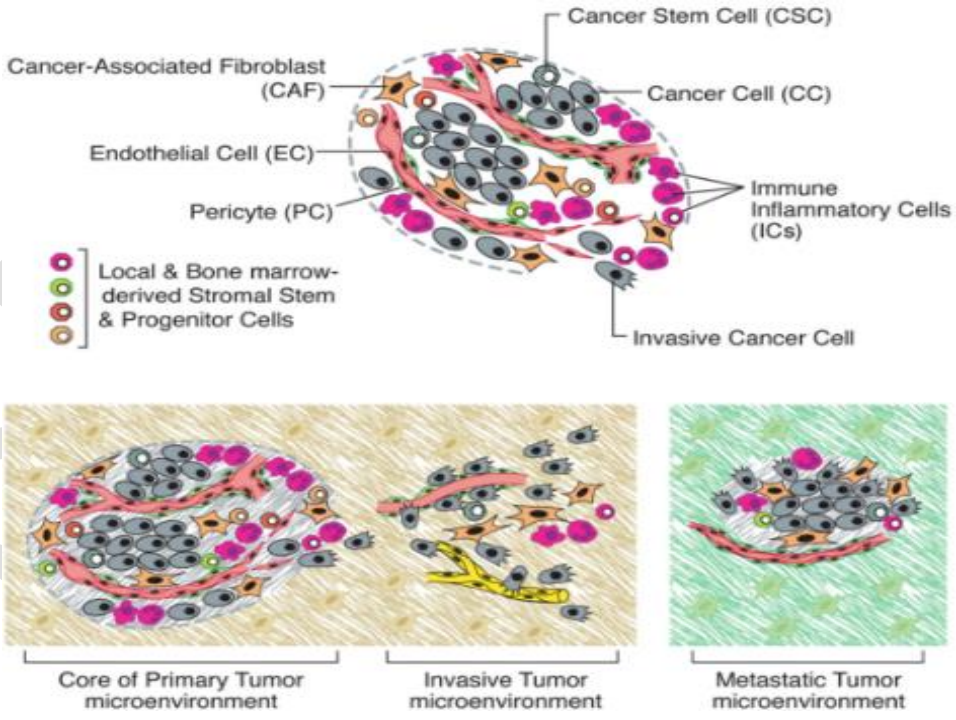


Figure-1: The Cells of the Tumor Microenvironment

- (Upper) An assemblage of distinct cell types constitutes most solid tumors. Both the parenchyma and stroma of tumors contain distinct cell types and subtypes that collectively enable tumor growth and progression. Notably, the immune inflammatory cells present in tumors can include both tumor-promoting as well as tumor-killing subclasses.

- (Lower) The distinctive microenvironments of tumors. The multiple stromal cell types create a succession of tumor microenvironments that change as tumors invade normal tissue and seed and colonize distant tissues. The abundance, histologic organization, and phenotypic characteristics of the stromal cell types and the extracellular matrix (hatched background) evolve during progression, enabling primary, invasive, and then metastatic growth. The surrounding normal cells of the primary and metastatic sites, shown only schematically, likely also affect the character of the various neoplastic microenvironments.

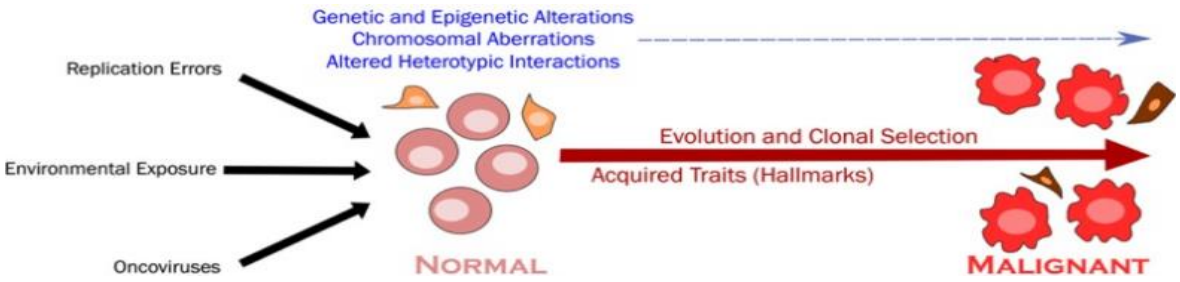


Figure-2: The transformation process

Different insults continuously act on cells leading to transformative alterations in (epi) genetics, chromosomal numbers and arrangements, and heterotypic interactions which undergo cycles of clonal selection leading to acquisition of cancer-competent traits, the hallmarks.

1. **Sustained proliferative signalling.**

   Cancer cells possess a characteristic ability to sustain chronic proliferation, enabling the cells to continuously divide. This refers to their capacity to activate signalling pathways that drive cell growth and division. Eg:- MEK-ERK Pathway, AKT/PI3K Pathway may get overstimulated in such cells, increasing gene expression and causing persistent production of proteins which can help in cell division.

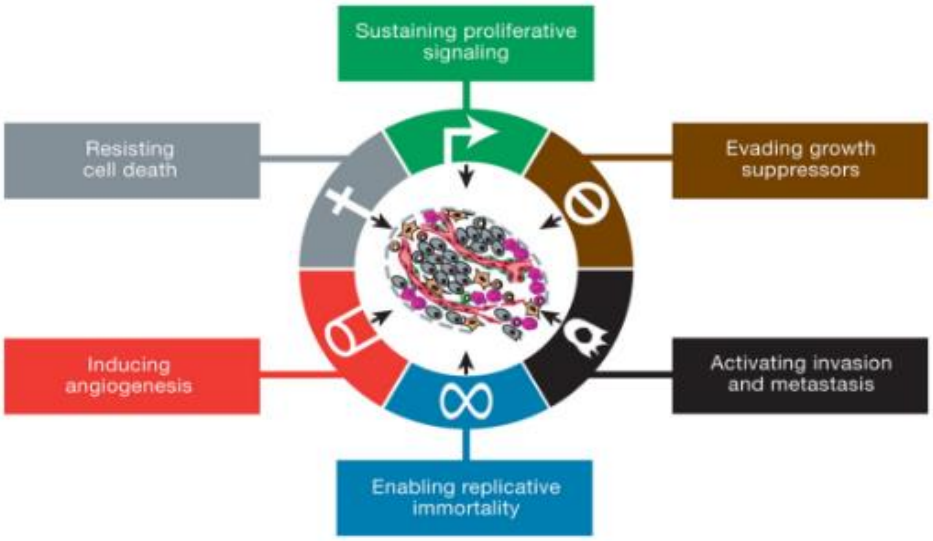


Figure-3: The six hallmarks of cancer

2. Evading growth suppressors

Growth suppressor genes typically function to inhibit cell growth and division. Several tumor-suppressive genes, encoding proteins with diverse mechanisms to inhibit cellular growth and proliferation, have been identified. The expression of genes involved in cell cycle control, apoptosis induction, and DNA repair following DNA damage. In cancer cells, these genes are often inactivated, either through mutations that impair their function or by mechanisms that prevent their expression. Example:- p53: "guardian of the genome," p53 is a crucial tumor suppressor gene. It is activated in response to cellular stress, such as DNA damage, and can trigger cell cycle arrest, DNA repair, or apoptosis (programmed cell death).

3. Resisting cell death

One of the key hallmarks of cancer cells is their ability to resist cell death. This allows them to survive and proliferate uncontrollably, leading to tumor growth and metastasis. Example:- a. Activation of Anti-Apoptotic Pathways: NF-κB: This transcription factor can activate genes that promote cell survival and inhibit apoptosis.

4. Enabling replicative immortality

This characteristic allows cancer cells to divide indefinitely, without undergoing the normal process of cellular aging. This is achieved by overcoming the "Hayflick Limit" that limits the number of times human cells can divide. Here are some ways cancer cells achieve replicative immortality: Telomere maintenance and Alternative lengthening of telomeres (ALT) pathway.

5. Inducing Angiogenesis

Cancer cells acquire the ability to induce angiogenesis, which is the formation of new blood vessels. This process not only provides the tumor with a continuous supply of oxygen and nutrients but also facilitates the removal of toxic metabolic wastes and allows the cancer to spread via the bloodstream (hematogenous metastasis).

6. Activating Invasion and Metastasis

Invasion and metastasis are a critical capability that cancer cells must acquire to progress from localized tumors to life-threatening diseases. Epithelial carcinomas of higher grades often show more invasive and metastatic behaviours, which are usually the cause of death in cancer patients rather than the primary tumor itself. The cancer cells undergo complex morphological changes and alterations in cell-cell or cell-matrix interactions that enable them to initiate the invasion-metastasis cascade. The cancer cells co-opt "Epithelial-Mesenchymal Transition" (EMT), a developmental program to facilitate invasion and metastasis. In this process, epithelial cells (which are usually stationary and form tight layers) transform into mesenchymal cells (which are more mobile and invasive). A key cell-to-cell adhesion molecule, E-cadherin, plays a major role in inhibiting invasion and metastasis. E-cadherin is frequently downregulated or inactivated in aggressive carcinomas, N-cadherin, typically found in migrating neurons and mesenchymal cells, is upregulated in invasive tumors.

EMERGING HALLMARKS: -

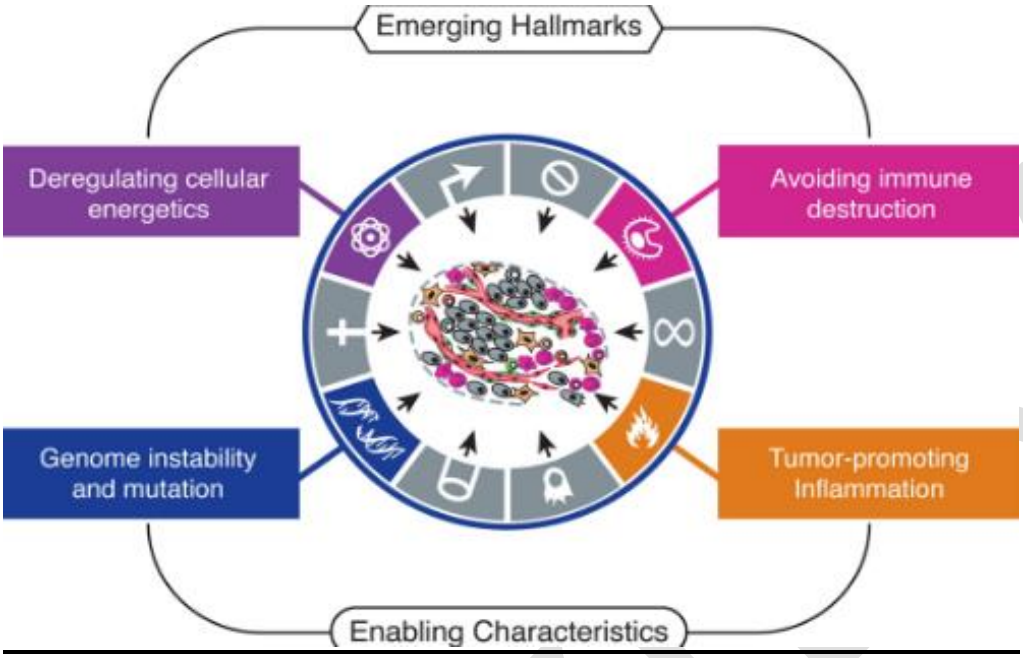


Figure-4: Emerging Hallmarks and Enabling Characteristics

1. Reprogramming Energy Metabolism:

Metabolic reprogramming is now recognized as one of the hallmarks of cancer, enabling tumor cells to meet their increased energy and biosynthetic demand.

Cancer cells often exhibit altered metabolic pathways to support their rapid growth and proliferation. This phenomenon is known as the **Warburg effect**, where cancer cells preferentially utilize glycolysis for energy production, even in the presence of oxygen. This leads to the production of lactate and the diversion of metabolic intermediates for biosynthetic processes, which are essential for cell growth and division.

2. Evading Immune Destruction:

The immune system plays a key role in detecting and eliminating cancer cells. However, tumors can develop mechanisms to evade immune destruction, which has been proposed as an emerging hallmark of cancer. Evasion of immune attack involves:

Downregulation of tumor antigens to avoid detection by the immune system, Recruitment of immunosuppressive cells like regulatory T cells and myeloid-derived suppressor cells to the tumor microenvironment.

Metabolic competition between cancer cells and T cells in the tumor microenvironment can also impair immune function. For example, cancer cells may deplete glucose and amino acids needed to support T cell activity.

SIGNALING PATHWAYS IN CANCER

An important focus of study within cancer research is called signal transduction. All the cells in the human body need external and internal signals to tell them what to do – grow, move, secrete something, live, die, etc. These messages are transmitted via a series of proteins, in an extremely complex and carefully regulated process. In fact, misbehaviour of these very pathways is often what causes and sustains cancer. Pathways involved in cell proliferation; differentiation can be dysregulated such that it is overactivated leading to uncontrolled cell division.

Some important signalling pathways in cancer research are-

1. RAS Pathway -

The RAS/RAF/MEK/ERK pathway transduces signals from the extracellular environment to the cell nucleus, regulating cell growth, division, differentiation, cell cycle, wound healing, tissue repair, and angiogenesis.

Mutations in the RAS family, particularly in KRAS, result in its constant activation in the GTP-bound state. This leads to persistent stimulation of downstream effectors without external growth signals, promoting continuous cell proliferation. It also creates a tumor microenvironment that supports tumor progression and immune escape.

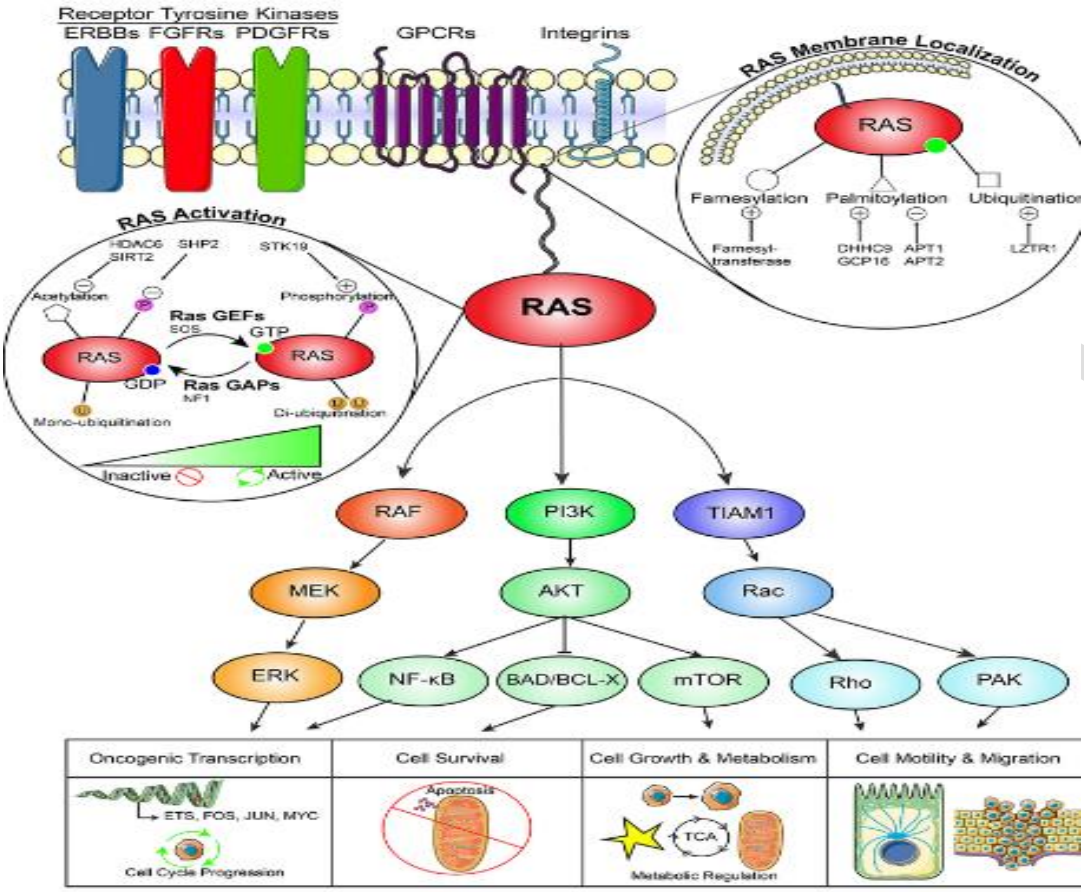


Figure-5: Ras Signalling Pathway

2. Wnt Pathway

The Wnt signaling pathway regulates cell fate determination, proliferation, migration, and polarity. It has two main branches: canonical (Wnt/β-catenin) and non-canonical pathway Mutations in Wnt pathway components (like APC, β-catenin) lead to deregulated β-catenin accumulation, causing continuous transcription of genes that promote cell proliferation and survival, thus contributing to tumor growth.

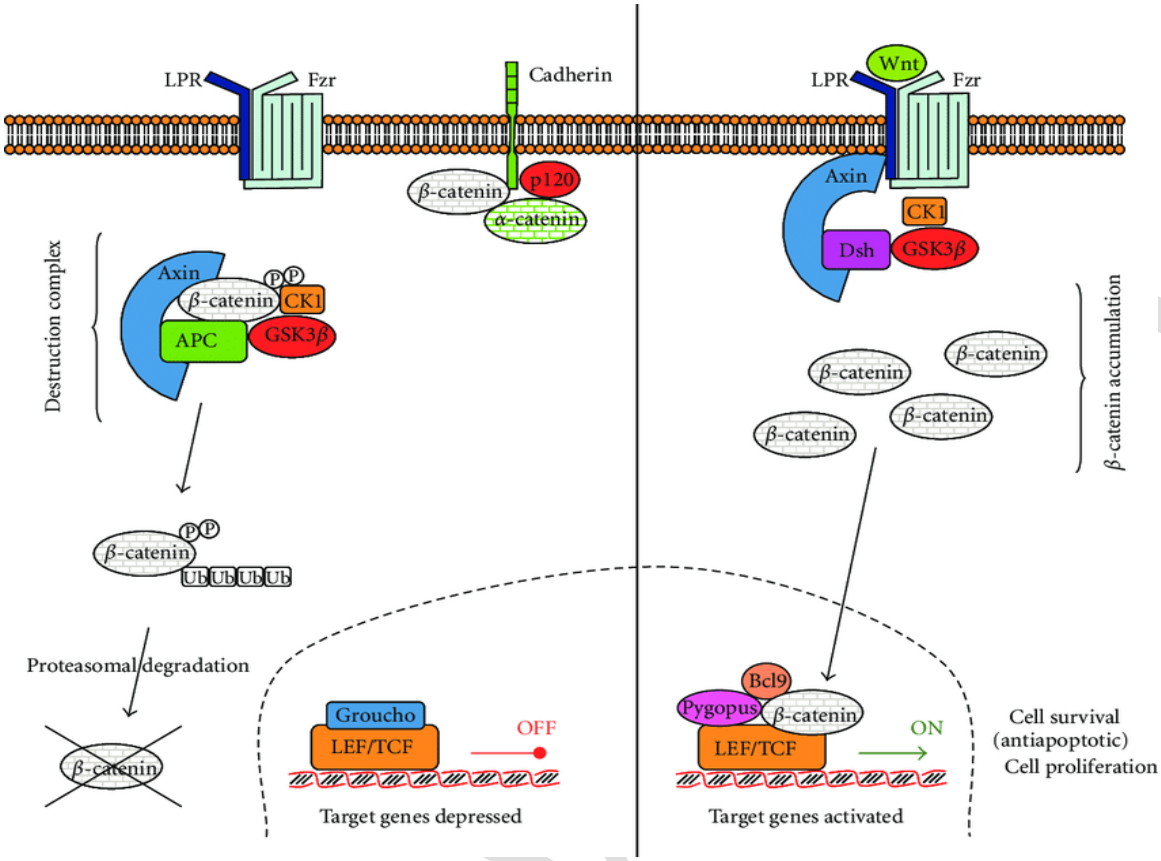


Figure-6: Wnt signalling pathway

3. PI3K/AKT/mTOR Pathway

This pathway is critical for regulating the cell cycle, promoting cell survival, and preventing apoptosis. The three main components, PI3K, AKT, and mTOR, work together to activate transcription factors and promote protein synthesis necessary for cell growth and division. Aberrant activation of the PI3K/AKT/mTOR pathway result in prolonged cell survival and resistance to apoptosis, contributing to tumor formation and growth. PI3K mutations (especially in the p110α subunit) and mTOR mutations are common in several cancers.

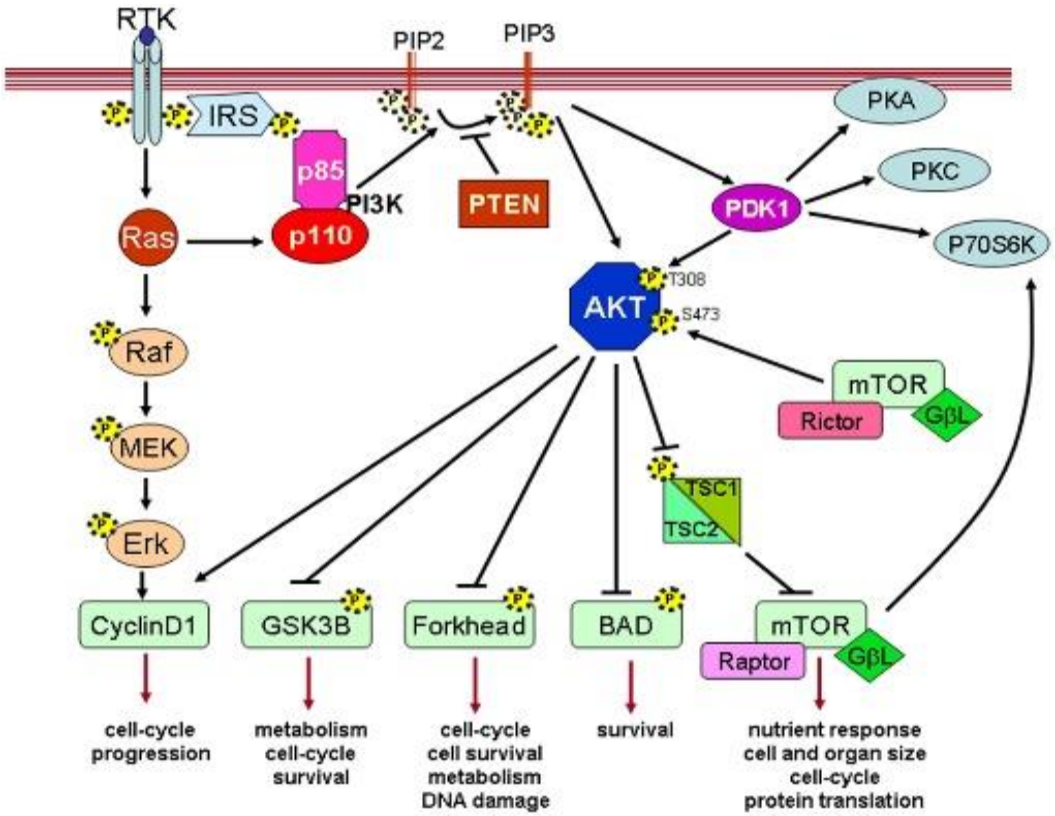


Figure-7: PI3K/AKT/mTOR Pathway

4. NF-κB Pathway

NF-κB regulates genes involved in immune and inflammatory responses. It is usually sequestered in the cytoplasm by inhibitory proteins and is activated by various stimuli, leading to the transcription of pro-inflammatory and survival genes. Deregulated activation of the NF-κB pathway results in excessive inflammatory responses, which can lead to chronic inflammation, promoting tumor growth and survival by creating a pro-inflammatory tumor microenvironment.

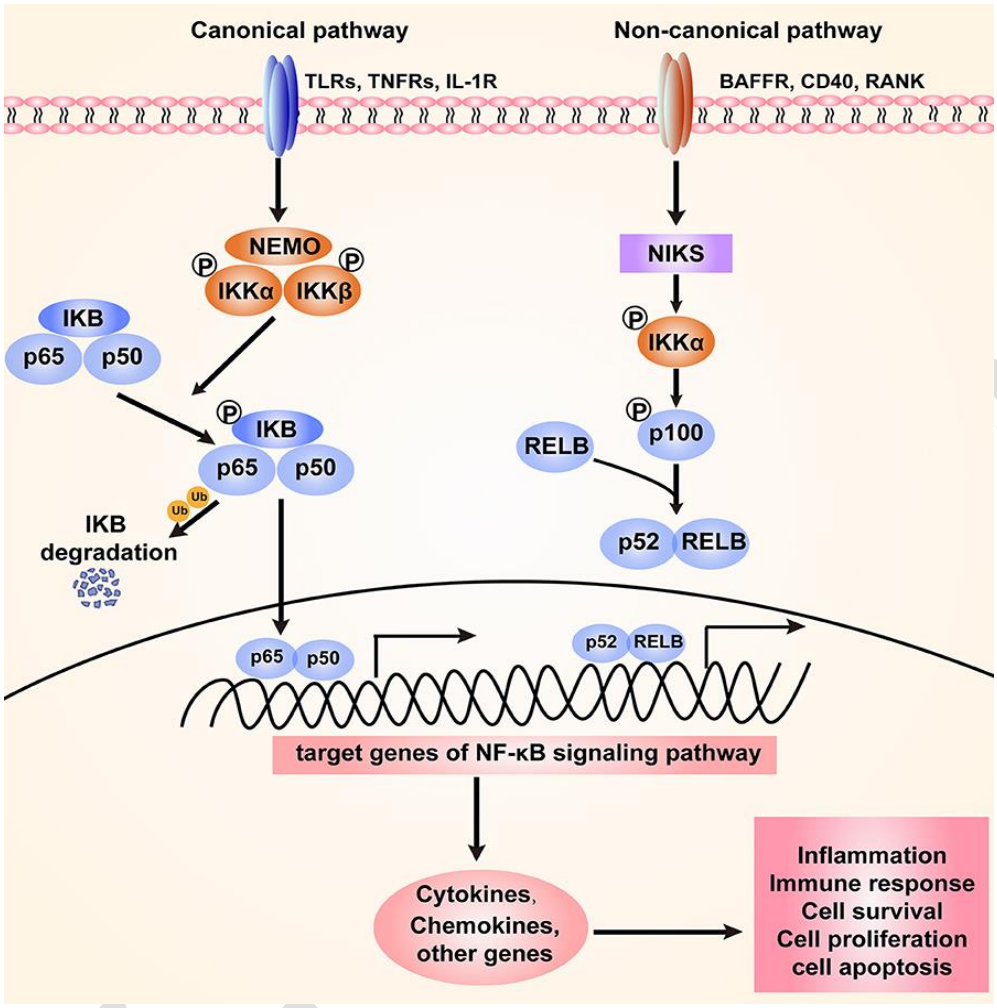


Figure-8: NF-κB Pathway

5. Calcium Signalling Pathway

Calcium (Ca^{2+}) is involved in several cellular processes such as gene transcription, apoptosis, autophagy, and cell proliferation. Abnormal calcium signalling can contribute to uncontrolled proliferation, invasiveness of tumour cells, and resistance to cancer therapies.

- Elevated intracellular Ca^{2+} levels are linked to enhanced cell migration and metastasis. Calcium signalling pathways, such as those mediated by TRP and other calcium channels, help regulate cancer cell motility, contributing to metastatic behaviour.

- Ca^{2+}/calmodulin (CaM) complexes activate multiple proteins, including CaM-dependent protein kinase IIs (CaMKIIs). This activation prolongs kinase activity, contributing to cancer cell survival and proliferation.
- Calcium signalling remodels the tumour microenvironment by regulating immune cells, promoting angiogenesis (formation of new blood vessels), and managing hypoxia (low oxygen conditions). For example, Ca^{2+} signalling affects the recruitment of macrophages, natural killer (NK) cells, and T cells, influencing the immune response in the TME.
- Various calcium channels and transporters, such as TRP channels, voltage-gated calcium channels (VGCCs), and stretch-activated Piezo channels, play essential roles in cancer cells. They help control calcium influx and efflux, which is crucial for maintaining the balance required for various signalling pathways in cancer.

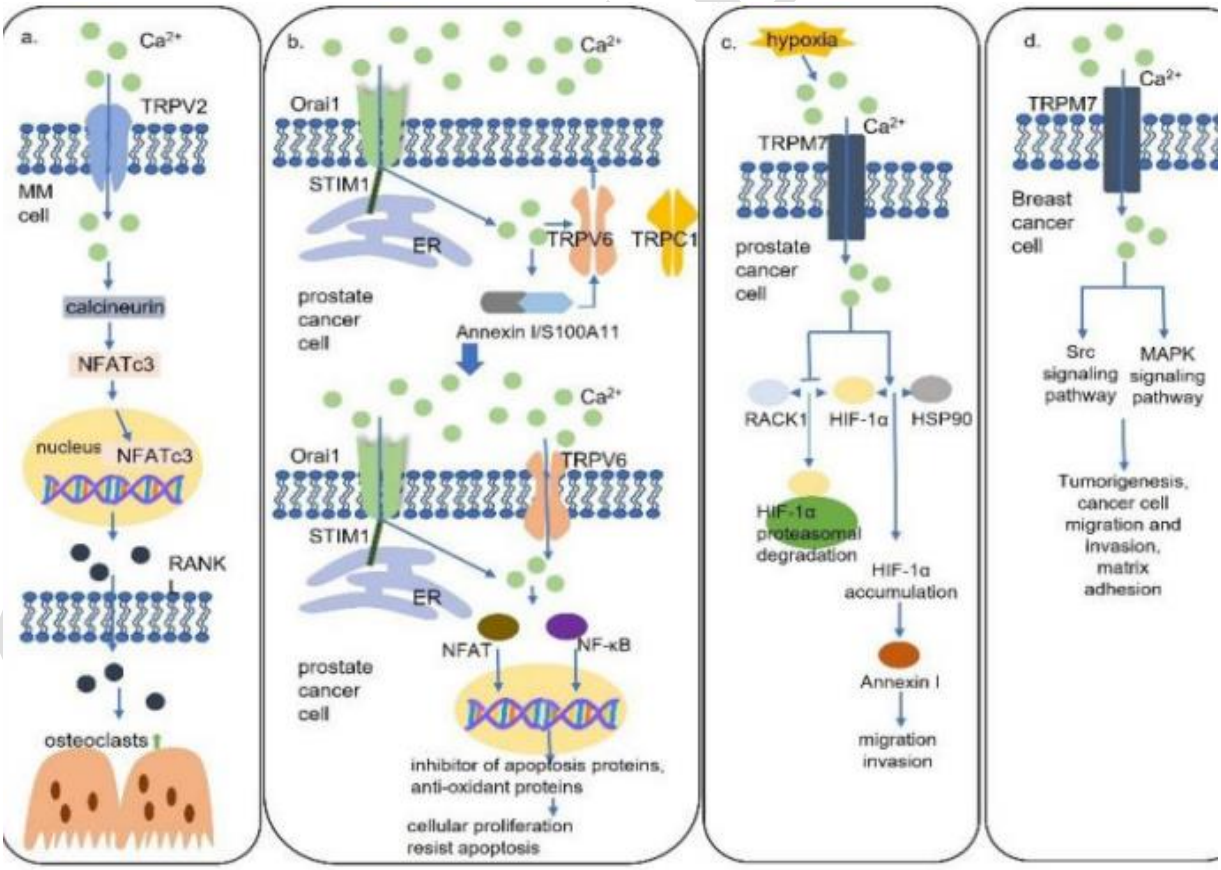


Figure-9: Calcium signalling

What are miRNAs?

miRNAs are small non-coding RNAs, with an average 22 nucleotides in length. Most miRNAs are transcribed from DNA sequences into primary miRNAs (pri-miRNAs) and processed into precursor miRNAs (pre-miRNAs) and mature miRNAs. In most cases, miRNAs interact with the 3′ UTR of target mRNAs to suppress expression. miRNAs have been shown to activate gene expression under certain conditions. Recent studies have suggested that miRNAs are shuttled between different subcellular compartments to control the rate of translation, and even transcription. miRNAs are critical for normal animal development and are involved in a variety of biological processes. Humans have been found to have about 2,500 microRNAs, which are engaged in almost every biological activity and an aberrant expression of miRNAs is associated with many human diseases, such as cancers.

Biogenesis:

MicroRNA biogenesis starts in the nucleus and ends at the cytoplasm. The biogenesis begins with DNA transcription, mediated by RNA Polymerase II or III. The first transcript has a hair-pin structure, with a dsRNA stem, protected at both the ends by capping and poly-A tailing. It consists of 100-10000 base pairs and is known as the primary or primordial (pri) miRNA.

Pri miRNA is further processed in the nucleus by DGCR8 and RNase III Drosha enzyme complex, leading to uncapping and poly-A tail removal. This leads to the formation of a ds-RNA molecule with a length of 65-100bp, known as the precursor (pre) miRNA. Pre miRNA is exported to the cytoplasm from the nucleus by exportin-5 and Ran-GTP complex.

In the cytoplasm, the pre miRNA is further processed by RNase type III Dicer enzyme. It cleaves the unpaired head of the hairpin structure and generates the miRNA duplex which is then interacts with RISC complex. Argonaut proteins in RISC leads to the cleavage of the duplex and it generates the mature miRNA.

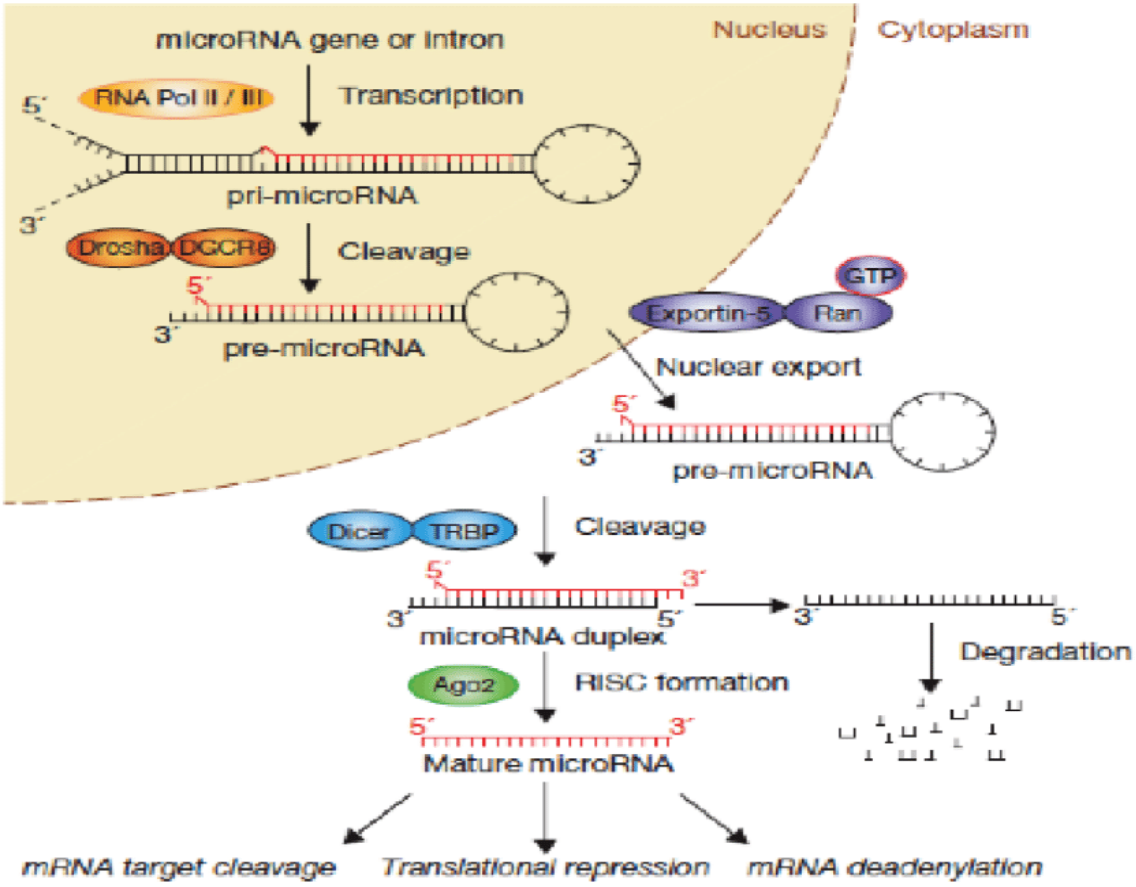


Figure-10: MicroRNA biogenesis

Oncomirs and tumour-suppressors:

Cancer-related miRNAs can be classified according to target mRNA function as tumor suppressor or oncogenic miRNAs (oncomiRs). This segregation is based on the ability of miRNA molecules to interfere in carcinogenesis-related processes, including mechanisms associated with cell migration and invasion, apoptosis and proliferation. Tumor suppressor miRNAs regulate the expression of mRNAs required for cell division or survival, whereas oncomiRs are more strongly expressed in cancer cells and down-regulate tumor suppressor genes, leading to enhanced cancer cell division**.**

miRNA Prediction Tools

MicroRNAs regulate gene expression by binding to microRNA responsive elements (mREs) on target mRNAs, resulting in significant alterations in a variety of physiological and pathological

processes. Thus, identifying miRNA-mRNA target interactions is critical for understanding the regulatory network mediated by miRNAs. Usually, computational prediction combined with experimental validation of these miRNA-mRNA interactions is the most effective approach to accomplish this goal. Therefore, computational techniques that assist in predicting which genes are targeted by miRNAs are known as microRNA (miRNA) prediction tools.

Several approaches underlie the development of miRNA target prediction algorithms.

These can be divided into two main categories:

1) **Algorithms based on the interaction between miRNA and mRNA or features of the mRNA sequence:**
    - **Seed Match:** The seed section of the miRNA (nucleotides 2-7) is critical for target recognition. The tools check for perfect or near-perfect complementarity between the seed sequence and the mRNA target region.
    - **3' UTR sequence:** Most programs use the 3' UTR dataset to look for a target site, because many studies have shown that this area is the most frequently targetable in miRNAs.
    - **Conservation:** The target site's evolutionary conservation across different species indicates its functional value. Tools look for the presence of the target location in similar organisms.
    - **Free Energy:** The energy required for miRNA-mRNA binding is an important consideration. Tools calculate the free energy change during interaction, with larger negative values suggesting stronger binding.
    - **Site Accessibility:** The mRNA target site should be accessible for miRNA binding. Some tools consider factors like RNA secondary structures that might hinder interaction.

2) **Statistical inference based on Machine Learning.**
    - Instead of using miRNA-mRNA properties like sequence or free energy, machine learning approaches identify putative miRNA targets by referencing miRNA-mRNA interactions that have been shown to have biological significance.
    - In general, machine learning is an artificial intelligence application that gives systems the capacity to automatically get better via experience; they "learn" from

sample datasets and apply the knowledge they have gained to forecast unknown data.

Some prediction tools and their features:

1) **TargetScan**

- With TargetScan, one can search for miRNAs by name, genes by name, or from miRNA families that are poorly, moderately, or widely conserved among various species.
- For conservation, the conservation of a 3' UTR is first determined followed by analysis of a specific k-mer (8mer, 7mer-m8, or 7mer-1A). Since one 3' UTR can contain multiple target sites, an aggregate PCT is provided. For each type of k-mer, the number is provided for that target and whether or not it is considered a conserved site or a poorly conserved site.

2) **PicTar**

- It is a computational tool designed to find common targets of microRNAs (miRNAs).
- To find miRNA targets, PicTar employs genome-wide alignments from eight distinct vertebrate species together with statistical testing.
- The tool has demonstrated a high degree of accuracy by correctly identifying known miRNA targets and experimentally validating seven predicted targets.

3) **miRANDA**

- The technique was first employed to locate targets in Drosophila, but it is not restricted to this use; it has also been utilized to predict targets in humans.
- **miRanda does a three-step analysis:**
  - To check for WC matches, the miRNA sequences provided as input are screened against the 3' UTRs provided by the user.
  - The free energy of each miRNA:mRNA target pair that surpasses a certain matching score is determined. Each target with a predicted free energy less than a threshold is then advanced to the last step.

- Finally, conservation of both binding site and position serves as a final filter. The remaining options are graded according on how closely they match the miRNA.

- A predicted target can be ranked high in the results if it receives a high individual score or has several anticipated sites.
- Unlike most miRNA target predictors, miRanda evaluates matching across the complete miRNA sequence. It considers the seed region by placing a higher priority on matches in that region. Matches can contain a restricted number of G-U wobble pairs, as well as insertions and deletions.

4) miRanda-mirSVR

- miRanda-mirSVR is an online tool that ***integrates two approaches***:
  - miRanda is used to discover potential target locations.
  - MirSVR (support vector regression) is used to rate them, using real-valued results.
- The findings are pre-computed, with no option to provide additional data, resulting in no user adjustment.
- MirSVR uses the real-valued outputs to generate a score that represents a miRNA's effect on expression.
- mirSVR was trained on nine miRNA transfection studies on HeLa cells and integrates a variety of additional essential parameters, including site accessibility, AU flanking content, target site position within the three UTRs, and UTR length.

5) miRWalk

- It was first introduced in 2010 by researchers at the University of Heidelberg in Germany.
- The major goal was to create a comprehensive database for predicting miRNA binding sites and facilitating the research of miRNA-target interactions.
- miRWalk employs a thorough prediction method that searches the entire gene sequence (5' UTR, CDS, and 3' UTR) for probable miRNA binding sites.

- It uses complementary base pairing between miRNAs and target mRNAs to predict binding sites.

6) **TargetMiner**
   - Targetminer is a bioinformatics tool designed for identifying potential microRNA targets in a given sequence.
   - It is based on machine learning and uses systematic identification of negative cases.
   - Key Features:
     - Machine learning-based model.
     - Focuses on minimizing false positives.

7) mirTargetLink2.0
   - It integrates data from high throughput experimental studies and computational predictions, this improves reliability and comprehensiveness of predicted interactions.
   - It provides users with a visualization interface to explore and analyze interaction networks between miRNAs and target genes by Machine learning algorithms.
   - Key Features:
     - Interactive Visualization
     - Hypothesis generation about miRNA function and regulatory mechanisms
     - Data integration

8) DIANA Tools
   - DIANA Tools (DNA Intelligent Analysis) is a comprehensive suite of web-based resources dedicated to microRNA (miRNA) analysis.
   - Among its functionalities, DIANA Tools stands out for its miRNA target prediction algorithms, particularly microT and microT-CDS.
   - Key Features:
     - Predicts targets in both the 3' UTR and CDS regions.
     - Incorporates miRNA binding site conservation and accessibility.
     - Includes a comprehensive database of experimentally validated targets.

9) StarBase

- StarBase (STAtic RNA-RNA interaction Balance Exploration) is a powerful online resource for exploring RNA-RNA interactions, particularly those involving microRNAs (miRNAs).
- Key Features:
    - miRNA-target interactions that are experimentally validated using CLIP-Seq (Cross-linking Immunoprecipitation Sequencing) data.
    - Support for Multiple Species including human, mouse, and zebrafish, making it useful for researchers working with various model organisms.
    - Interactive tools for visualizing and analysing RNA interactions

10) RNA22

- Pattern-based miRNA target prediction tool that first identifies putative target sites and then evaluates them for complementarity and accessibility. It uses statistical methods to discover recurring patterns in the input miRNA sequences.
- Key Features:
    - Does not rely on conserved sequences.
    - Evaluates miRNA-mRNA binding at a whole-transcriptome level.
    - Provides a web interface for easy access.

11) miRDB

- Database for miRNA target prediction and functional annotations. It uses a machine learning-based approach to predict targets based on thousands of miRNA-target interactions.
- Key Features:
    - Employs a support vector machine (SVM) model trained on high-throughput experimental data.
    - Provides target predictions for various species.
    - All the predicted targets have target prediction scores between 50 - 100. The higher the score, the more confidence we have in this prediction.

12) RNAhybrid

- Computational tool used for predicting miRNA target sites by finding the energetically most favorable hybridization sites between miRNA and mRNA.
- Key Features:

- Calculates the minimum free energy of hybridization.
- Allows prediction of non-canonical binding sites.
- Can be run locally or online.

The major tools used and focused on, for collecting the experimental data were TarBase and miRTarBase and for collecting the prediction data, mirDIP was used.

1. **<u>TarBase:</u>**

   i. It is database that stores a collection of experimentally validated microRNA/miRNA target interactions. The interactions stored in this database has been tested by various test methods like Luciferase Reporter Assays, Western blotting, qRT-PCR & CLIP. It The data provided by TarBase is used by scientists and researchers all over the world for biological studies.

   ii. It is an enormous database of over 6 million entries available online, which is three times larger than miRTar-Base and six times larger than TarBase-v8.0. This makes it the largest collection of experimentally validated miRNA-gene interactions to date. Utilizing mi-croCLIP, a cutting-edge CLIP-Seq analysis framework that integrates deep learning classifiers under a super learning scheme for CLIP-Seq-guided miRNA interaction discovery, is a first for TarBase. Additionally, using a standardized data processing pipeline, TarBase-v9.0 incorporates interactions from direct miRNA–target chimeras generated by the de novo analysis of CLASH and qCLASH experiments, as well as miRNA–gene pairs derived from the analysis of miRNA-specific transfection / knockdown RNA-Seq experiments. TarBase v9.0 is a significant release as it includes around 34,000 pairs of interactions between host mRNAs and virally-encoded miRNAs, interactions that result in target-directed miRNA degradation (TDMD) events, and millions of precise miRNA-binding positions that can be resolved down to the cell type. (Skoufos et al., 2024)

   iii. The TarBase interface has undergone a complete redesign, adding a number of new features to increase its usefulness and versatility. With a plethora of filtering options, including cell lines, experimental settings, cell types, experimental

procedures, species, and tissues of interest, the recently developed interface enables users to create complex searches. Moreover, a multitude of fine-tuning features have been smoothly incorporated into the platform, allowing users to further refine their search results according to expression levels, microRNA confidence, microCLIP and DIANA-microT 2023 scores, and more. Furthermore, the platform now allows for unlimited local retrieval of the interactions as well as all related metadata. (Skoufos et al., 2024)

iv. Additionally, TarBase-v9.0 connects to other biological databases with ease. TarBase has been included into Ensembl since version six. Ensembl genome browser allows viewing interactions with specific binding sites; miRBase and Ensembl provide access to gene and miRNA information, respectively. The RNACentral knowledgebase lists TarBase as another expert non-coding RNA resource. Lastly, publications' details may be obtained directly via the connection to PubMed, and DIANA-microT-CDS 2023 also provides access to predicted binding sites and miRNA-target interaction scores. (Skoufos et al., 2024)

v. Primarly, during this study, interactions section of TarBase has been used to collect the data of microRNA targets, with primary interactions and which have been obtained through direct experimental method. The values with high confidence value had been filtered and extracted for later studies.

vi. For both basic and applied RNA research projects, meticulous generation, curation, and indexing of miRNA–gene interactions across many experimental situations and species is crucial. The continuous dedication to methodically cataloguing millions of empirically supported miRNA targets for more than 15 years is demonstrated by TarBase-v9.0. With about 6 million entries and 2 million distinct miRNA-target combinations supported by 37 experimental procedures (both high- and low-yield) across 172 tissues and cell types, the most recent version is available. Crucially, it lists interactions that result in target-directed miRNA degradation (TDMD) events for the first time, as well as millions of miRNA-binding locations with cell-type

precision that are generated by the de novo analysis of cutting-edge NGS techniques. Users are able to access this information via a redesigned interface that adds a ton of new features and improves adaptability and user-friendliness. (Skoufos et al., 2024)

vii. The greatest database of miRNA-gene interactions that has been supported by experimentation is TarBase v9.0. This abundance of data can confirm in silico expected interactions or in some circumstances even replace them. (Skoufos et al., 2024)

2. <u>miRTarBase:</u>

i. More than 470,000 miRNA-target interactions (MTIs) have been added to the database miRTarBase as a whole. These MTIs are gathered by hand after NLP of the text is performed in order to filter research articles pertaining to functional investigations of miRNAs. Typically, reporter assay, western blot, microarray, and next-generation sequencing tests are used to experimentally confirm the gathered MTIs. When compared to other comparable, previously created databases, the miRTarBase offers the most current collection while having the most certified MTIs.

ii. This tool is used for integrative analysis of high-throughput data, microRNA & target gene information, text-mining technique to prescreen literature and to study miRNA and 3'UTR-related SNPs or DRVs, regulatory factors of microRNA, circulating cell-free microRNAs, etc. Some notable features of the current version include word cloud of miRNA-disease information, miRNA-Target site viewer in CLIP-seq data, Clinical microRNA and gene expression profiles from TCGA, etc.

iii. miRTarBase search option has been used in this study and the data had been extracted based on the CLIP-Seq.

iv. 2011 saw the initial release of miRTarBase. The number of MTIs that have been experimentally validated has grown significantly over the past ten years and now stands at a noteworthy level. Simultaneously, the miRTarBase web interface has undergone continuous updates and improvements to offer users a more effective and superior access experience. This latest update includes more biological data and a variety of miRNA expression patterns in addition to providing more thorough information on MTIs. Such data is displayed in an intuitive, beautiful, and succinct online interface. More effectively, MTIs were retrieved from related papers obtained from the PubMed literature database using an enhanced scoring system. The specificity and heterogeneity of miRNAs were shown by the miRNA expression profiles across extracellular vesicles, blood, and other organs, as well as by SNPs and DRVs in miRNAs and gene UTRs. These findings also supported the identification of miRNA biomarkers. (Huang et al., 2022)

v. Version 9.0 has considerably enhanced the number of MTIs retrieved from research publications and CLIP-seq data over version 8.0. From 13 389 research publications and CLIP-seq data, a total of 19 912 394 empirically validated MTIs between 4630 miRNAs and 27 172 mRNAs (target genes) were manually curated. To support the numerous MTIs reported, version 9.0 has integrated 440 CLIP-seq data from 44 different independent investigations. (Huang et al., 2022)

vi. Circular RNA and an upstream transcription factor can mediate the expression of miRNAs, which can control gene expression through a posttranscriptional regulatory mechanism. To learn more about the roles and biomarkers of miRNAs, a thorough analysis of their expression levels in extracellular vesicles or tissues is necessary. According to recent research, SNPs and variants influence the miRNA–mRNA binding affinity and miRNA production rather than being controlled by regulators. In order to better explore the regulation mechanism of miRNAs, this update endeavors to create regulatory networks connecting targets, regulators, and miRNAs. Additionally, we show the distribution of miRNA expression in extracellular vesicles and throughout tissues, as well as SNPs and DRVs carried in

miRNAs or gene 3UTRs that may be linked to the genesis of disease. (Huang et al., 2022)

vii. It is interesting to note that papers referencing miRTarBase are scattered throughout a wide range of disciplines; nonetheless, for all miRTarBase versions, "biochemistry and molecular biology" and "oncology" alternate as the top two citation categories, with "genetics and heredity" typically following closely behind. Understanding the course of disease requires an understanding of miRNA regulation and target interaction. By offering a new web query interface, the revised miRTarBase facilitates the integration of regulators and targets and provides known, empirically confirmed MTIs to study miRNA regulation. (Huang et al., 2022)

viii. Furthermore, miRTarBase constructed an osteoarthritis-specific miRNA interactome expressing the gene in cartilage influenced by miRNA by combining the publicly available, empirically confirmed miRNA and target interaction database. A versatile search interface for miRNA, the target gene, and associated disorders is offered by miRTarBase. For instance, when we search for hsa-miR-195-5p, a list of the main diseases that are associated with it appears; in this instance, glioblastoma cancer was the primary focus for this miRNA, which has been shown in independent studies and TCGA data to have a significant negative correlation with the progression of the disease. Together, we may find novel MTIs within our illness of interest and learn more about miRNA regulation by filtering known MTIs that are included in miRTarBase. (Huang et al., 2022)

3. <u>mirDIP:</u>

i. The revised version of mirDIP that is being given collects miRNAtarget interactions from several updated sources and then annotates them with the integrated score that was previously stated. In addition, MirDIP 5.2 includes context annotation for miRNA-gene connections for both healthy and diseased tissues. This allows for more in-depth examination of condition-specific miRNA

interaction networks. It also includes a curation of new miRNAs from RNAseq studies published in the literature, as well as interaction predictions for mir-GeneDB and miRBase miRNAs, increasing the quantity and quality of miRNAs. (Hauschild et al., 2023)

ii. 46 364 047 predictions for 27 936 genes and 2734 miRNAs are included in mirDIP 5.2. The only resource that offers interactions for the high-quality data from mirGeneDB is mirDIP 5.2. (Hauschild et al., 2023)

iii. To facilitate a range of workflows, mirDIP offers several query types:

- The user can directly search for binary or quantile-normalized miRNA context associations by using the Search Tissues via miRNAs and miRNAs (scale) options. A score that indicates how well supported a miRNA's expression is in a given environment is generated by combining data from the chosen datasets for each miRNA. The number of selected sources in which the miRNA is expressed in the context divided by the number of selected sources that measure the miRNA in the context yields the score of a miRNA in a context for a binary value inquiry. (Hauschild et al., 2023)
- Context-specific interactions are chosen by combining the relevant mirDIP interaction information with both miRNA and gene contexts through the use of the Search Tissues and Interactions option. (Hauschild et al., 2023)
- The Tissue Matrix option displays the findings in an accumulated matrix with one column for each tissue and enables querying of context-specific miRNA and interactions. (Hauschild et al., 2023)

iv. The context of each miRNA-gene interaction in the mirDIP database is rated according to data on the measurement of both miRNA and gene expression across several datasets. Value 0 denotes that at least one of the two molecules has no expression in the context across all of its sources, whereas value 1 shows that both entities' expression is unanimous among their respective sources (i.e., all the datasets that assessed the gene and miRNA in that context). (Hauschild et al., 2023)

v. Additionally, mirDIP v.5.2 integrates gene expression for 278 tissue/cell type and illness settings, including 90 distinct disease states for 27 576 genes and 188 distinct normal tissues and cell types (with differing degrees of tissue specificity). The range of miRNA expression in normal tissues is 1 to 206, however in illness settings it can range from 1 to a maximum of 92 situations. On the other hand, the expression of various microRNAs in sickness varies from 183 to 2529 miRNAs, whereas in normal conditions it varies from 1 to 2526 miRNAs. (Hauschild et al., 2023)

GOAL OF THE STUDY:

- To understand the regulation of oncomeric microRNAs in GBM tumors

OBJECTIVES:

- Identify the common target genes regulated by selected oncomiRs in GBM
- Bioinformatic analysis of microRNA target expression at mRNA and protein levels
- Network and pathway analysis using microRNA targets

2.0 METHODS:

1. Literature Review:

We conducted thorough research of scientific papers to identify microRNAs (miRNAs) involved in GBM. A number of 8 miRNAs were selected for the study, namely, miR-21, miR-155, miR93, miR-221, miR-222, miR-196a, miR-10b and miR-182.

2. Use of Experimentally Verified Databases:

Utilized Bioinformatic tools such as **TarBase** and **miRTarBase** to select the target genes of the miRNAs. TARBASE v9.0 is a database consisting of experimentally supported miRNA targets on protein-coding transcripts. We extracted genes targeted by different microRNAs by selecting, homo sapiens as species, direct as experimental type and primary interactions only. miRTARBASE is a bioinformatic tool consisting of experimentally verified data of miRNA-target

interactions, validated by reporter assay, western blot, microarray and next-generation sequencing experiments. We used Clip-seq (NGS) and Microarray as filters to obtain data for targets regulated by OncomiRs involved in GBM, in humans.

3. **Prediction of miRNA Targets**:

   Employed **miRDIP** prediction tool to predict the target genes of the selected miRNAs. miRDIP (microRNA Data Integration Portal) is a tool that integrates data from multiple microRNA target prediction and validation databases. We downloaded the data for all the 8 miRNA's predicted target genes, with medium and high score.

4. **Common Target Identification**:

   We identified 195 common targets shared among all selected miRNAs by analyzing and integrating the data from the aforementioned online tools.

5. **Categorization of Targets into Hallmarks of Cancer**:

   Used **TISIDB** (Tumor-Immune System Interaction Database) to categorize the 195 common targets into various hallmarks of cancer and mapped targets to specific cancer pathways, such as proliferation, invasion, apoptosis, etc. TISIDB is a user-friendly web portal for a comprehensive investigation of tumour-immune interactions, which integrates multiple types of data resources in onco-immunology.

6. **Expression Analysis at mRNA level in GBM**:

   Utilized **GEPIA2** to check the level of regulation of the identified targets in GBM tissue samples, with the help of boxplots. GEPIA2 is a highly cited resource for gene expression analysis based on tumour and normal samples from the **TCGA (The Cancer Genome Atlas)** and the **GTEx** databases. It features 1,98,619 isoforms and 84 cancer subtypes, has extended gene expression quantification from the gene level to the transcript level, and supports analysis of a specific cancer subtype, and comparison between subtypes. In addition, GEPIA2 has adopted new analysis techniques of gene signature quantification inspired by single-cell sequencing studies. Focused on the targets showing significant differential expression in GBM.

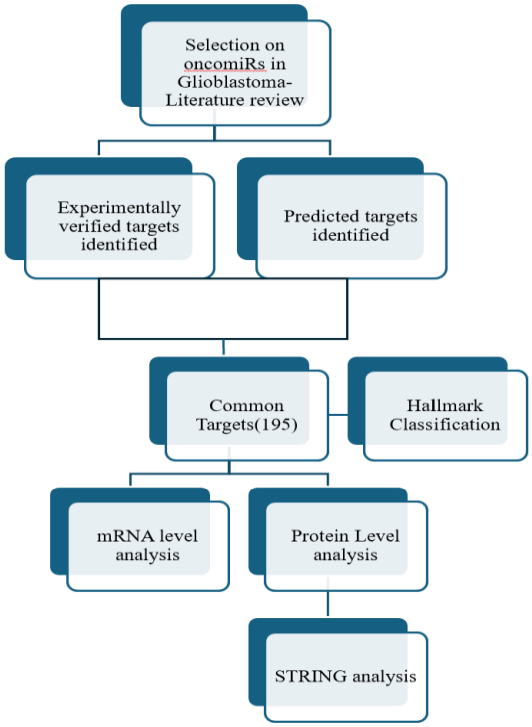
7. **Protein-Level Analysis**:

Performed protein expression validation using **UALCAN (**The **U**niversity of **AL**abama at Birmingham **CAN**cer data analysis) portal, to assess the levels of selected targets at the protein level in GBM. UALCAN is a comprehensive, user-friendly, and interactive web resource for analysing cancer OMICS data. It is designed to, **a)** provide easy access to publicly available cancer OMICS data (TCGA, MET500, CPTAC and CBTTC), **b)** allow users to identify biomarkers or to perform in silico validation of potential genes of interest, **c)** provide graphs and plots depicting expression profile and patient survival information for protein-coding, miRNA-coding and lincRNA-coding genes, **d)** evaluate epigenetic regulation of gene expression by promoter methylation, **e)** perform pan-cancer gene expression analysis, **f)** Provide additional information about the selected genes/targets by linking to GeneCards, Pubmed, TargetScan, The human protein atlas, DRUGBANK, Open Targets and the GTEx. These resources allow researchers to gather valuable information and data about the genes/targets of interest, **g)** provide clinical proteomic consortium data analysis including total/phospho-proteins and **h)** provide pediatric brain tumour gene expression and protein expression analysis. 58 target genes showcased down-regulation at protein level

8. **Network and Interaction Analysis**:

Used **STRING** analysis to evaluate the protein-protein interaction (PPI) networks of the selected targets. STRING database systematically collects and integrates protein-protein interactions-both physical interactions as well as functional associations. The data originates from several sources: automated text mining of the scientific literature, computational interaction predictions from co-expression, conserved genomic context, databases of interaction experiments and known complexes/pathways from curated sources. We selected interconnected targets based on their relevance in the PPI network, aiming to identify key nodes in GBM pathways, out of which 45 were observed to be statistically significant. The target interactions were observed and analysed based on the KEGG pathways, wiki pathways, reactome pathways, molecular function, cellular component and disease-gene associations.

The experimental design of the study is shown below in **Figure 11.**



3.0 Results:

The number of experimentally verified targets as obtained from Tarbase and miRTarbase for miRNAs-miR-21-5p, miR-93-5p, miR-155-5p, miR-10b-5p, miR-221-3p, miR-222-3p, miR-182-5p, miR-196a-5p, are:

Tarbase:

List of microRNAs	Number of targets identified from Tarbase
miR-21-5p	7997
miR-93-5p	6101
miR-155-5p	4379
miR-10b-5p	3351
miR-221-3p	6719
miR-222-3p	2934
miR-182-5p	7228
miR-196a-5p	9599

miRTarbase:

List of microRNAs	Number of targets identified from miRTarbase
miR-21-5p	140
miR-10b-5p	302
mir-155-5p	216
miR-93-5p	634
miR-221-3p	30
miR-222-3p	109
miR-196a-5p	103
miR-182-5p	111

The number of predicted targets as obtained from miRdip for the same set of miRNAs are:

microRNA	No. of targets identified with medium Score	No. of targets identified with high Score
miR-21-5p	7570	1114
miR-10b-5p	6570	952
miR-155-5p	6643	977
miR-93-5p	6329	931
miR-221-3p	7619	1121
miR-222-3p	7608	1119
miR-196a-5p	6573	967
miR-182-5p	6884	1013

List of 195 common target genes of 8 microRNAs is found to be:

ABCA1	EFCAB14	MFHAS1	SOD2
ADAM10	EIF5	MPRIP	SON
ADAR	ENTPD7	MTDH	SORT1
AFF4	FAM222B	MTPN	SP1

AKAP11	FBXO45	NAA50	SPIN1
ANKFY1	FEM1B	NAP1L1	SPTY2D1
APLP2	FGF2	NF1	SSH1
ARHGAP11A	FNDC3B	NHLRC2	SSX2IP
ARHGAP5	FOSL2	NR2C2	STRN
ARHGEF12	FOXJ3	NRP1	SYNE2
ARRDC3	FOXO3	NSD1	TBL1XR1
ASH1L	FRS2	NUCKS1	TCF12
ATP2B1	FRYL	NUFIP2	TEAD1
ATXN1	GAN	NUP58	TET2
ATXN1L	GFPT1	PAK2	TFRC
BCAT1	GIT2	PAPOLA	TGFBR1
BCL2L11	GLS	PJA2	THBS1
BICC1	GRSF1	PODXL	TMED7
BICD1	GSPT1	POU2F1	TMEM245
BMPR2	GTF3C4	PPP1CC	TNKS2
BRWD1	HELZ	PRDM2	TNPO1
C5orf24	HIF1AN	PRRC1	TNRC6B
CALU	HIPK1	PTEN	TNRC6C
CANX	HMGA2	PTPN14	TRAM1
CASK	HNRNPU	PTPRJ	TRIO
CBX5	IREB2	PURB	TSC22D2
CCNT2	ITGB8	QKI	TTN
CDC27	KBTBD2	QSER1	TULP4
CDK6	KBTBD6	RAB11FIP2	TXNRD1
CDV3	KDM5B	RAPGEF2	UHMK1
CELF1	KMT2A	RBM12	USP46
CEMIP2	KMT2C	RCOR1	USP9X

CHORDC1	LANCL1	RHOBTB3	VANGL1
CLTC	LCOR	RIC1	WAC
CNOT6	LEPROT	RLIM	WAPL
CNST	LHFPL2	RNF213	WDR26
CREBL2	LPGAT1	RNF44	WDR33
CREBRF	LPP	RREB1	WEE1
CUX1	LRRC58	SCARB2	XIAP
CYLD	MAP1B	SCD	YAP1
DAG1	MAP3K2	SERTAD2	YOD1
DCAF7	MAP4	SESN3	ZFP36L1
DCP2	MAPK1	SETD2	ZMYND11
DDX21	MAT2A	SH3GLB1	ZNF644
DDX3X	MAVS	SIN3A	ZNF652
DDX6	MBD2	SIX4	
DICER1	MBNL1	SKIL	
DR1	MCL1	SLC35F5	
E2F3	MDM2	SLC38A1	
E2F7	MET	SNTB2	

According to mRNA level performed by GEPIA2, the total upregulated and downregulated gene targets are 147 and 29 respectively as shown in Table below.

Targets	**GBM**	**Targets**	**GBM**	**Targets**	**GBM**
ABCA1	up	LEPROT	up	TMEM245	up
ADAM10	up	LHFPL2	up	TNKS2	up
ADAR	up	LPGAT1	down	TNPO1	up
AFF4	up	LPP	up	TNRC6B	down
AKAP11	down	LRRC58	up	TNRC6C	down
ANKFY1	up	MAP1B	down	TRAM1	up
APLP2	up	MAP3K2		TRIO	up

ARHGAP11A	up	MAP4	no change	TSC22D2	up
ARHGAP5	slight down	MAPK1	up	TTN	down
ARHGEF12	no change	MAT2A	up	TULP4	down
ARRDC3	up	MAVS	up	TXNRD1	up
ASH1L	down	MBD2	up	UHMK1	up
ATP2B1	down	MBNL1	up	USP46	down
ATXN1	down	MCL1	up	USP9X	up
ATXN1L	up	MDM2	up	VANGL1	up
BCAT1	up	MET	no change	WAC	down
BCL2L11	up	MFHAS1	up	WAPL	-
BICC1	no change	MPRIP	down	WDR26	up
BICD1	up	MTDH	up	WDR33	up
BMPR2	up	MTPN	up	WEE1	up
BRWD1	down	NAA50	up	XIAP	up
C5orf24	up	NAP1L1	up	YAP1	up
CALU	up	NF1	no change	YOD1	up
CANX	up	NHLRC2	no change	ZFP36L1	up
CASK	up	NR2C2	no change	ZMYND11	down
CBX5	up	NRP1	up	ZNF644	up
CCNT2	up	NSD1	up	ZNF652	up
CDC27	up	NUCKS1	up		
CDK6	up	NUFIP2	up		
CDV3	up	NUP58	NA		
CELF1	up	PAK2	up		
CEMIP2	-	PAPOLA	up		
CHORDC1	down	PJA2	up		
CLTC	up	PODXL	up		
CNOT6	up	POU2F1	up		
CNST	down	PPP1CC	up		

CREBL2	up	PRDM2	down		
CREBRF	No change	PRRC1	up		
CUX1	up	PTEN	up		
CYLD	up	PTPN14	up		
DAG1	up	PTPRJ	up		
DCAF7	up	PURB	up		
DCP2	up	QKI	up		
DDX21	up	QSER1	up		
DDX3X	up	RAB11FIP2	down		
DDX6	up	RAPGEF2	down		
DICER1	no change	RBM12	up		
DR1	up	RCOR1	up		
E2F3	up	RHOBTB3	up		
E2F7	up	RIC1	up		
EFCAB14	up	RLIM	up		
EIF5	up	RNF213	up		
ENTPD7	up	RNF44	down		
FAM222B	up	RREB1	up		
FBXO45	no change	SCARB2	up		
FEM1B	slightly up	SCD	down		
FGF2	up	SERTAD2	up		
FNDC3B	up	SESN3	up		
FOSL2	up	SETD2	up		
FOXJ3	up	SH3GLB1	up		
FOXO3	no change	SIN3A	up		
FRS2	up	SIX4	up		
FRYL	no change	SKIL	up		
GAN	no change	SLC35F5	up		
GFPT1	up	SLC38A1	down		

GIT2	no change	SNTB2	up		
GLS	down	SOD2	up		
GRSF1	up	SON	up		
GSPT1	up	SORT1	no change		
GTF3C4	up	SP1	up		
HELZ	up	SPIN1	no change		
HIF1AN	up	SPTY2D1	up		
HIPK1	up	SSH1	up		
HMGA2	up	SSX2IP	down		
HNRNPU	up	STRN	up		
IREB2	up	SYNE2	up		
ITGB8	up	TBL1XR1	up		
KBTBD2	up	TCF12	up		
KBTBD6	slightly up	TEAD1	up		
KDM5B	up	TET2	up		
KMT2A	down	TFRC	up		
KMT2C	up	TGFBR1	up		
LANCL1	down	THBS1	up		
LCOR	down	TMED7	up		

The hallmark classification data for all 195 identified common gene targets is as mentioned below:

Term	**Adjusted P-value**	**Genes**
SUSTAINING PROLIFERATIVE SIGNALING	0.19	VANGL1;MAVS;DICER1;MAPK1;DDX3X;CUX1;SORT1; CREBL2;MAP3K2;SIN3A;POU2F1;TFRC;BCL2L11;CDK6; SSX2IP;AFF4;PTPRJ;PTEN;THBS1;TGFBR1;MTPN;NRP1; BMPR2;E2F3;TBL1XR1;HIPK1;ARHGEF12;FRS2;FGF2;D DX6;ITGB8;NF1;SCD;PPP1CC;FOXO3;KMT2A;TXNRD1; YAP1;WEE1;PAK2;ADAM10;MDM2;RAPGEF2;SP1;TET2;

		ARHGAP5;CNOT6;ARHGAP11A;MET;TEAD1;MCL1;NAP1L1;CLTC
GENOME INSTABILITY	0.06	KMT2C; WEE1; CLTC; DICER1; CDK6; MDM2; CUX1; TET2; SETD2; PTEN; MET; NAP1L1; SESN3; THBS1; NUCKS1; XIAP
EVADING GROWTH SUPPRESSORS	0.13	KMT2C;SOD2;MAVS;DICER1;MAPK1;DDX3X;CUX1;SORT1;PTPN14;CREBL2;MAP3K2;SIN3A;STRN;TFRC;MAP4;QKI;BCL2L11;CDK6;SERTAD2;PTPRJ;SKIL;MFHAS1;PTEN;THBS1;TGFBR1;NRP1;BMPR2;E2F3;HMGA2;TBL1XR1;KBTBD6;FGF2;CYLD;NF1;GLS;PPP1CC;FOXO3;CELF1;TXNRD1;YAP1;E2F7;PAK2;SP1;MDM2;RAPGEF2;TET2;ARHGAP5;ARHGAP11A;MET;TEAD1;MCL1
EVADING IMMUNE DESTRUCTION	0.30	FOXO3; ENTPD7; MAVS; PAK2; MAPK1; CUX1; TMED7; ARHGAP5; ADAR; PTEN; MET; KMT2A
SUSTAINED ANGIOGENESIS	1.48E-05	DICER1;MAPK1;CUX1;SETD2;ZFP36L1;STRN;QKI;THBS1;TGFBR1;NRP1;BMPR2;HMGA2;RNF213;HIPK1;ARHGEF12;FRS2;FGF2;CYLD;ITGB8;NF1;PPP1CC;YAP1;SP1;MDM2;MTDH;MET;KMT2A;DDX6
TISSUE INVASION AND METASTASIS	0.001	KMT2C;VANGL1;DICER1;MAPK1;DDX3X;PODXL;SORT1;CUX1;MAP3K2;STRN;XIAP;MAP4;ASH1L;DAG1;SSX2IP;AFF4;PTPRJ;PTEN;SSH1;THBS1;TGFBR1;USP9X;BMPR2;NRP1;TBL1XR1;HMGA2;ARHGEF12;FRS2;FGF2;CYLD;ITGB8;NF1;PPP1CC;RHOBTB3;YAP1;PAK2;SP1;CASK;ADAM10;RAPGEF2;LPP;MDM2;ARHGAP5;TET2;MAP1B;MET;KMT2A;GIT2
TUMOR-PROMOTING INFLAMMATION	0.30	XIAP; FOXO3; ENTPD7; CANX; MAPK1; TMED7; ARHGAP5; CYLD; ITGB8; KMT2A; THBS1; TFRC
RESISTING CELL DEATH	0.001	SOD2;SIX4;DICER1;MAPK1;DDX3X;CUX1;SORT1;MAP3K2;SIN3A;STRN;XIAP;TFRC;MAP4;DAG1;BCL2L11;CDK6;PTEN;SSH1;THBS1;TGFBR1;BMPR2;E2F3;HMGA2;FGF2;DDX6;CYLD;ITGB8;NF1;SESN3;FOXO3;SCARB2;E2F7;

		PAK2;ADAM10;MDM2;SP1;SON;MTDH;MET;KMT2A;MCL1;CLTC
REPROGRAMMING ENERGY METABOLISM	0.10	FOXO3; SOD2; ARRDC3; BCL2L11; MAPK1; TET2; DDX3X; MDM2; GFPT1; ZFP36L1; MET; PTEN; SCD; GLS; TGFBR1
REPLICATIVE IMMORTALITY	0.0001	KMT2C; VANGL1; DICER1; MAPK1; TNKS2; CDK6; SKIL; PTEN; THBS1; TGFBR1; NRP1; BMPR2; E2F3; TBL1XR1; FGF2; FOXO3; E2F7; SP1; MDM2; KMT2A

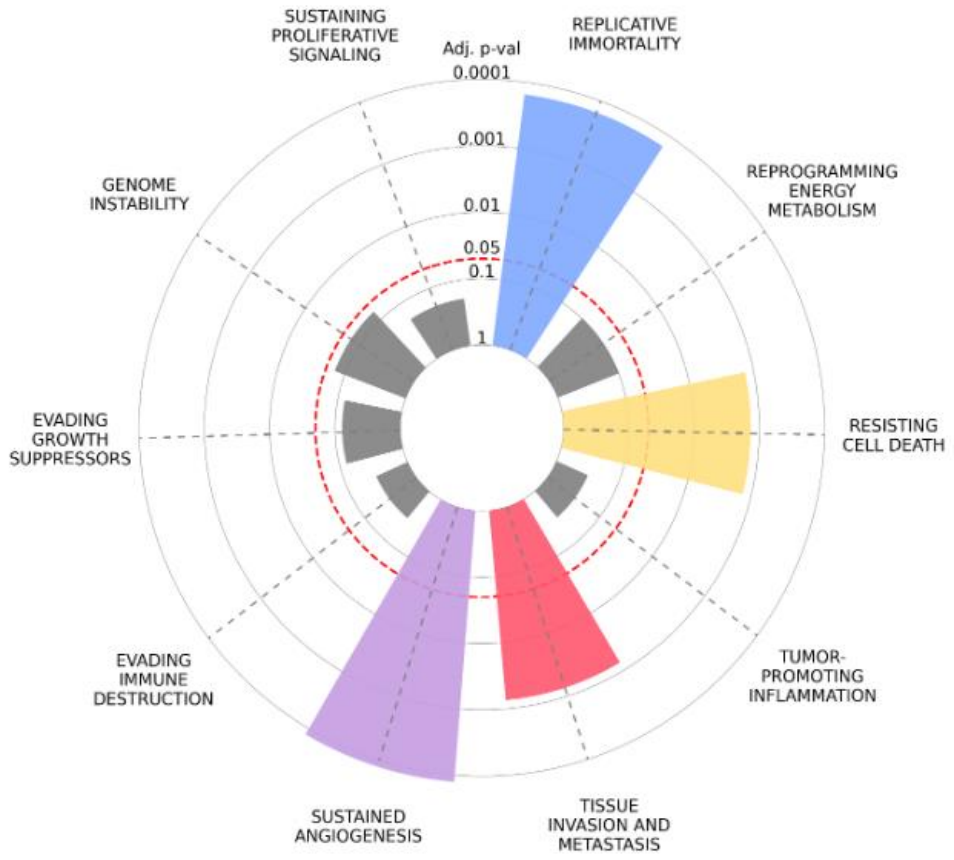


Figure-12: Hall marks classification of 195 targets

On performing protein level analysis for 99 gene targets using UALCAN, the upregulated and downregulated gene targets were:

Gene	Protein level	Statistical significance value	Gene	Protein level	Statistical significance value
CEMIP2	down	2.26E-01	TRIO	down	8.11E-23
AKAP11	down	4.15E-01	TSC22D2	down	1.05E-23
ATXN1	down	9.14E-01	USP9X	down	1.60E-10
PRDM2	down	5.99E-01	WDR26	down	5.47E-08
SPIN1	down	4.19E-01	YOD1	down	2.18E-06
FEM1B	down	2.61E-01	EIF5	down	1.06E-02
ATXN1L	down	5.12E-01	FOXO3	down	2.31E-04
DCP2	down	6.34E-01	UHMK1	-	-
PTPRJ	down	4.14E-01	KMT2A	Up	2.07E-04
SNTB2	down	2.72E-01	GSPT1	up	1.75E-08
ATP2B1	down	1.94E-25	GTF3C4	up	1.16E-11
CHORDC1	down	3.25E-08	HELZ	Up	4.59E-11
CNST	down	3.13E-07	HIPK1	Up	8.41E-04
GLS	down	7.94E-27	HNRNPU	Up	1.12E-10
MPRIP	down	6.47E-10	IREB2	up	5.97E-07
RAB11FIP2	down	3.08E-19	KDM5B	Up	2.07E-04
RAPGEF2	down	3.45E-46	KMT2C	Up	1.52E-08
SLC38A1	down	3.62E-01	LHFPL2	Up	1.01E-01
SSX2IP	down	1.33E-09	LPP	Up	4.25E-06
TTN	down	4.16E-10	RNF44	-	-
USP46	down	1.30E-28	LCOR	-	-
ARHGEF12	down	4.07E-13	ARHGAP11A	-	-
CREBRF	down	6.54E-03	ARRDC3	-	-
FBXO45	Down	2.37E-05	BCL2L11	-	-

GAN	down	1.82E-04	C5orf24	-	-
MET	down	6.20E-05	LEPROT	-	-
NF1	down	2.73E-25	MDM2	-	-
SORT1	down	1.26E-06	SCD	Up	2.75E-03
ARHGAP5	down	6.74E-11	CELF1	No change	4.86E-01
APLP2	down	1.58E-13	DAG1	No change	5.68E-01
BMPR2	down	5.41E-26	ITGB8	down	6.02E-01
CASK	down	2.46E-10	KBTBD2	Up	2.86E-02
CLTC	down	1.54E-33	MFHAS1	down	1.81E-01
CYLD	down	5.81E-45	NAP1L1	up	4.60E-02
DCAF7	down	2.68E-10	NSD1	Up	9.92E-02
FRS2	down	5.17E-05	SESN3	No change	9.92E-02
GRSF1	down	3.02E-03	FAM222B	No change	9.33E-01
MAPK1	down	1.02E-09	TULP4	No data	-
MAT2A	down	3.64E-06	SERTAD2	No data	-
NUCKS1	down	1.56E-07	SIX4	No data	-
PAK2	down	7.36E-03	SKIL	No data	-
PJA2	down	2.40E-04	SLC35F5	No data	-
PPP1CC	down	1.37E-20	TET2	No data	-
PTEN	Down	1.13E-15	TMED7	No data	-
PURB	down	4.42E-04	TNKS2	No data	-
RHOBTB3	down	3.26E-03	CREBL2	not identified	-
SSH1	down	2.02E-05	E2F7	not identified	-

STRN	down	6.23E-10	ENTPD7	not identified	-
TMEM245	down	1.05E-09	CCNT2	up	2.73E-01
			EFCAB14	up	4.83E-02

On studying the network and interaction of these gene targets using STRING software, we concluded that:

1. Molecular function

Term description	**Matching proteins in your network**
Lys63-specific de-ubiquitinase activity	USP9X, YOD1, CYLD
Protein kinase binding	PAK2, FOXO3, PPP1CC, PJA2, RAB11FIP2, PTEN, BMPR2, RHOBTB3, CYLD, FRS2, TTN, CLTC
Enzyme binding	MAPK1, SORT1, STRN, PAK2, MET, YOD1, FOXO3, PPP1CC, PJA2, CNST, RAB11FIP2, PTEN, BMPR2, RHOBTB3, CYLD, FRS2, TTN, CLTC
Purine ribonucleotide binding	MAPK1, EIF5, RAPGEF2, MAT2A, PAK2, MET, CHORDC1, TRIO, BMPR2, RHOBTB3, ARHGAP5, ATP2B1, TTN, CASK
Protein binding	MAPK1, EIF5, SORT1, STRN, RAPGEF2, MAT2A, PAK2, SSH1, USP9X, MET, CHORDC1, YOD1, FOXO3, SSX2IP, PPP1CC, PJA2, CNST, NUCKS1, RAB11FIP2, PTEN, BMPR2, RHOBTB3, ARHGAP5, PURB, MPRIP, ARHGEF12, CYLD, ATP2B1, FRS2, TTN, CLTC, CASK, APLP2

2. Wiki pathway

Term description	Matching proteins in your network
Hepatocyte growth factor receptor signaling	MAPK1, MET, PTEN
Splicing factor NOVA regulated synaptic proteins	ATP2B1, CASK, APLP2
Neural crest cell migration in cancer	SORT1, PAK2, TRIO
MET in type 1 papillary renal cell carcinoma	MAPK1, STRN, PAK2, MET
Glioblastoma signaling pathways	MAPK1, MET, FOXO3, NF1, PTEN
EGFR tyrosine kinase inhibitor resistance	MAPK1, MET, FOXO3, NF1, PTEN
Synaptic signaling pathways associated with autism spectrum disorder	MAPK1, NF1, PTEN
Endometrial cancer	MAPK1, FOXO3, PTEN
Melanoma	MAPK1, NF1, PTEN
Fragile X syndrome	MAPK1, NF1, PTEN, CLTC

3. KEGG pathway

Term description	Matching proteins in your network
EGFR tyrosine kinase inhibitor resistance	MAPK1, MET, FOXO3, NF1, PTEN
Central carbon metabolism in cancer	MAPK1, MET, GLS, PTEN
Endometrial cancer	MAPK1, FOXO3, PTEN
Renal cell carcinoma	MAPK1, PAK2, MET
Non-small cell lung cancer	MAPK1, MET, FOXO3
Adherens junction	MAPK1, MET, SSX2IP
Melanoma	MAPK1, MET, PTEN
Neurotrophin signaling pathway	MAPK1, SORT1, FOXO3, FRS2

Axon guidance	MAPK1, PAK2, SSH1, MET, BMPR2, ARHGEF12
MicroRNAs in cancer	MAPK1, MET, GLS, PTEN, BMPR2

4. Reactome pathway

Term description	Matching proteins in your network
Frs2-mediated activation	MAPK1, FRS2
ALK mutants bind TKIs	STRN, CLTC
RHOV GTPase cycle	PAK2, USP9X, CLTC
RHOU GTPase cycle	PAK2, USP9X, CLTC
Signaling by ALK fusions and activated point mutants	MAPK1, STRN, FRS2, CLTC
RHOJ GTPase cycle	PAK2, TRIO, ARHGAP5
Oncogenic MAPK signaling	MAPK1, PPP1CC, NF1, MPRIP
PI3K/AKT Signaling in Cancer	STRN, MET, FOXO3, PTEN, FRS2
Semaphorin interactions	PAK2, MET, ARHGEF12
Negative regulation of the PI3K/AKT network	MAPK1, STRN, MET, PTEN, FRS2

4.0 Discussion:

This study aimed to identify miRNAs that could serve as potential therapeutic targets for glioblastoma multiforme (GBM). To achieve this, a comprehensive analysis of miRNA targets and their functional relevance was conducted. A literature search was conducted and identified predominant miRNAs implicated in GBM such as miR-21-5p, miR-93-5p, miR-155-5p, miR-10b-5p, miR-221-3p, miR-222-3p, miR-182-5p, miR-196a-5p. Subsequently, the experimentally verified targets of these miRNAs were extracted from miRtarbase, tarbase whereas the predicted targets were extracted from miRdip database. Overlapping experimental and predicted targets were subjected to hallmark classification to assess their tumor relevance in GBM. A total of 195 gene targets were identified, encompassing various hallmarks of cancer as shown in **figure-12**. Of them, genome instability, sustained angiogenesis, tissue invasion and metastasis, resisting cell death, and replicative immortality were found to be statistically significant. To evaluate the impact of miRNA-mediated regulation, mRNA and protein expression levels were analysed. While 147 and 29 targets were upregulated and downregulated respectively at the mRNA level according to

GEPIA 2. Protein expression analysis using UALCAN showed 45 of them were highly downregulated according to their statistical values. These downregulated gene targets were further investigated to elucidate their biological functions and involvement in GBM pathogenesis using STRING analysis. Mainly the role of these proteins based on data from wiki pathways, KEGG pathways, reactome pathways were analysed and briefly discussed here. Wiki Pathways analysis revealed that these miRNA target proteins have potential role in in Hepatocyte growth factor receptor signalling, Neural crest cell migration in cancer, MET in type 1 papillary renal cell carcinoma, Glioblastoma signalling pathways, EGFR tyrosine kinase inhibitor resistance, Endometrial cancer, Melanoma. In KEGG pathways analysis EGFR tyrosine kinase inhibitor resistance, Central carbon metabolism in cancer, Endometrial cancer, Renal cell carcinoma, non-small cell lung cancer, Melanoma, Neurotrophin signalling pathway and MicroRNAs in cancer. In Reactome pathways analysis Signalling by ALK fusions and activated point mutants, Oncogenic MAPK signalling, PI3K/AKT Signalling in Cancer, Negative regulation of the PI3K/AKT network were identified. Together, the pathway analysis revealed the involvement of key proteins in cancer and GBM-related pathways. The present study revealed MAPK1, NF1, FOXO3, MET and PTEN as key hub players in glioblastoma that are regulated by these oncomeric miRNAs (**figure-13**). The outcome of this study may help in identifying miRNA panels for various therapeutic applications in GBM, in future.

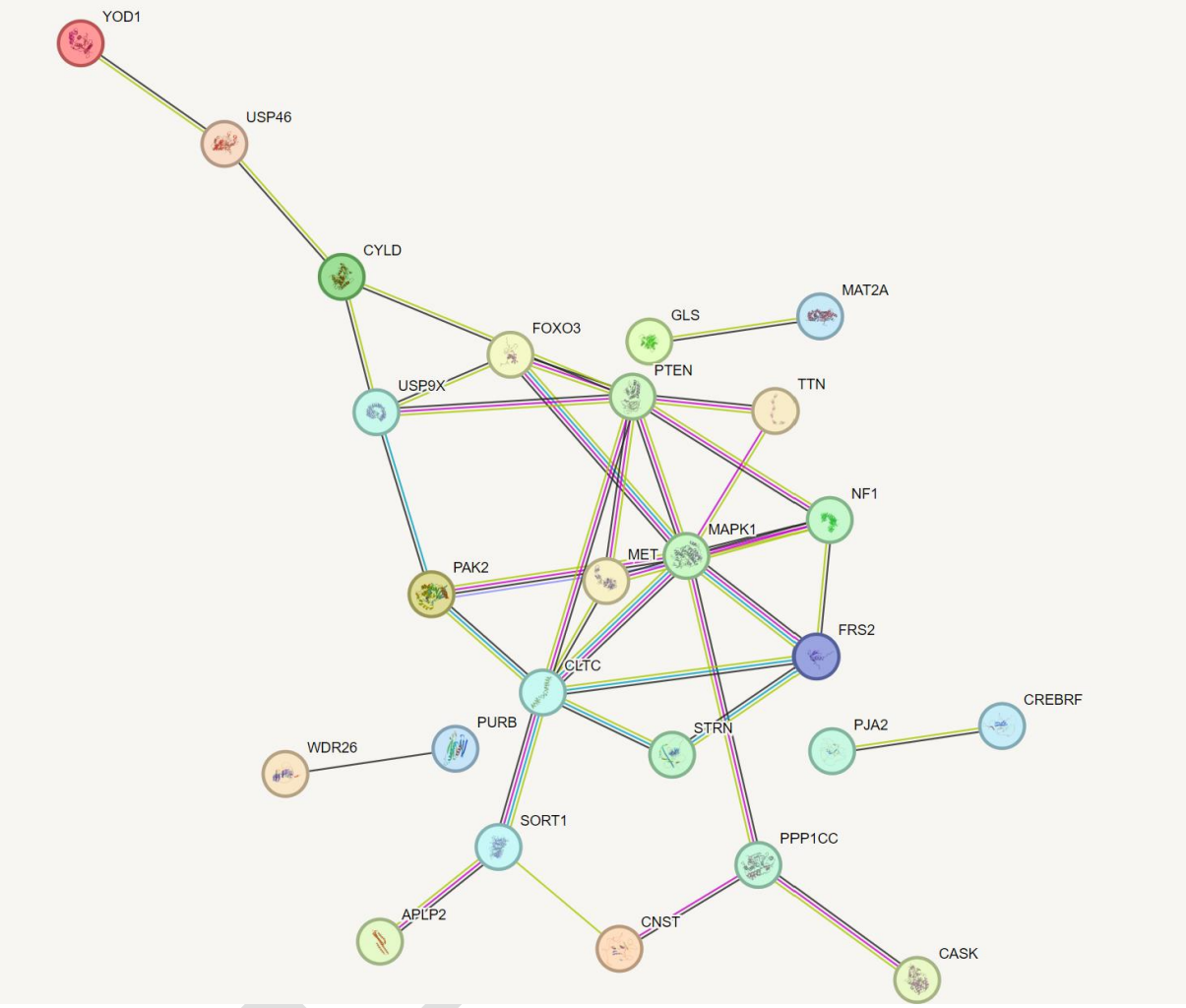


Figure-13: Gene target interactions network

5.0 References

1. O'Brien, J., Hayder, H., Zayed, Y., & Peng, C. (2018). Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation. *Frontiers in endocrinology*, *9*, 402. https://doi.org/10.3389/fendo.2018.00402
2. Jorge, A. L., Pereira, E. R., Oliveira, C. S., Ferreira, E. D. S., Menon, E. T. N., Diniz, S. N., & Pezuk, J. A. (2021). MicroRNAs: understanding their role in gene expression and cancer. *Einstein (Sao Paulo, Brazil)*, *19*, eRB5996. https://doi.org/10.31744/einstein_journal/2021RB5996
3. Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *Cell*, *144*(5), 646–674. https://doi.org/10.1016/j.cell.2011.02.013
4. Fouad, Y. A., & Aanei, C. (2017). Revisiting the hallmarks of cancer. *American journal of cancer research*, *7*(5), 1016–1036.
5. Gutschner, T., & Diederichs, S. (2012). The hallmarks of cancer: a long non-coding RNA point of view. *RNA biology*, *9*(6), 703–719. https://doi.org/10.4161/rna.20481
6. Carnero, A., & Paramio, J. M. (2014). The PTEN/PI3K/AKT Pathway in vivo, Cancer Mouse Models. *Frontiers in oncology*, *4*, 252. https://doi.org/10.3389/fonc.2014.00252
7. Riffo-Campos, Á. L., Riquelme, I., & Brebi-Mieville, P. (2016). Tools for Sequence-Based miRNA Target Prediction: What to Choose?. International journal of molecular sciences, 17(12), 1987. https://doi.org/10.3390/ijms17121987
8. 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
9. Peterson, S. M., Thompson, J. A., Ufkin, M. L., Sathyanarayana, P., Liaw, L., & Congdon, C. B. (2014). Common features of microRNA target prediction tools. Frontiers in genetics, 5, 23. https://doi.org/10.3389/fgene.2014.00023
10. Chen, L., Heikkinen, L., Wang, C., Yang, Y., Sun, H., & Wong, G. (2019). Trends in the development of miRNA bioinformatics tools. Briefings in bioinformatics, 20(5), 1836–1852. https://doi.org/10.1093/bib/bby054

11. Skoufos, G., Kakoulidis, P., Tastsoglou, S., Zacharopoulou, E., Kotsira, V., Miliotis, M., Mavromati, G., Grigoriadis, D., Zioga, M., Velli, A., Koutou, I., Karagkouni, D., Stavropoulos, S., Kardaras, F. S., Lifousi, A., Vavalou, E., Ovsepian, A., Skoulakis, A., Tasoulis, S. K., Georgakopoulos, S. V., … Hatzigeorgiou, A. G. (2024). TarBase-v9.0 extends experimentally supported miRNA-gene interactions to cell-types and virally encoded miRNAs. Nucleic acids research, 52(D1), D304–D310. https://doi.org/10.1093/nar/gkad1071
12. Huang, H. Y., Lin, Y. C., Cui, S., Huang, Y., Tang, Y., Xu, J., Bao, J., Li, Y., Wen, J., Zuo, H., Wang, W., Li, J., Ni, J., Ruan, Y., Li, L., Chen, Y., Xie, Y., Zhu, Z., Cai, X., Chen, X., … Huang, H. D. (2022). miRTarBase update 2022: an informative resource for experimentally validated miRNA-target interactions. Nucleic acids research, 50(D1), D222–D230. https://doi.org/10.1093/nar/gkab1079
13. Hauschild, A. C., Pastrello, C., Ekaputeri, G. K. A., Bethune-Waddell, D., Abovsky, M., Ahmed, Z., Kotlyar, M., Lu, R., & Jurisica, I. (2023). MirDIP 5.2: tissue context annotation and novel microRNA curation. Nucleic acids research, 51(D1), D217–D225. https://doi.org/10.1093/nar/gkac1070
14. Tang, Z., Kang, B., Li, C., Chen, T., & Zhang, Z. (2019). GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic acids research*, *47*(W1), W556–W560. https://doi.org/10.1093/nar/gkz430
15. Ru, B., Wong, C. N., Tong, Y., Zhong, J. Y., Zhong, S. S. W., Wu, W. C., Chu, K. C., Wong, C. Y., Lau, C. Y., Chen, I., Chan, N. W., & Zhang, J. (2019). TISIDB: an integrated repository portal for tumor-immune system interactions. *Bioinformatics (Oxford, England)*, *35*(20), 4200–4202. https://doi.org/10.1093/bioinformatics/btz210